Multimodal Architectures in KDD

Akash Choudhuri

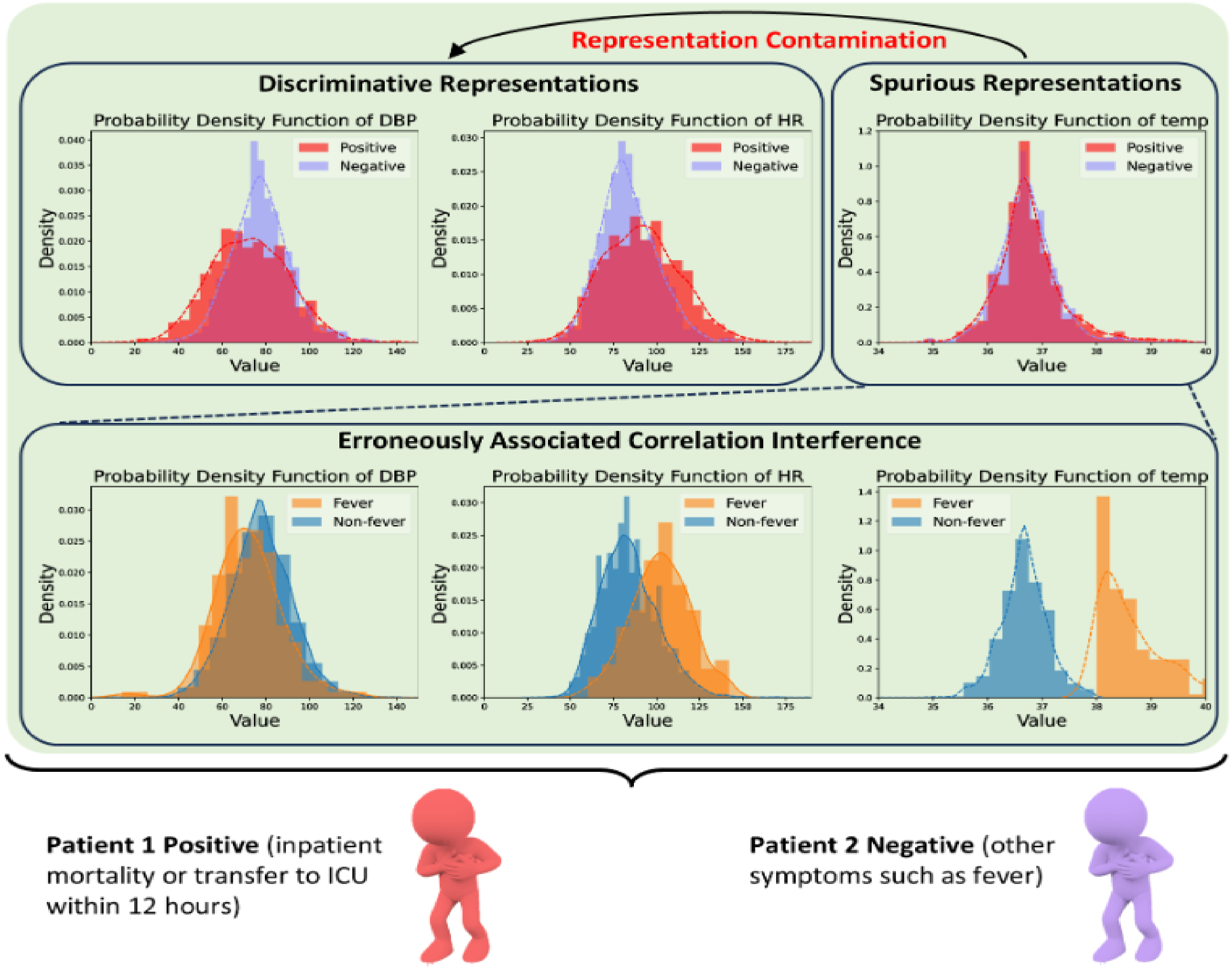
CLEAR: Addressing Representation Contamination in Multimodal Healthcare Analytics

Ge Su, Kaiping Zheng, Tiancheng Zhao, Jianwei Yin

KDD 2025

Motivation

- Prior works overlook the issue of **latent contamination**
- Example:
  - Patient 1:
    - Positive for (12 hrs):
      - In-Hospital Mortality
      - ICU Transfer
    - Shows:
      - Low diastolic blood pressure
      - High Heart Rate
  - Patient 2:
    - Negative
    - Shows:
      - Abnormal Temperatures



Elevated temp can lead to low blood pressure and an increased HR

Objectives

- Model representation contamination (Counterfactual Prompt Learning)
- Perform representation calibration (Adaptive Dynamic Imputation)
- Enhancement for constructing highly discriminative representations (Multimodal Representation Fusion)

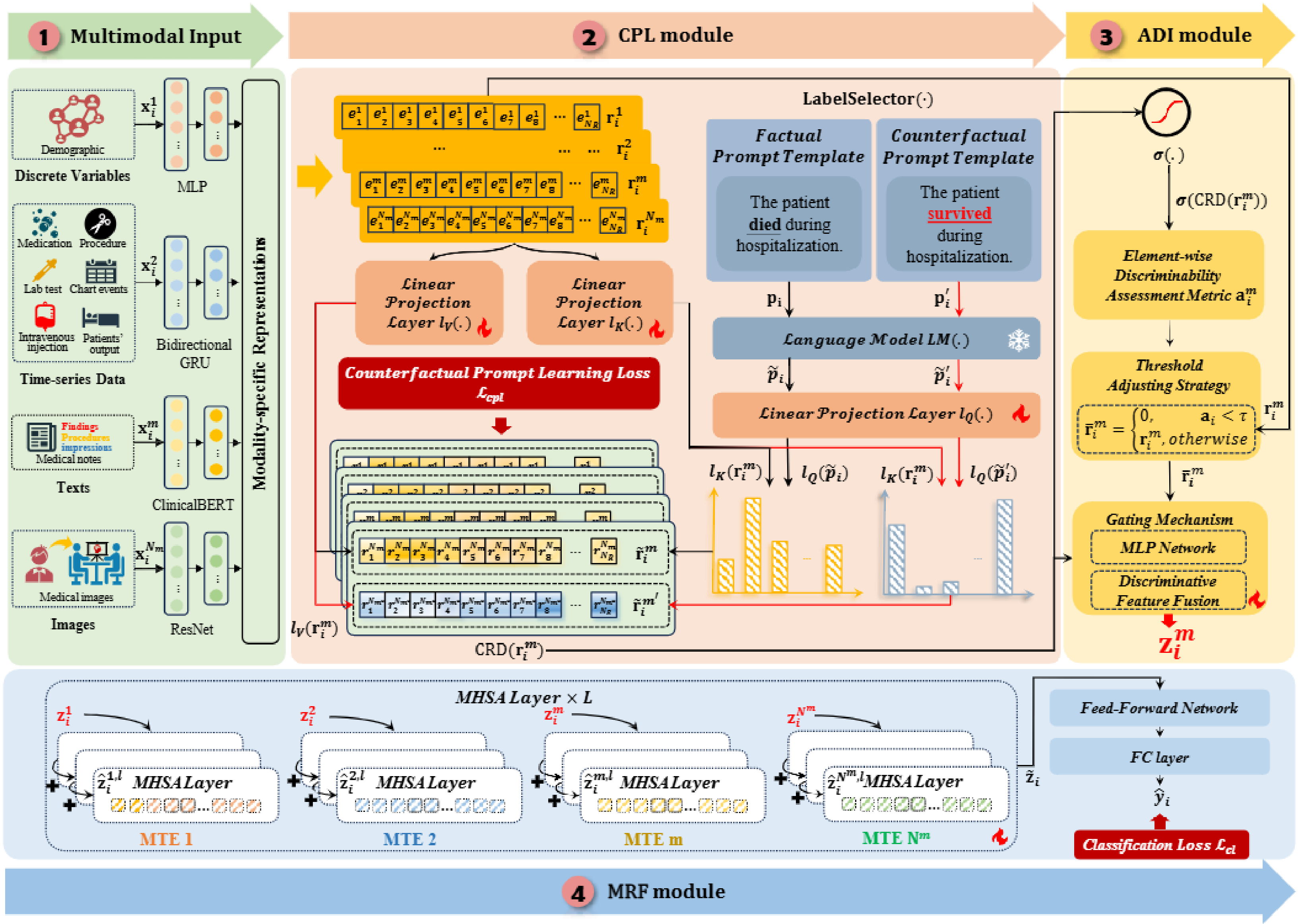
CLEAR: **C**ounterfactual disparity **L**earning model for **E**xplicit Multimod**a**l EH**R** analysis

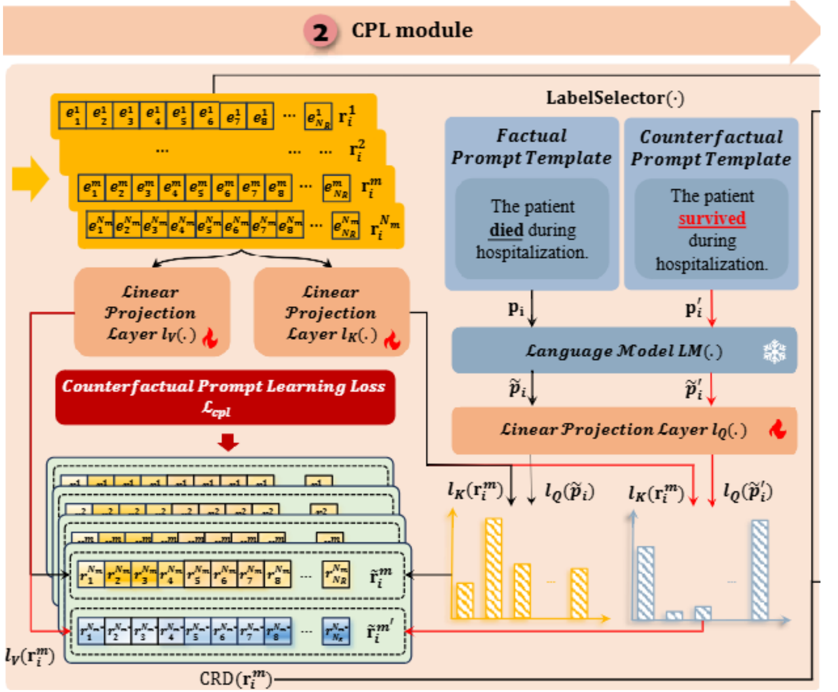
Problem Formulation

- **Multimodal Data:** $\mathcal{D} = \{\mathbf{x}_i, \mathbf{y}_i\}_{i=1}^{N_p}$ $\quad \mathbf{x}_i = \{\mathbf{x}_i^1, \mathbf{x}_i^2, ..., \mathbf{x}_i^{N_m}\}$ $\quad \mathbf{y}_i \in \{0, 1\}^{|N_c|}$
- **Objective:** Utilize multimodal data along with their diagnostic labels to train a general multimodal model CLEAR for supporting effective medical decision-making
- Training Pipeline:

$$\underbrace{\mathcal{M}((\mathrm{X}, \mathrm{Y}); (\Theta_C, \Theta_{\mathcal{A}}, \Theta_{\mathcal{M}}))}_{\text{CLEAR}} = \underbrace{\mathcal{M}_C(\tilde{\mathrm{R}}, \tilde{\mathrm{R}}' | (\mathrm{X}, \mathrm{P}, \mathrm{P}'); \Theta_C)}_{\text{CPL}}$$
$$\rightarrow \underbrace{\mathcal{M}_{\mathcal{A}}(\mathrm{Z} | (\tilde{\mathrm{R}}, \tilde{\mathrm{R}}'); \Theta_{\mathcal{A}})}_{\text{ADI}} \rightarrow \underbrace{\mathcal{M}_{\mathcal{M}}((\mathrm{Y}, \hat{\mathrm{Y}}) | \mathrm{Z}; \Theta_{\mathcal{M}})}_{\text{MRF}}$$

Architecture





Counterfactual Prompt Learning

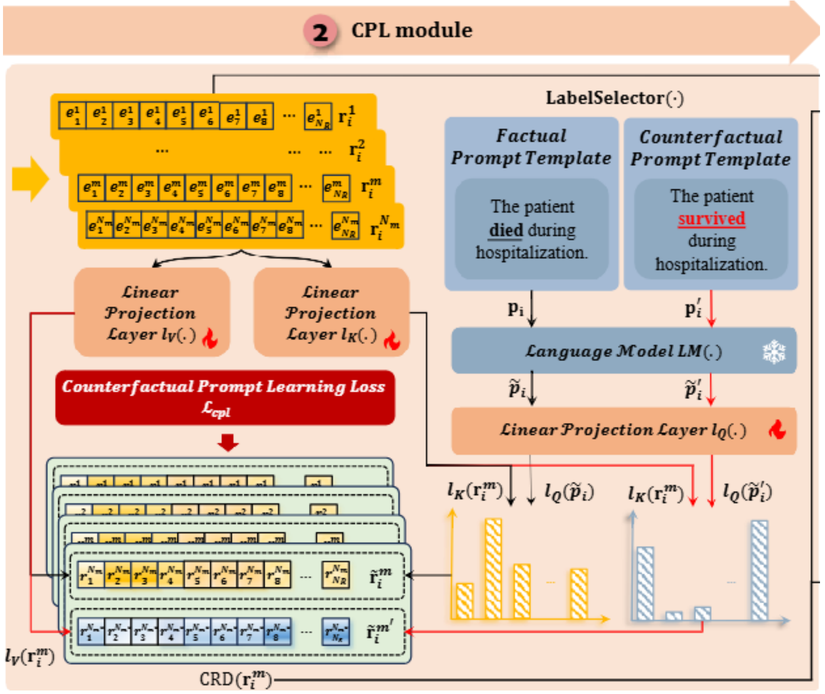
- Modality-specific encoder: $\mathrm{r}_i^m = f_m(\mathrm{x}_i^m; \Theta_m)$
- Objective: Harness the counterfactual representation discrepancy projected by factual and counterfactual prompts to distinguish between discriminative and spurious representations
- Construct factual and counterfactual prompts:

$$\mathrm{p}_i = \mathrm{LabelSelector}(\mathrm{PTs}, \mathrm{y}_i)$$
$$\mathrm{p}'_i = \mathrm{LabelSelector}(\mathrm{PTs}, \neg\mathrm{y}_i)$$

- Map the prompts into latent space:

$$\tilde{\mathrm{p}}_i = \mathrm{ClinicalBERT}(\mathrm{p}_i)$$
$$\tilde{\mathrm{p}}'_i = \mathrm{ClinicalBERT}(\mathrm{p}'_i)$$

Counterfactual Prompt Learning



- Project correlations between modality representations and both prompt representations
- MHA:

$$\tilde{\mathbf{r}}_i^m = \text{softmax}\left(\frac{l_Q(\tilde{\mathbf{p}}_i)\, l_K(\mathbf{r}_i^m)^\top}{\sqrt{d_k}}\right) l_V(\mathbf{r}_i^m)$$

$$\tilde{\mathbf{r}}_i^{m\prime} = \text{softmax}\left(\frac{l_Q(\tilde{\mathbf{p}}_i^\prime)\, l_K(\mathbf{r}_i^m)^\top}{\sqrt{d_k}}\right) l_V(\mathbf{r}_i^m)$$

- Counterfactual representation discrepancy:

$$\text{CRD}(\mathbf{r}_i^m) = \tilde{\mathbf{r}}_i^m \ominus \tilde{\mathbf{r}}_i^{m\prime}$$

- Minimize the overall loss:

$$\mathcal{L}_{cpl} = \frac{1}{N_B}\sum_{i=1}^{N_B}\left(\sum_{m=1}^{N_m} ||\tilde{\mathbf{r}}_i^m - \tilde{\mathbf{r}}_i^{m\prime}||_F + \mathcal{L}_{ce}(l_{dc}(\text{Concat}(\text{CRD}(\mathbf{r}_i^m))), \mathbf{y}_i)\right)$$

Counterfactual Representation Discrepancy: It is the element-wise difference between 2 representations projected by the factual prompt P and projected by the counterfactual prompt P':

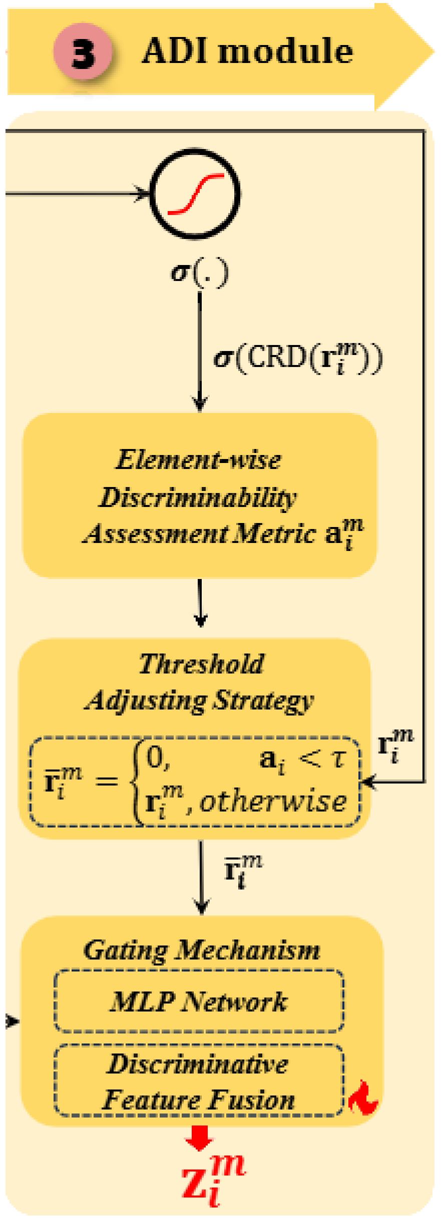
$$\text{CRD}(\text{R}) = \underbrace{\text{Enc}(\tilde{\text{R}}|(\text{R},\text{P});\Theta_{\text{Enc}})}_{\text{Representations projected by }\mathbf{P}} - \underbrace{\text{Enc}(\tilde{\text{R}}^\prime|(\text{R},\text{P}^\prime);\Theta_{\text{Enc}})}_{\text{Representations projected by }\mathbf{P}^\prime}$$

Adaptive Dynamic Imputation

- Performs representation calibration and enhancement element-wise
- Map to binary value: $\mathbf{a}_i^m = \sigma(\mathrm{CRD}(\mathbf{r}_i^m))$
- Threshold with representations with high discriminative discrepancies are retained, and those with low discrepancies are suppressed: $$\bar{\mathbf{r}}_i^m = \begin{cases} 0, & \mathbf{a}_i^m < \tau \\ \mathbf{r}_i^m, & otherwise \end{cases}$$
- Gating-Merge:

$$\mathbf{z}_i^m = \mathbf{g} \odot \bar{\mathbf{r}}_i^m + (1 - \mathbf{g}) \odot \mathrm{CRD}(\mathbf{r}_i^m)$$

$$\mathbf{g} = \mathrm{MLP}(\mathrm{Concat}(\bar{\mathbf{r}}_i^m, \mathrm{CRD}(\mathbf{r}_i^m)))$$



Multimodal Representation Fusion

- Establish both intra-modality correlations and inter-modality associations
- Assuming $\hat{z}_m^i$ is interaction within a modality, intra-modality:

$$\hat{z}_i^{m,l} = \mathrm{LayerNorm}(\hat{z}_i^{m,l-1} + \mathrm{MHSA}_l^{intra}(\hat{z}_i^{m,l-1})), 1 \le l \le L$$

- Concatenate all modality information: $\hat{z}_i^w = [\hat{z}_i^1; \hat{z}_i^2; \dots; \hat{z}_i^{N_m}]$
- For inter-modality:

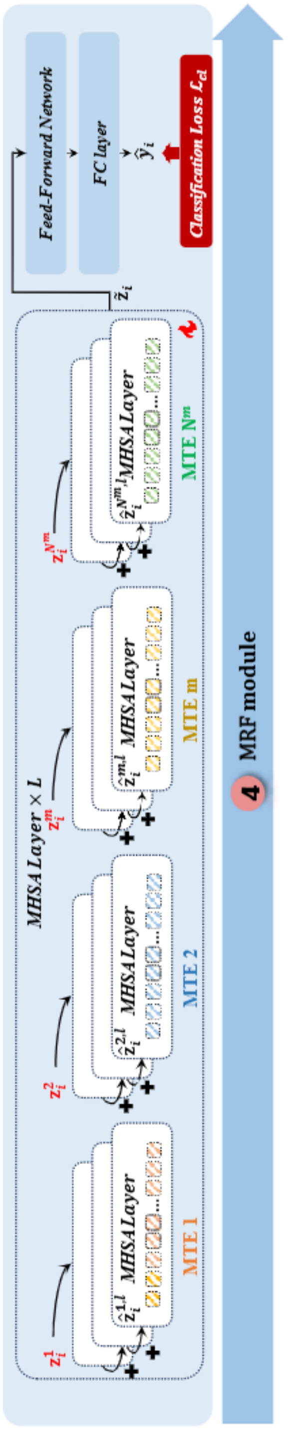
$$\tilde{z}_i^{w,l} = \mathrm{LayerNorm}(\tilde{z}_i^{w,l-1} + \mathrm{MHSA}_l^{inter}(\tilde{z}_i^{w,l-1})), 0 \le l \le L-1 \quad (19)$$

$$z_i^{w,l} = \mathrm{LayerNorm}(\tilde{z}_i^{w,l} + \mathrm{FFN}(\tilde{z}_i^{w,l})) \quad (20)$$

- Final prediction and loss: $\hat{y}_i = \sigma(z_i^{w,L-1} W_{final})$

$$L_{cl} = -\frac{1}{N_B}\sum_{i=1}^{N_B}(y_i \log \hat{y}_i + (1-y_i)\log(1-\hat{y}_i))$$

$$\mathcal{L}_{COOP} = \lambda \mathcal{L}_{cpl} + (1-\lambda)\mathcal{L}_{cl}$$



Experiments

- Data:
  - MIMIC-IV
  - ODIR
- Tasks:
  - Critical Outcome Prediction and Hospitalization Prediction (MIMIC-IV)
  - Ocular Phenotype Prediction (ODIR)

Dataset	MIMIC-IV	ODIR
# samples	418,100	5,000
# modalities	5	3
# modalities features	[65, 7, 704, 10320, 768]	[2, 768*2, 1000*2]
# tasks	2	1
% positive labels	[47.76, 6.63]	32.48
% train:val:test	8:1:1	7:1:2

Table 3: Experimental results of critical outcome prediction on the MIMIC-IV dataset.

Critical Outcome Prediction (Binary Classification)				
Method	AUPRC	F1	Precision	Recall
MLP [56]	0.3931 (0.0017)	0.2622 (0.0238)	0.6100 (0.0095)	0.1674 (0.0193)
M3Care [63]	0.6702 (0.0142)	0.5446 (0.0119)	**0.7865** (0.0207)	0.4165 (0.0111)
Dyhealth [66]	0.6310 (0.0030)	0.5269 (0.0160)	0.7499 (0.0214)	0.4068 (0.0250)
UTDE [64]	0.7423 (0.0056)	0.6484 (0.0222)	0.7460 (0.0409)	0.5733 (0.0517)
HetMed [26]	0.6934 (0.0307)	0.6215 (0.0207)	0.7327 (0.0569)	0.5425 (0.0408)
CLEAR (Ours)	**0.7844** (0.0008)	**0.6902** (0.0056)	0.7399 (0.0103)	**0.6467** (0.0126)

Hospitalization Prediction (Binary Classification)				
Method	AUPRC	F1	Precision	Recall
MLP [56]	0.7946 (0.0015)	0.7284 (0.0222)	0.7198 (0.0316)	0.7426 (0.0752)
M3Care [63]	0.9858 (0.0007)	0.9292 (0.0018)	**0.9607** (0.0023)	0.8998 (0.0024)
Dyhealth [66]	0.9746 (0.0024)	0.8975 (0.0037)	0.9375 (0.0065)	0.8607 (0.0075)
UTDE [64]	0.9852 (0.0053)	0.9278 (0.0164)	0.9385 (0.0146)	0.9175 (0.0253)
HetMed [26]	0.9633 (0.0006)	0.9071 (0.0028)	0.9573 (0.0068)	0.8620 (0.0039)
CLEAR (Ours)	**0.9940** (0.0002)	**0.9579** (0.0009)	0.9570 (0.0077)	**0.9588** (0.0084)

Ocular phenotype Prediction (Multi-label Classification)										
Method	N	D	G	C	A	H	M	O	macro-AUC	micro-AUC
MLP [56]	0.8774 (0.0086)	0.6963 (0.0477)	0.5989 (0.0340)	0.8595 (0.0369)	0.6846 (0.0493)	0.4990 (0.0010)	0.8563 (0.0197)	0.6996 (0.0201)	0.7214 (0.0024)	0.7664 (0.0171)
M3Care [63]	0.9868 (0.0103)	0.9981 (0.0003)	0.8673 (0.0779)	0.8258 (0.0556)	0.8872 (0.0208)	0.7783 (0.0804)	0.9620 (0.0407)	0.9022 (0.0233)	0.9010 (0.0038)	0.9599 (0.0042)
Dyhealth [66]	0.9547 (0.0225)	0.8854 (0.0121)	0.8955 (0.0330)	0.9331 (0.0517)	0.9383 (0.0600)	0.8328 (0.0520)	0.9224 (0.0885)	0.7782 (0.0836)	0.8926 (0.0188)	0.9285 (0.0101)
HetMed [26]	0.7526 (0.0111)	0.7545 (0.0073)	0.7865 (0.0069)	0.9885 (0.0010)	0.8117 (0.0114)	0.7027 (0.0028)	0.9889 (0.0003)	0.5204 (0.0355)	0.7882 (0.0093)	0.8641 (0.0028)
CLEAR (Ours)	**0.9998** (0.0001)	**0.9991** (0.0004)	**0.9933** (0.0024)	**0.9999** (0.0002)	**0.9996** (0.0005)	**0.9912** (0.0014)	**0.9948** (0.0038)	**0.9977** (0.0013)	**0.9969** (0.0008)	**0.9986** (0.0006)

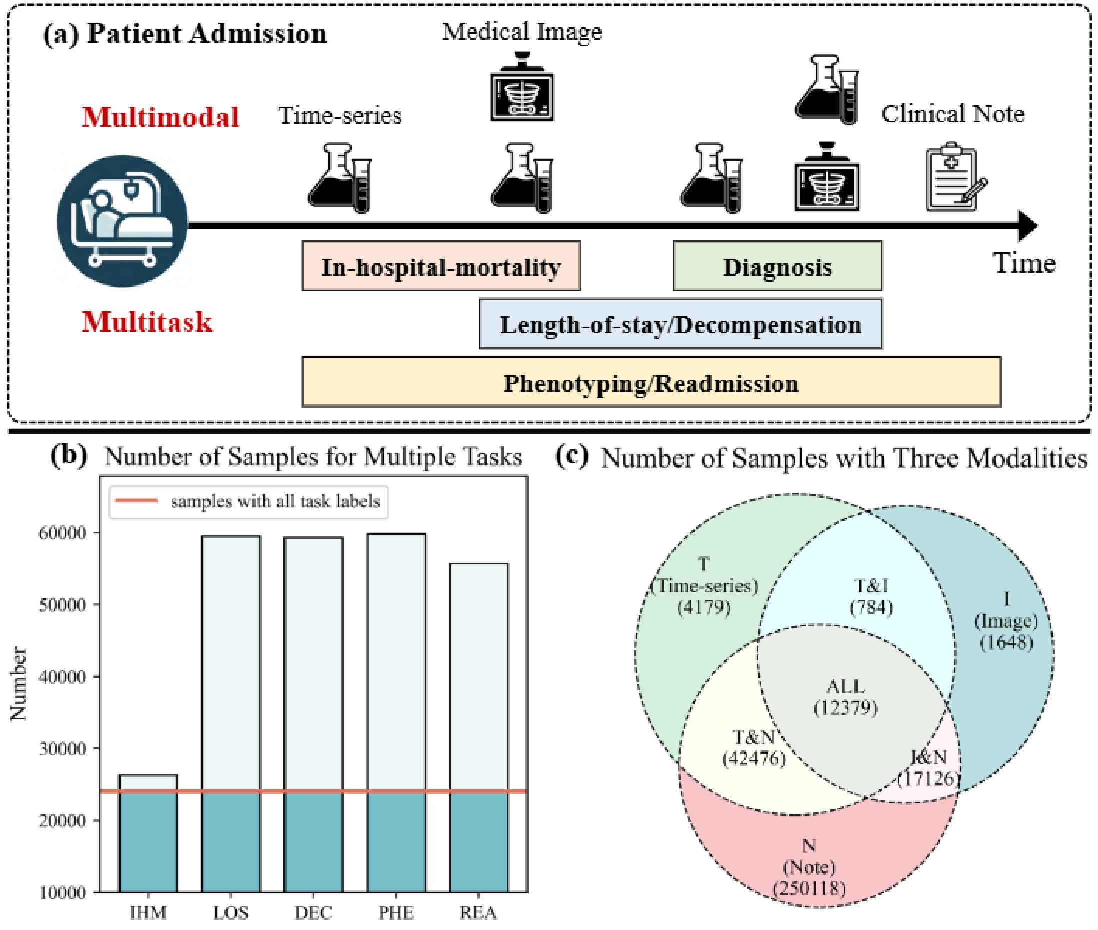
FlexCare: Leveraging Cross-Task Synergy for Flexible Multimodal Healthcare Prediction

Muhao Xu, Zhenfeng Zhu, Youru Li, Shuai Zheng, Yawei Zhao, Kunlun He

KDD 2024

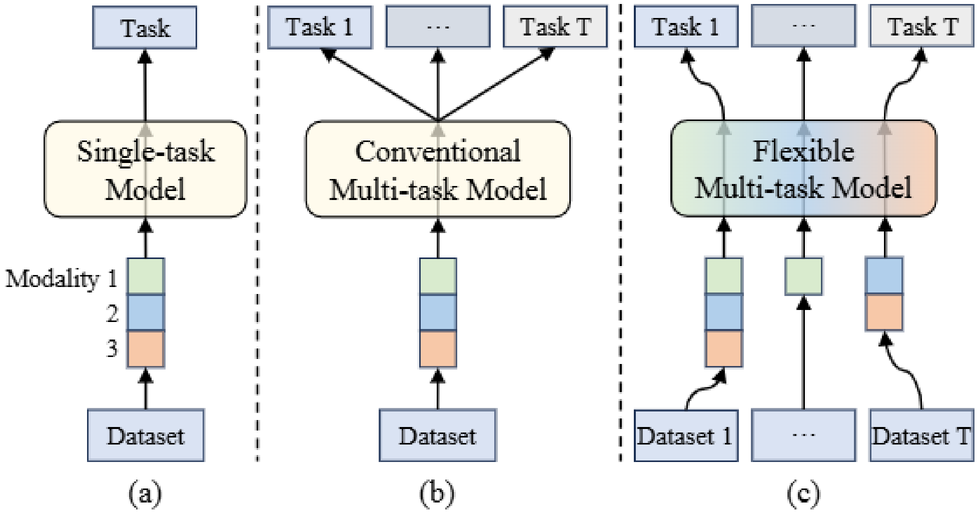
Motivation

- Current multitask models for healthcare prediction typically necessitate complete labels for all tasks
- Such demand for data is exceedingly stringent



Objectives

- Develop a flexible model capable of supporting multimodal inputs and adapting to various heterogeneous tasks, without requiring comprehensive labels for each sample across all tasks
- Deal with the information disparities among modalities and tasks comprehensively within a multitask framework



FlexCare: Flexible Multimodal Health**care** Prediction

Problem Formulation

- Datasets for a set of tasks: $\{\mathcal{D}_\tau\}_{\tau=1}^{T}$

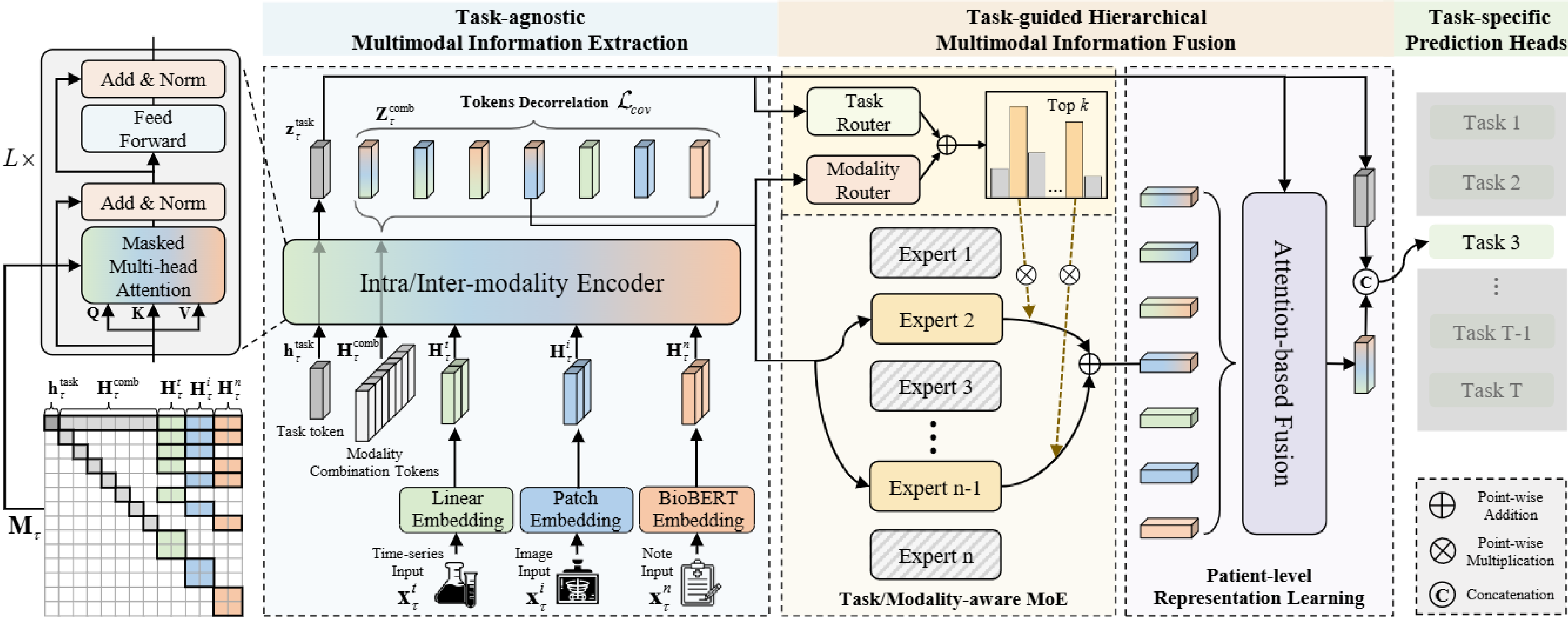
$\mathcal{D}_\tau = \left\{(\mathrm{X}_\tau^{(n)}, y_\tau^{(n)})\right\}_{n=1}^{N_\tau}$, where N_τ is the number of samples, $\mathrm{X}_\tau^{(n)}$ and $y_\tau^{(n)} \in \mathcal{Y}_\tau$ are the multimodal input and ground truth of the n-th sample, respectively. $\mathcal{Y}_\tau$ is the set of labels for the τ-th task.

Definition 2 (Patient multimodal data). Given $\mathcal{M} = \{t, i, n\}$ a set of modalities (i.e., time-series data, image, note), the input of n-th sample can be defined as: $\mathrm{X}_\tau^{(n)} = \{\mathrm{X}_\tau^{(n),m}\}_{m\in\mathcal{M}}$. Considering the absence of some modalities, the incomplete input is $\mathrm{X}_\tau^{(n)} = \{\mathrm{X}_\tau^{(n),m}\}_{m\in\mathcal{M}^{(n)}}$, where $\mathcal{M}^{(n)}$ are the modalities actually present in the n-th sample, and $\mathcal{M}^{(n)} \subseteq \mathcal{M}$. Note that $|\mathcal{M}^{(n)}| \geq 1$, because at least one modality is present for each sample.

Definition 3 (Modality combination). The modality combination set represents all patterns of unimodal or multimodal combination, defined as: $C = 2^{\mathcal{M}} \backslash \emptyset$ (i.e., all nonempty subsets of $\mathcal{M}$). When $|\mathcal{M}| = 3$, the number of modality combination set $|C| = 7$.

- Overall Objective: Given the multimodal datasets for different tasks, learn a unified task-adaptive predictive function $\hat{y}_\tau^{(n)} = f_\theta(\mathrm{X}_\tau^{(n)}, \tau)$

Architecture

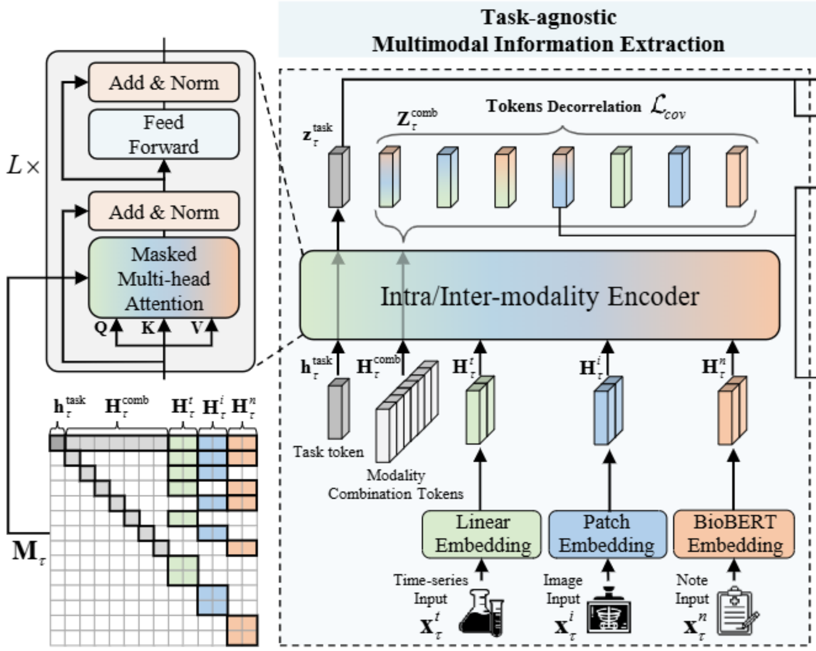


3 parts:

- **Task-agnostic multimodal information extraction:** Leverages unimodal feature extractors and a multimodal encoder
- **Task-guided hierarchical multimodal fusion:** Hierarchical fusion from modality-level to patient-level embeddings
- **Task-specific prediction heads:** Configured with individual predictors for each task

Task-agnostic Multimodal Information Extraction

- Encode to embedding: $\mathbf{H}_\tau^m = f_m(\mathbf{X}_\tau^m) + \mathbf{p}^m$
- Inject task-specific information into the model, a learnable task token is allocated for each task category $\mathbf{h}_\tau^{\text{task}} \in \mathbb{R}^d$
- Stack inter-modality token: $\mathbf{H}_\tau^{\text{comb}} = \{\mathbf{h}_\tau^c\}_{c \in C}$
- Stack everything together to get a multimodal sequence: $\mathbf{H}_\tau^0 = [\mathbf{h}_\tau^{\text{task}}, \mathbf{H}_\tau^{\text{comb}}, \mathbf{H}_\tau^t, \mathbf{H}_\tau^i, \mathbf{H}_\tau^n]$
- Get correlation and multimodal fusion:

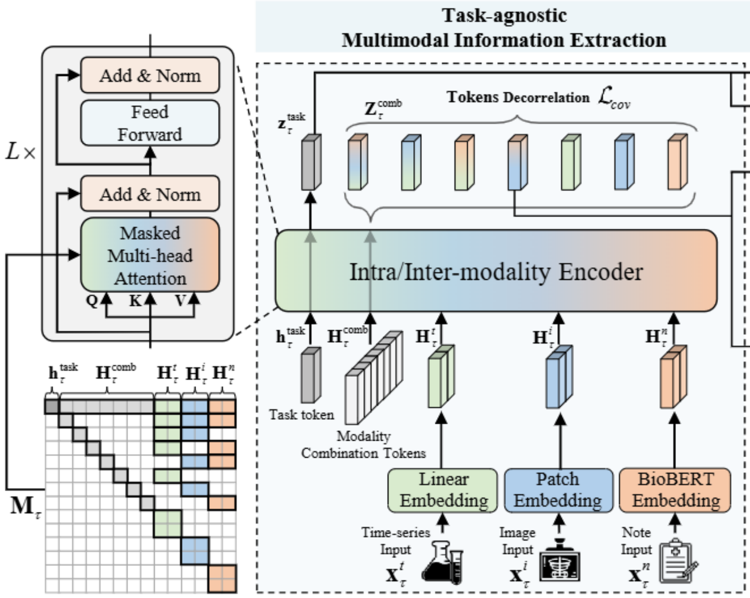


$$\widetilde{\mathbf{H}}_\tau^l = \mathrm{LN}(\mathbf{H}_\tau^{l-1} + \text{M-MHSA}(\mathbf{H}_\tau^{l-1}, \mathbf{H}_\tau^{l-1}, \mathbf{H}_\tau^{l-1}, \mathbf{M}_\tau)),$$
$$\mathbf{H}_\tau^l = \mathrm{LN}(\widetilde{\mathbf{H}}_\tau^l + \mathrm{FFN}(\widetilde{\mathbf{H}}_\tau^l)),$$

$$\text{M-MHSA}(\mathbf{Q}, \mathbf{K}, \mathbf{V}, \mathbf{M}) = \mathrm{Softmax}(\frac{\mathbf{Q}\mathbf{W}^Q(\mathbf{K}\mathbf{W}^K)^\top}{\sqrt{d}} + \mathbf{M})\mathbf{V}\mathbf{W}^V$$

Task-agnostic Multimodal Information Extraction

$$\widetilde{\mathrm{H}}_\tau^l = \mathrm{LN}(\mathrm{H}_\tau^{l-1} + \text{M-MHSA}(\mathrm{H}_\tau^{l-1}, \mathrm{H}_\tau^{l-1}, \mathrm{H}_\tau^{l-1}, \mathbf{M}_\tau)),$$
$$\mathrm{H}_\tau^l = \mathrm{LN}(\widetilde{\mathrm{H}}_\tau^l + \mathrm{FFN}(\widetilde{\mathrm{H}}_\tau^l)),$$



- **Additional modality mask:** Enables modality combination tokens to precisely target information relevant to diverse modality combination patterns:

$$\mathbf{M}_\tau(i,j) = \begin{cases} 0, & \phi(j) \in \phi(i) \text{ or } \phi(j) = \phi(i) \\ -\infty, & \text{otherwise} \end{cases},$$

where ϕ : index $\mapsto (\mathcal{M}|\mathcal{C})$ defines a function that maps the token to the modality m or modality combination c it belongs to. $\phi(j) \in \phi(i)$ indicates that the token $\mathrm{H}_{\tau,j}$ originates from modality m, the token $\mathrm{H}_{\tau,i} = \mathrm{h}_\tau^c$ from modality combination c, and $m \in c$. $\phi(j) = \phi(i)$ indicates that the token $\mathrm{H}_{\tau,j}$ and token $\mathrm{H}_{\tau,i}$ is from the same modality m or modality combination c.

- As multimodal information extraction is task-agnostic, task token aggregates info from other tokens unidirectionally $\mathrm{z}_\tau^{\text{task}} = \mathrm{h}_\tau^{\text{task},L}$ $\quad \mathrm{Z}_\tau^{\text{comb}} = \mathrm{H}_\tau^{\text{comb},L}$

- To avoid feature redundancy, regularizer:

$$\mathcal{L}_{cov} = \frac{1}{N_\tau}\sum_{n=1}^{N_\tau} C_\tau^{(n)} \qquad C_\tau = \frac{1}{(|C|-1)^2}\sum_{i\neq j}[\mathrm{Cov}(\mathrm{Z}_\tau^{\text{comb}})]_{i,j}^2. \quad \mathrm{Cov}(\mathrm{Z}) = \frac{1}{d-1}\sum_{j=1}^{d}(\mathrm{z}_{:,j} - \bar{\mathrm{z}})(\mathrm{z}_{:,j} - \bar{\mathrm{z}})^{\mathsf{T}}, \text{where} \quad \bar{\mathrm{z}} = \sum_{j=1}^{d} \mathrm{z}_{:,j},$$

Task-guided Hierarchical Multimodal Fusion & Prediction Heads

- Task/Modality-aware MOE: Weighted average of the selected k experts:

$$R(\mathbf{z}_\tau^c, \mathbf{z}_\tau^{\text{task}}) = \text{Softmax}(\text{TopK}(\mathbf{z}_\tau^c \mathbf{W}_1^R + \mathbf{z}_\tau^{\text{task}} \mathbf{W}_2^R, k)),$$

$$\mathbf{s}_\tau^c = \sum_{i=1}^{N^e} R(\mathbf{z}_\tau^c, \mathbf{z}_\tau^{\text{task}})_i E_i(\mathbf{z}_\tau^c),$$

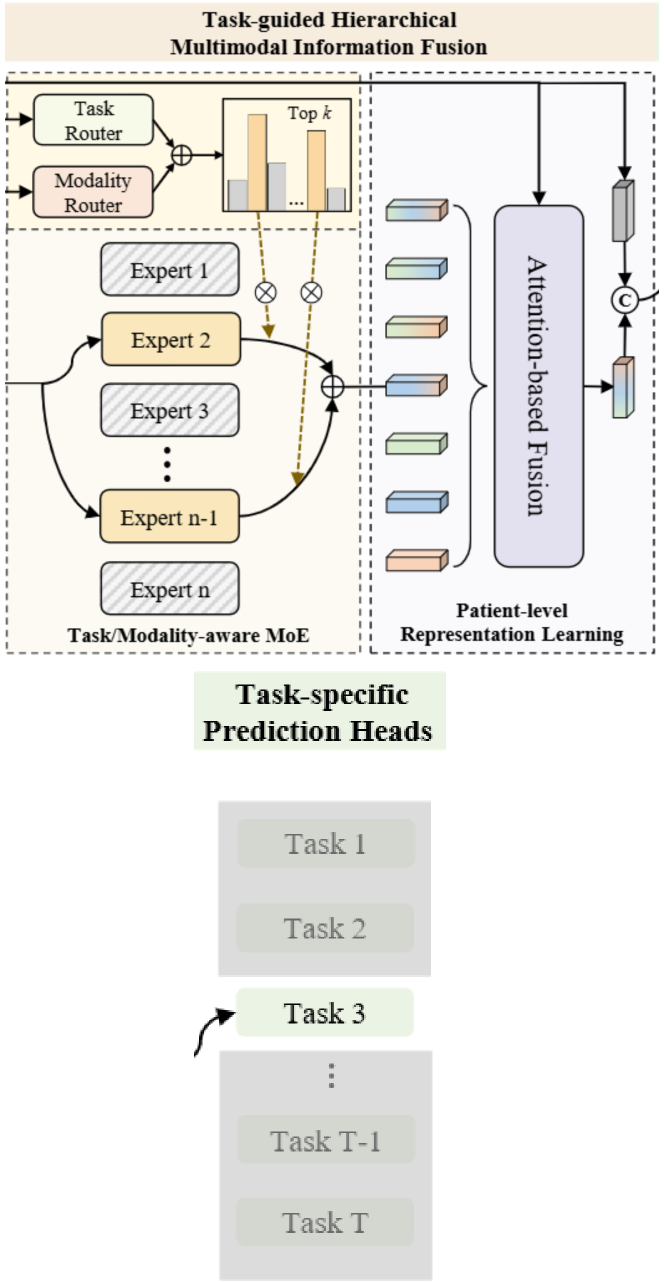
$$\text{TopK}(\mathbf{v}, k)_i = \begin{cases} \mathbf{v}_i, & \text{if } \mathbf{v}_i \text{ is within the top-k elements of } \mathbf{v} \\ -\infty, & \text{otherwise} \end{cases},$$

- Patient-level embeddings:

$$\mathbf{s}_\tau^p = [\mathbf{z}_\tau^{\text{task}} || \text{LN}(\sum_{c \in C} \alpha^c \mathbf{s}_\tau^c)], \quad \alpha^c = \frac{\exp(\widetilde{\alpha}^c/\varepsilon)}{\sum_{c \in C} \exp(\widetilde{\alpha}^c/\varepsilon)},$$

$$\widetilde{\alpha}^c = \text{Tanh}([\mathbf{z}_\tau^{\text{task}} || \mathbf{s}_\tau^c] \mathbf{W}_1^A) \cdot \mathbf{W}_2^A,$$

- Final Prediction: $\hat{y}_\tau = P_\tau(\mathbf{s}_\tau^p)$
- Overall Loss: $\mathcal{L}_\tau = \mathcal{L}_\tau^{pred} + \beta \mathcal{L}_{cov}, \quad \mathcal{L}_\tau^{pred} = \frac{1}{N_\tau} \sum_{n=1}^{N_\tau} \ell_\tau(y_\tau^{(n)}, \hat{y}_\tau^{(n)})$



Experiments

- Data: MIMIC-IV, MIMIC-CXR JPG, MIMIC-IV NOTE
- Tasks:
  - in-hospital mortality (IHM)
  - length-of-stay (LOS)
  - decompensation(DEC)
  - phenotyping (PHE)
  - readmission (REA)
  - diagnosis (DIA)

Task	# Number	Missing rate per modality		
		Time-series	Image	Note
IHM	26,318	0%	76.40%	7.49%
LOS	59,495	0%	85.16%	8.27%
DEC	59,269	0%	85.15%	8.24%
PHE	59,798	0%	81.94%	8.30%
REA	55,712	0%	82.38%	8.06%
DIA	132,576	76.34%	0%	32.56%

Result

Task	Metric	MedFuse [6]	MT [22]	M3Care [37]	MMF [15]	MultiModN [30]	FlexCare-st	FlexCare
IHM	AUROC	0.8772 (0.003)	0.8726 (0.002)	0.8732 (0.006)	0.8804 (0.001)	0.8751 (0.002)	0.8749 (0.004)	**0.8823 (0.002)**
	AUPRC	0.5158 (0.006)	0.5133 (0.008)	0.5148 (0.017)	0.5136 (0.010)	0.5055 (0.006)	0.5116 (0.007)	**0.5372 (0.006)**
LOS	ma-F1	0.1487 (0.006)	0.1531 (0.006)	0.1549 (0.007)	**0.1554 (0.006)**	0.1503 (0.010)	0.1492 (0.005)	0.1479 (0.005)
	mi-F1	0.6289 (0.005)	0.6298 (0.003)	0.6267 (0.004)	0.6282 (0.004)	0.6307 (0.006)	0.6317 (0.001)	**0.6358 (0.003)**
DEC	AUROC	0.9396 (0.002)	0.9409 (0.001)	0.9406 (0.004)	0.9435 (0.001)	0.9470 (0.001)	0.9420 (0.002)	**0.9538 (0.001)**
	AUPRC	0.4782 (0.006)	0.4792 (0.010)	0.4911 (0.011)	0.4981 (0.008)	0.4922 (0.005)	0.4926 (0.010)	**0.5123 (0.006)**
PHE	ma-AUROC	0.8340 (0.001)	0.8362 (0.001)	0.8429 (0.001)	**0.8446 (0.001)**	0.8424 (0.000)	0.8417 (0.000)	0.8393 (0.005)
	mi-AUROC	0.8785 (0.001)	0.8769 (0.001)	0.8830 (0.001)	**0.8845 (0.000)**	0.8826 (0.000)	0.8820 (0.000)	0.8803 (0.004)
REA	AUROC	0.7598 (0.002)	0.7585 (0.002)	0.7618 (0.001)	0.7627 (0.002)	0.7622 (0.001)	0.7604 (0.002)	**0.7680 (0.002)**
	AUPRC	0.3618 (0.003)	0.3481 (0.008)	0.3562 (0.003)	0.3482 (0.006)	0.3526 (0.004)	0.3517 (0.003)	**0.3702 (0.004)**
DIA	ma-AUROC	0.6651 (0.007)	0.6715 (0.005)	0.6756 (0.006)	0.6692 (0.002)	0.6717 (0.005)	0.6750 (0.005)	**0.6845 (0.006)**
	mi-AUROC	0.8920 (0.002)	0.8960 (0.002)	0.8955 (0.001)	0.8960 (0.001)	0.8944 (0.001)	0.8948 (0.001)	**0.8984 (0.001)**