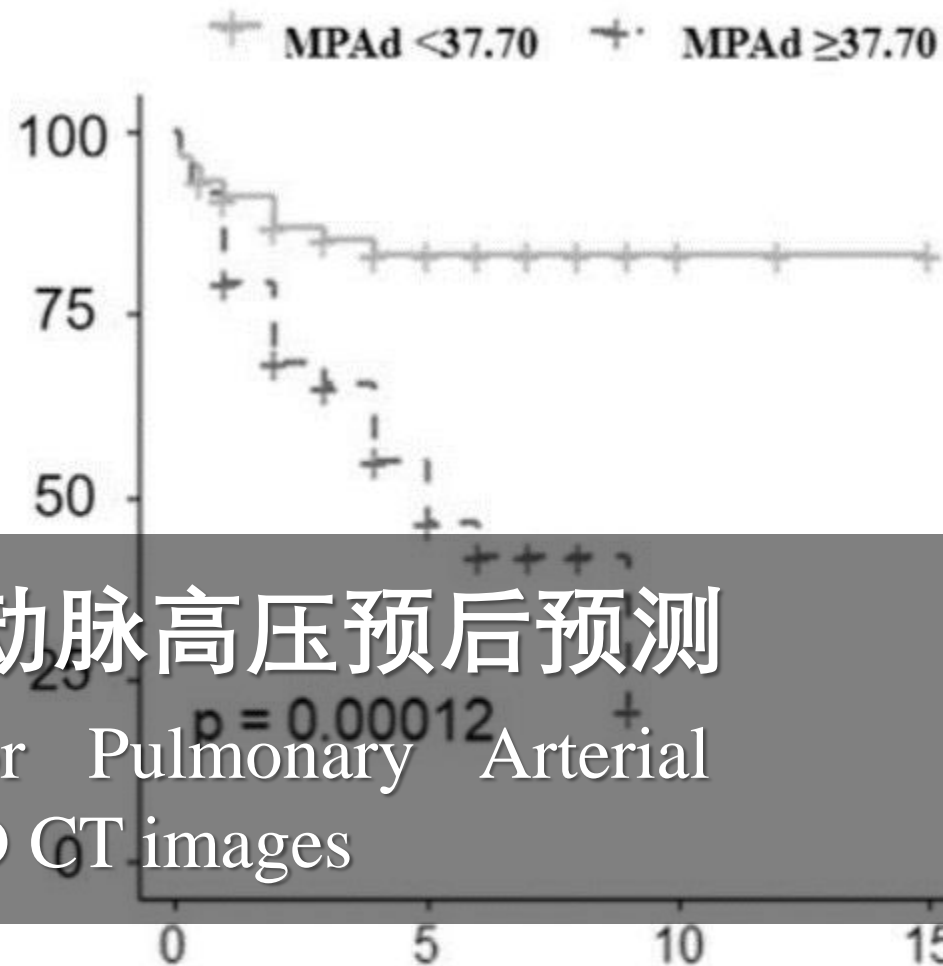


# $P^2$ -Net: 先验提示学习用于肺动脉高压预后预测

$P^2$ -Net: Priori Prompt Learning for Pulmonary Arterial Hypertension Prognosis Prediction on 3D CT images

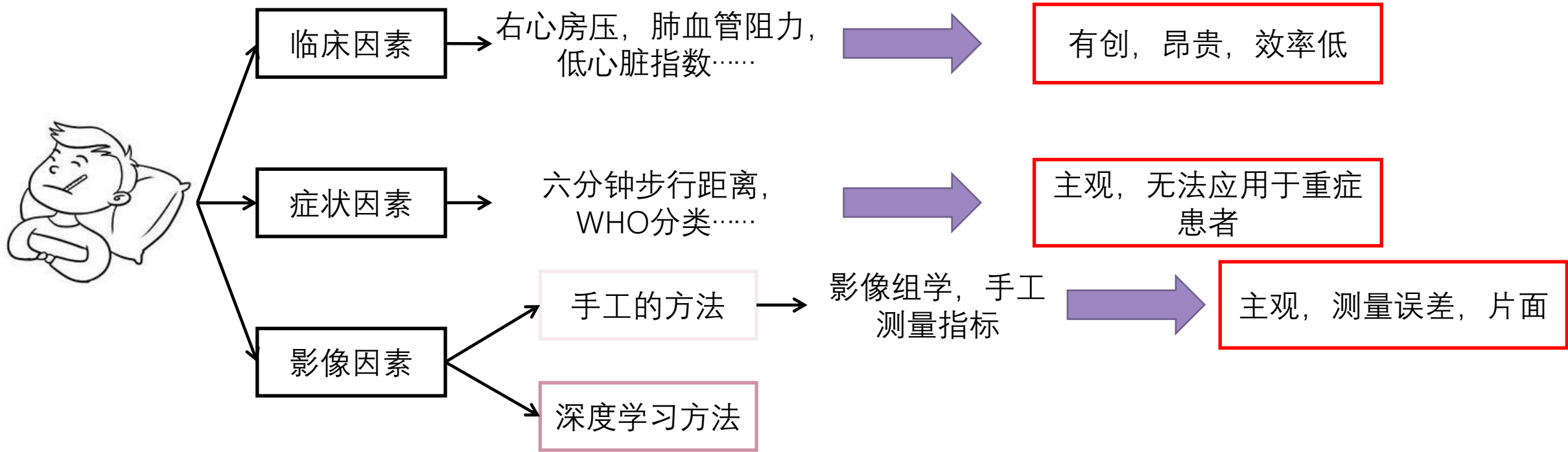
Survival probability



	Time, years			
Number at risk	0	5	10	15
MPAd <37.70	91	27	4	1
MPAd ≥37.70	49	13	0	0

# 什么是预后预测，意义是什么

寻找与病人预后表现相关的特征（因素），从而判断病人所属生存组别，针对不同组别选择治疗方案



# 基于深度学习的方法-COX 比例风险回归

$h_0(t)$ 基础风险函数，表示疾病在  $t$ 时刻的风险

$$h(X, t; \beta) = h_0(t)e^{DL(x)}$$

$h(t, X)$ 风险函数，表示病人在  $t$ 时刻的风险

$e^{DL(x)}$ 表示病人各因素对风险的影响

联合概率分布（似然）

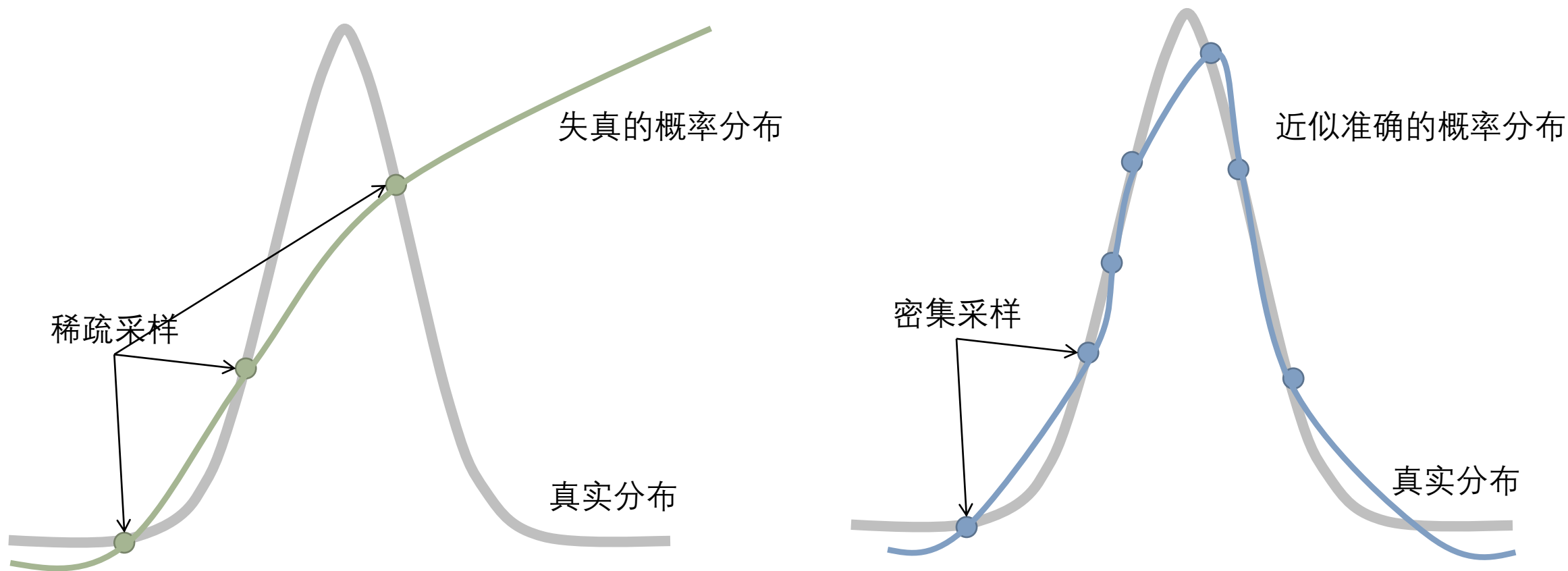
$$L(\beta) = \prod_{i=1}^I \frac{h(X_i, t; \beta)}{\sum_{k=i}^I h(X_k, t; \beta)} = \prod_{i=1}^I \frac{e^{DL(x_i)}}{\sum_{k=i}^I e^{DL(x_k)}}$$

极小化负对数似然

$$\min - \sum_{i=1}^I (DL(x_i) - \ln \sum_{k=i}^I e^{DL(x_k)})$$

基于概率分布的损失函数，需要对数据分布密集采样，获得有效的统计

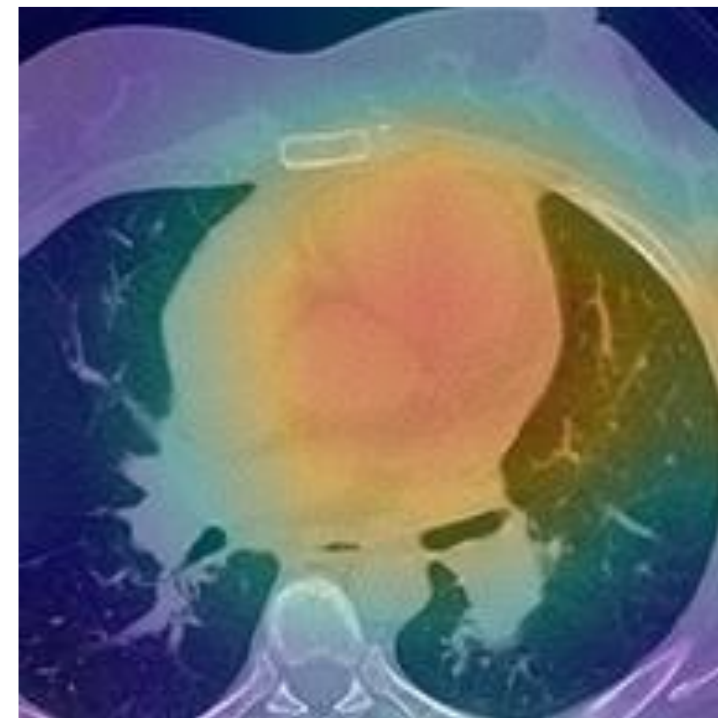
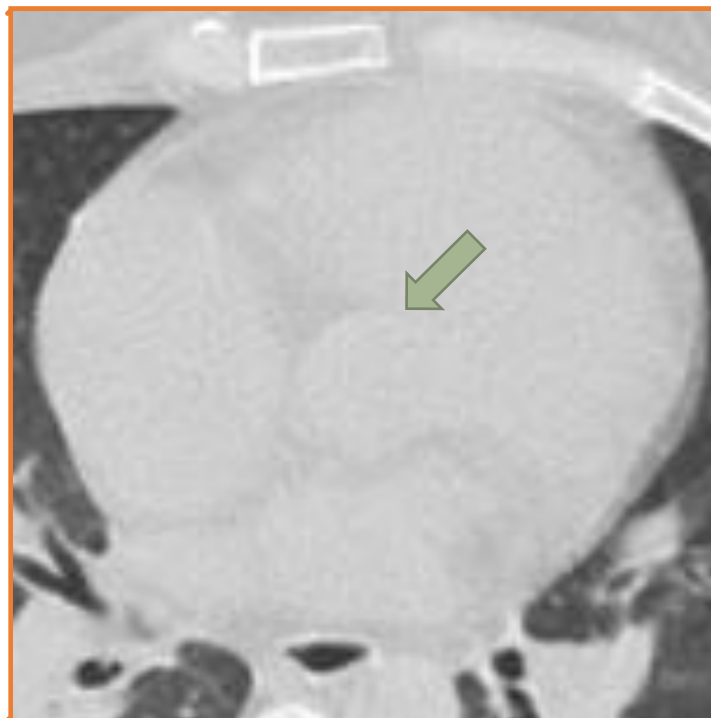
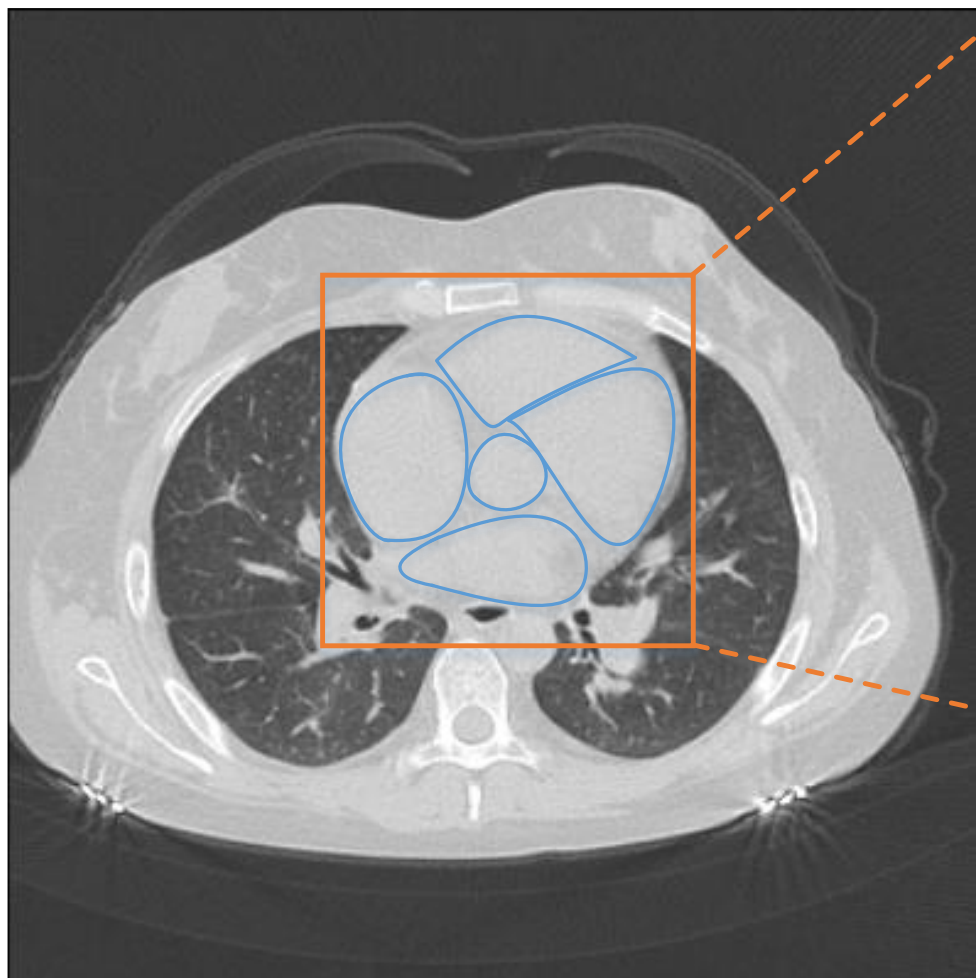
## 基于深度学习的方法-COX 比例风险回归



**挑战1：统计失效导致训练不稳定，难以获得有效的收敛。** 损失函数依靠概率分布，每次计算均需要大量统计数据，基于batch的损失函数计算方式受限于GPU存储空间只能获得少量的统计数据，导致概率分布统计失效，使得损失函数失真，训练不稳定。

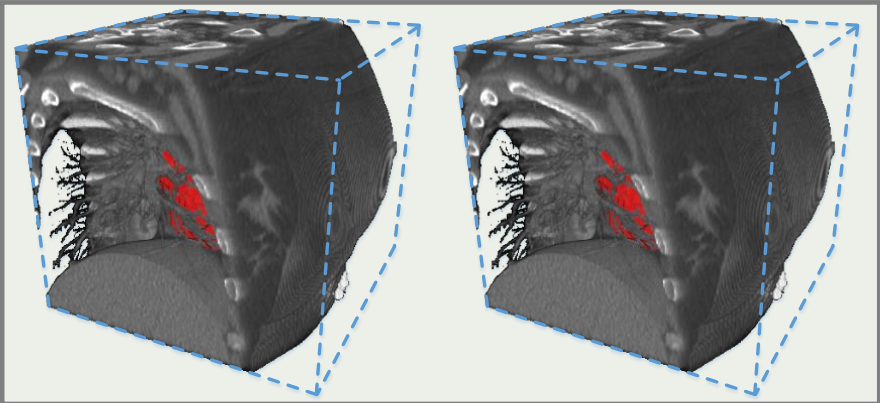
## 基于深度学习的方法-低效的特征提取

预后网络的CAM

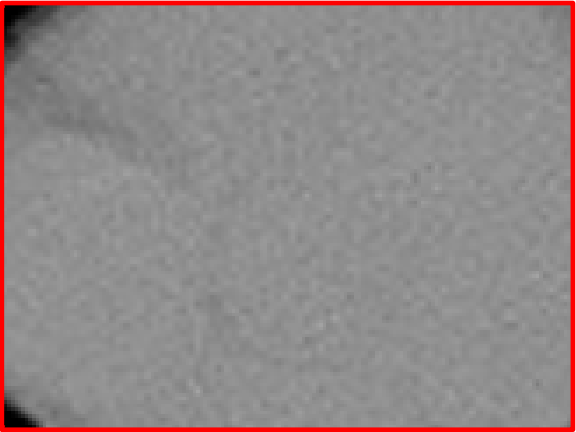
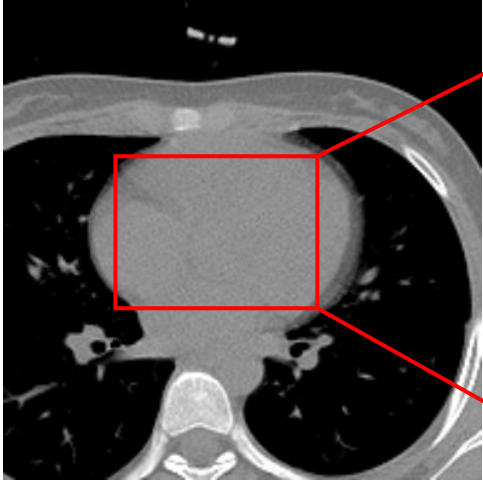
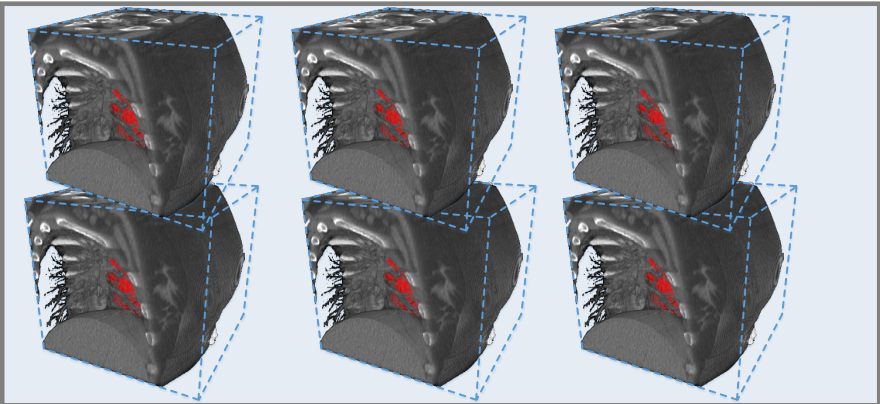
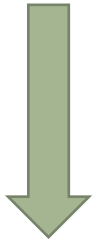


心脏内部对比度低，感兴趣结构边缘模糊，难以聚焦感知任务感兴趣特征。

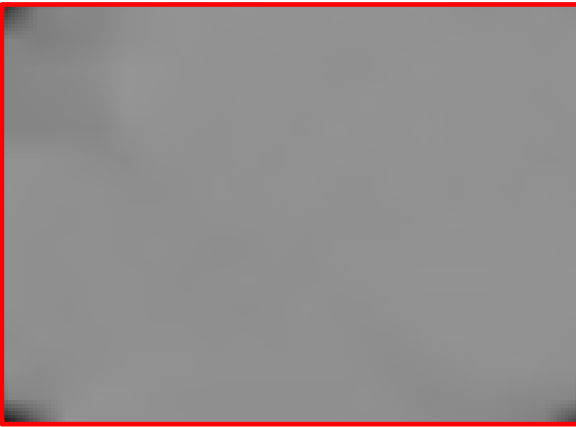
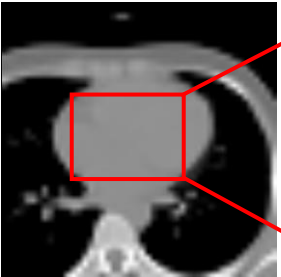
# 已有的策略-缩小尺寸



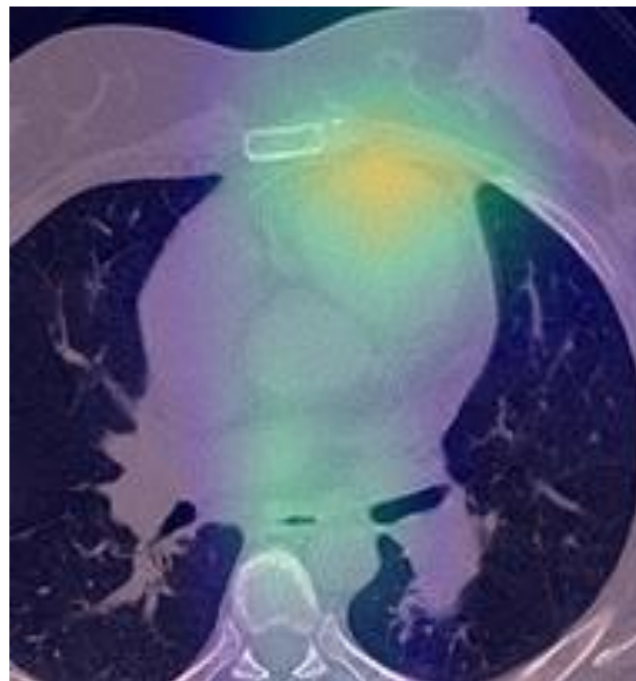
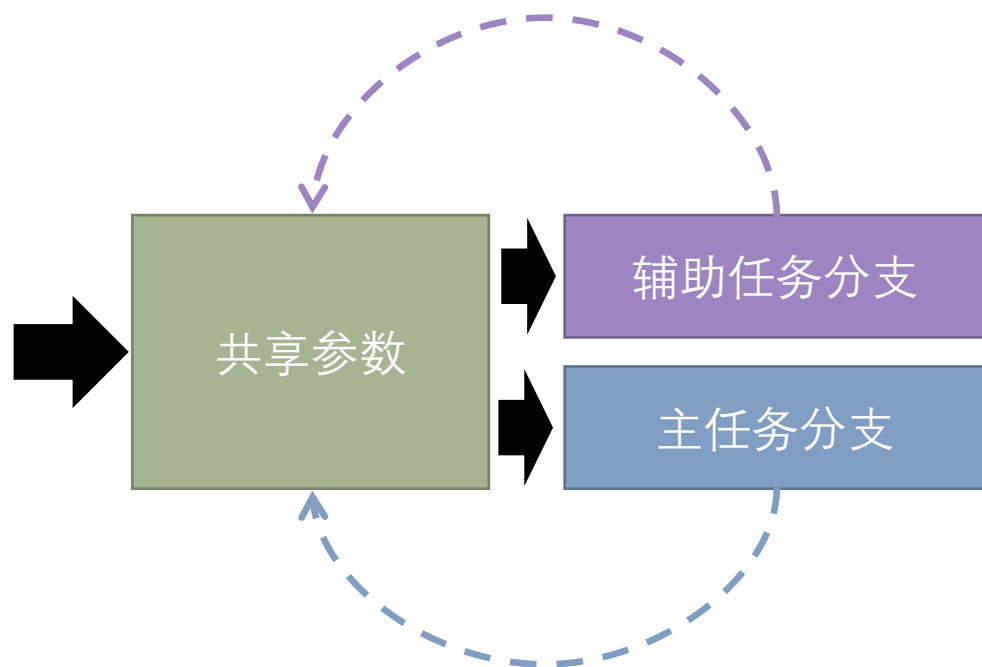
缩小尺寸



严重丢失本就不清晰的边界信息



## 已有的策略-辅助任务学习



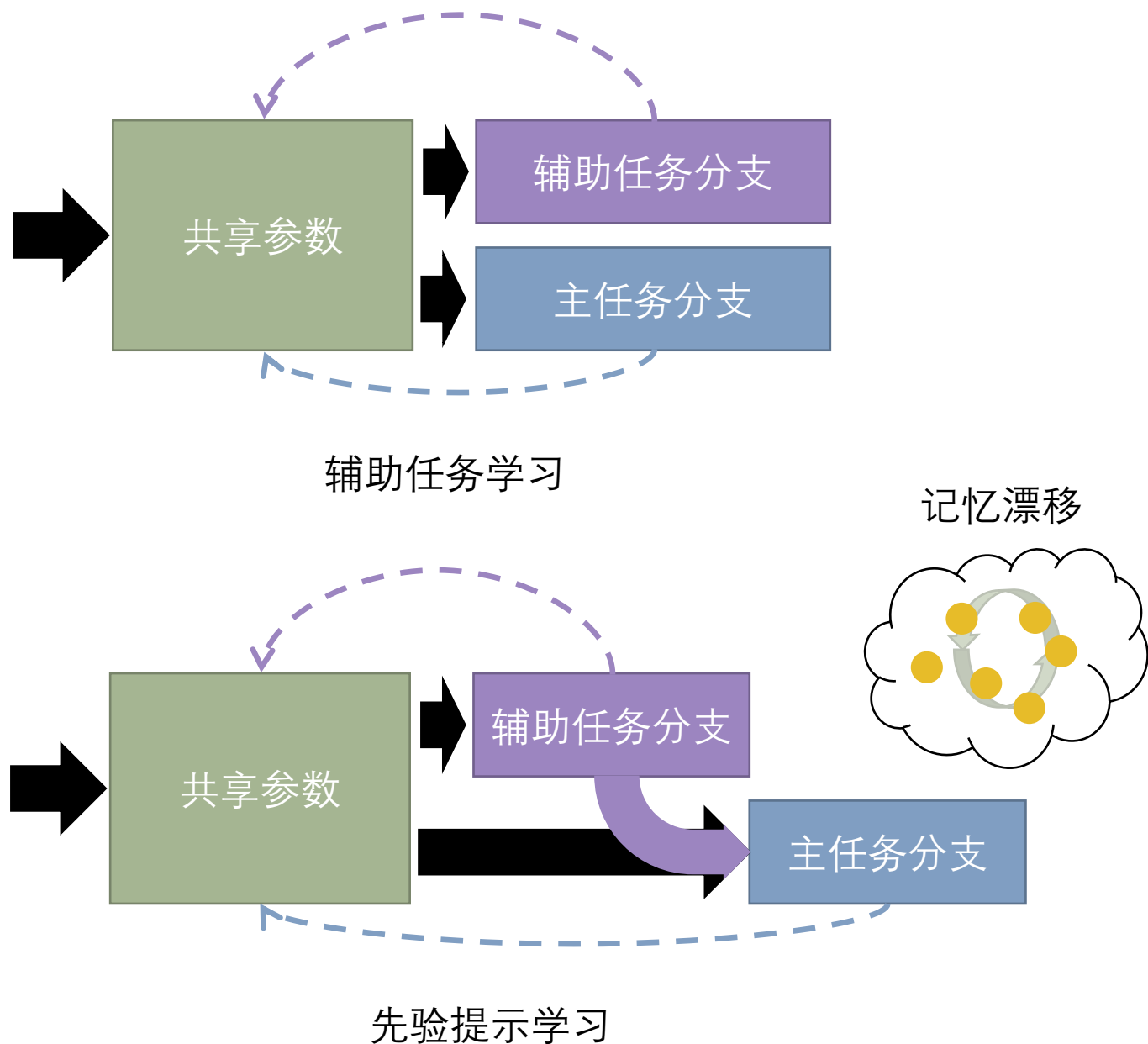
MPA估计分支的注意力图：**更加聚焦，关注重要区域**

### 限制：

1. 主导任务的不稳定性。
2. 低效率的知识嵌入。仅通过梯度引导底层共享参数的学习，无法嵌入高级的全局知识。



# 基于先验提示学习和记忆漂移的PAH预后预测

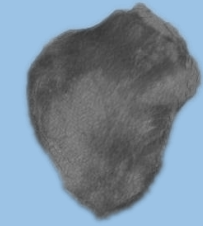


## 优势:

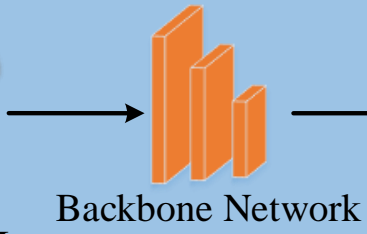
- 1. 隐先验嵌入。**通过梯度，利用共享表征偏好，引导网络关注与预后相关的特征区域，隐式地嵌入先验知识
- 2. 显先验嵌入。**直接将辅助任务预测结果作为提示输入主任务分支，显式地嵌入高级先验知识
- 3. 密集输出空间采样。**利用记忆库，保存大量的输出结果，从而无代价的获得大量的采样数据，计算具有统计意义的损失值
- 4. 动态记忆更新。**随着训练动态地更新记忆库，把老的样本丢弃，从而使记忆库中样本分布随输出分布变化。



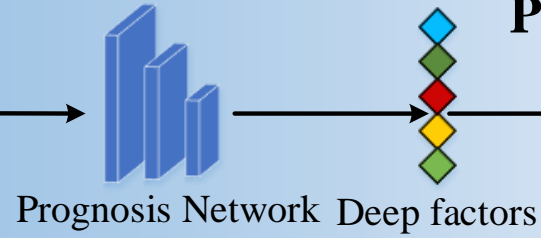
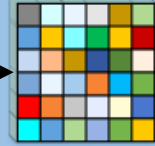
# $P^2$ -Net



Cardiac ROI

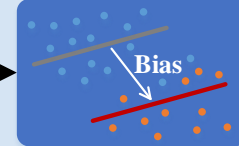


Feature maps



## \*Prognosis prediction branch *for prognostic factors estimation*

### Priori bias regression



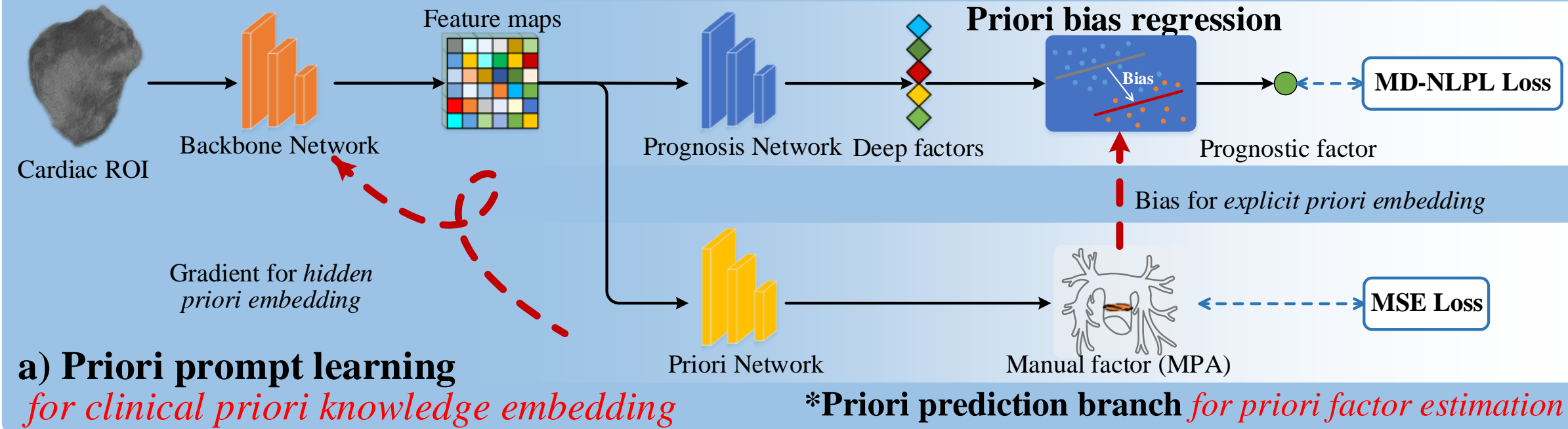
Bias

Prognostic factor

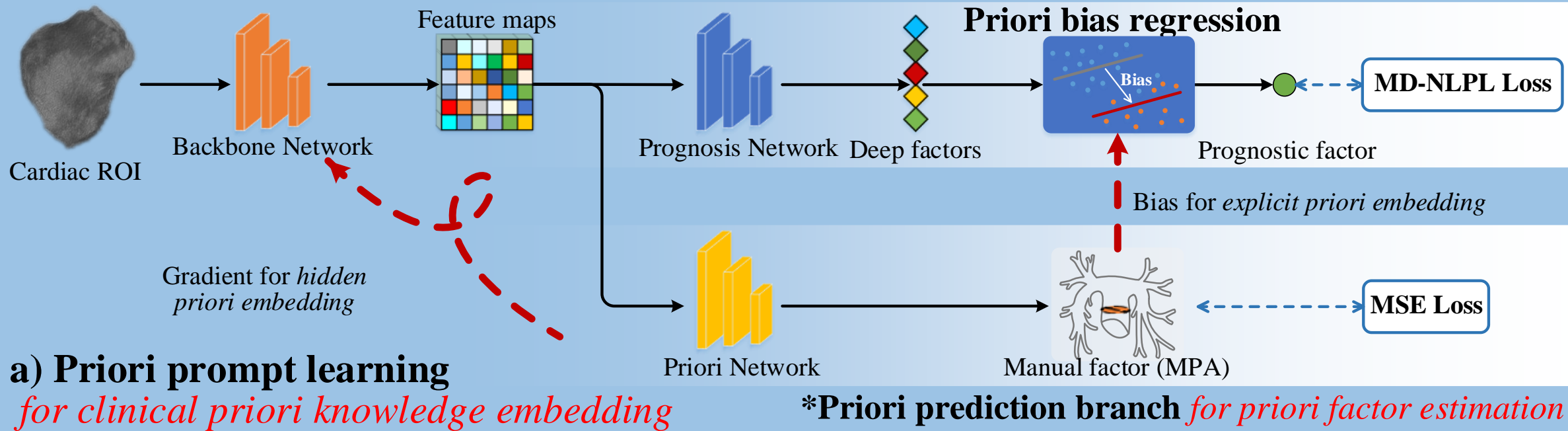


MD-NLPL Loss

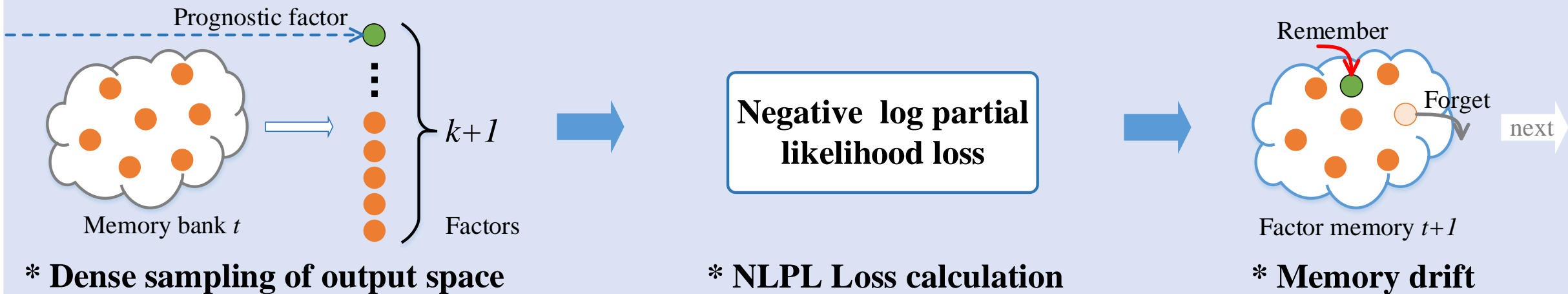
# $P^2$ -Net



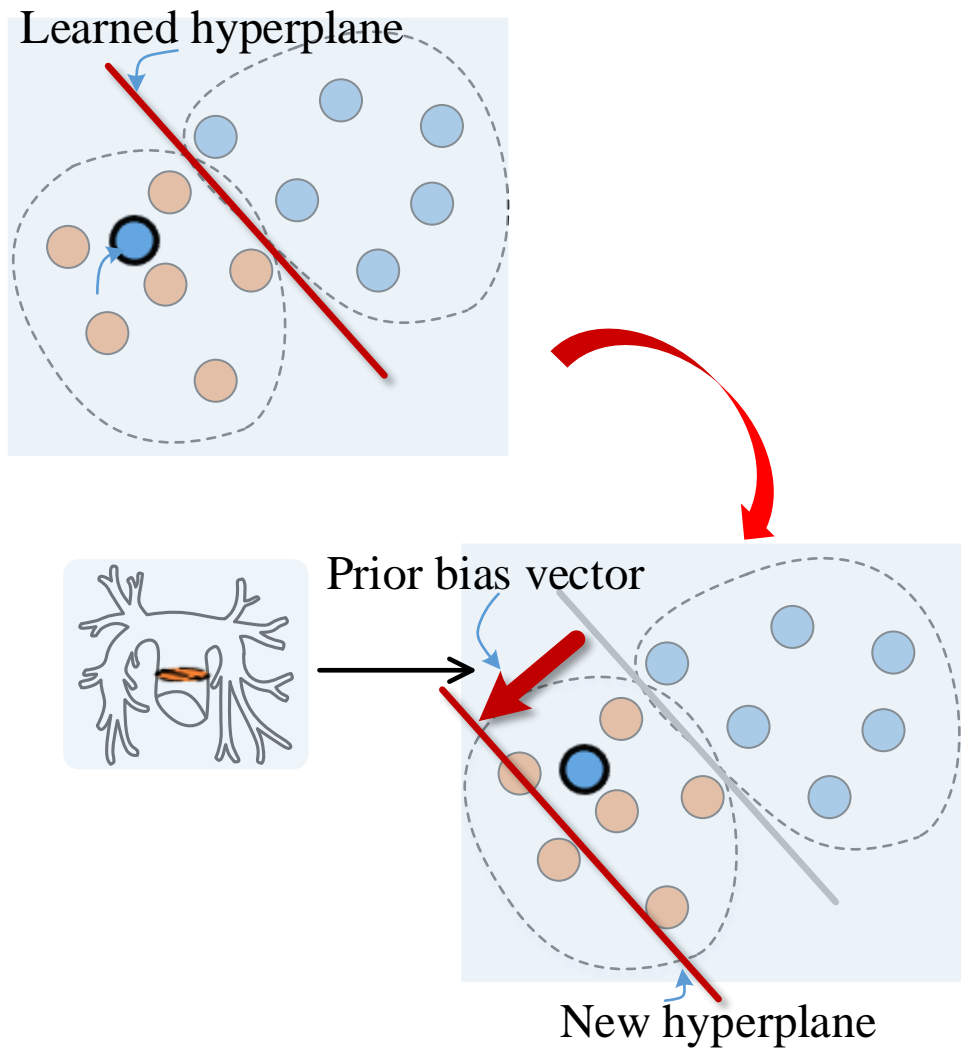
# $P^2$ -Net



## b) Memory Drift (MD-) NLPL Loss for optimization on representative statistics



# $P^2$ -Net



$$y = \mathbf{W}^T \times \mathbf{D} + b$$

COX回归公式

$$y = \mathbf{W}^T \times \mathbf{D} + y^{aux}$$

先验偏置回归

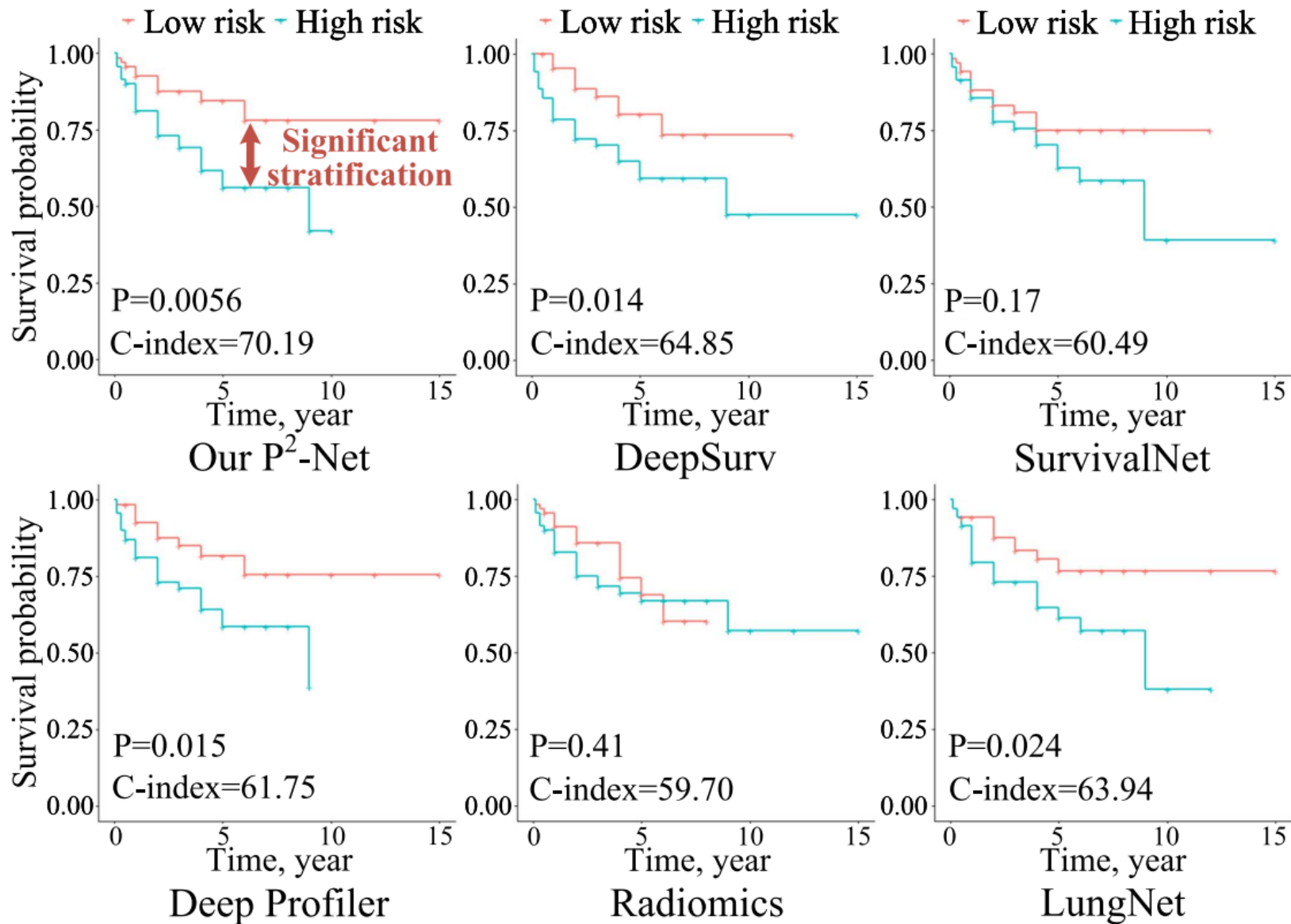
↑  
辅助任务输出

# 对比分析

Method	MD	HR (95% CI) ↑	C-index (%) ↑	logrank p ↓	Time-dependent AUC (%) ↑			
					TP2	TP5	TP8	TP11
MPA (Manual)	-	1.12 (1.04-1.20)	61.69	$1.2 \times 10^{-2}$	60.89	<b>71.69</b>	<b>73.22</b>	<b>80.51</b>
LPA (Manual)	-	1.13 (1.03-1.24)	61.29	$1.0 \times 10^{-1}$	55.18	63.63	65.72	71.71
RPA (Manual)	-	1.10 (1.01-1.24)	58.59	$1.1 \times 10^{-1}$	53.22	60.87	63.61	70.09
MPA/AAo (Manual)	-	1.51 (0.31-7.49)	51.39	$5.5 \times 10^{-1}$	54.11	55.79	57.96	62.36
MPA/DAo (Manual)	-	1.57 (0.49-5.08)	52.18	$7.0 \times 10^{-1}$	53.30	58.88	60.61	64.47
Radiomics [7]	-	1.05 (0.82-1.34)	59.70	$4.1 \times 10^{-1}$	55.35	49.43	41.60	44.94
DeepSurv [9]	×	1.50 (1.14-1.98)	60.38	$1.7 \times 10^{-1}$	52.81	57.67	59.91	66.40
SurvialNet [12]	×	1.50 (1.11-2.01)	60.47	$2.3 \times 10^{-1}$	59.73	56.64	56.94	59.46
LungNet [11]	×	1.69 (1.27-2.24)	63.20	$2.1 \times 10^{-2}$	64.07	62.93	64.49	67.86
(Multi-task) Deep Profiler [10]	×	1.14 (0.80-1.61)	52.65	$7.8 \times 10^{-1}$	49.70	52.21	51.35	63.34
(Multi-task) Cross-Stitch [41]	×	1.50 (1.06-2.13)	59.67	$2.9 \times 10^{-2}$	61.01	63.81	58.02	66.91
(Multi-task) Multi-Head [40]	×	1.59 (1.16-2.20)	0.58	$1.2 \times 10^{-1}$	56.92	62.31	64.47	69.93
DeepSurv [9]	✓	1.65 (1.17-2.33)	64.85	$1.4 \times 10^{-2}$	65.96	66.04	62.91	66.20
SurvialNet [12]	✓	1.59 (1.11-2.27)	60.49	$1.7 \times 10^{-1}$	58.98	59.83	61.98	66.73
LungNet [11]	✓	1.68 (1.28-2.20)	63.94	$2.4 \times 10^{-2}$	65.39	63.89	65.01	68.61
(Multi-task) Deep Profiler [10]	✓	1.55 (1.05-2.27)	61.75	$1.5 \times 10^{-2}$	64.96	63.80	62.24	68.48
(Multi-task) Cross-Stitch [41]	✓	1.84 (1.29-2.62)	64.87	$3.0 \times 10^{-2}$	63.77	67.22	58.22	65.74
(Multi-task) Multi-Head [40]	✓	1.64 (1.16-2.31)	64.93	$6.2 \times 10^{-2}$	65.79	64.45	54.94	58.46
<b>Our P<sup>2</sup>-Net wo MD</b>	×	<b>1.97 (1.39-2.79)</b>	<b>65.24</b>	<b><math>2.9 \times 10^{-3}</math></b>	<b>66.57</b>	68.92	70.65	76.92
<b>Our P<sup>2</sup>-Net</b>	✓	<b>2.14 (1.52-3.02)</b>	<b>70.19</b>	<b><math>5.6 \times 10^{-3}</math></b>	<b>72.41</b>	<b>76.15</b>	<b>74.98</b>	<b>78.94</b>

我们的方法在各项指标上都表现出了更高的性能。

# 对比分析



# 模型分析

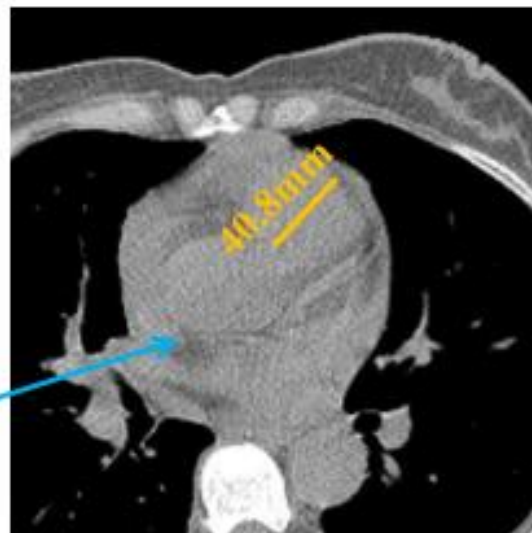
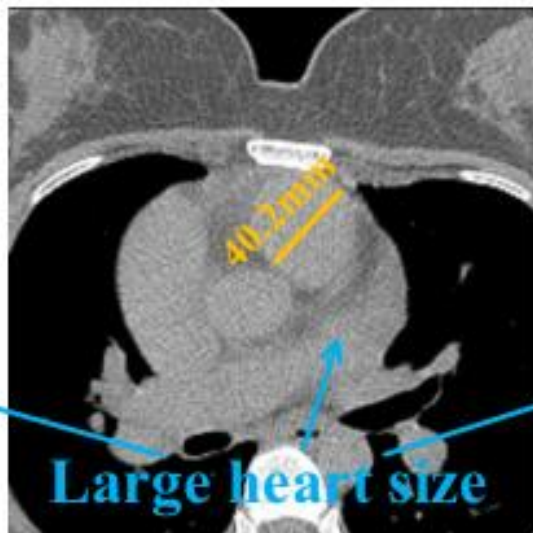
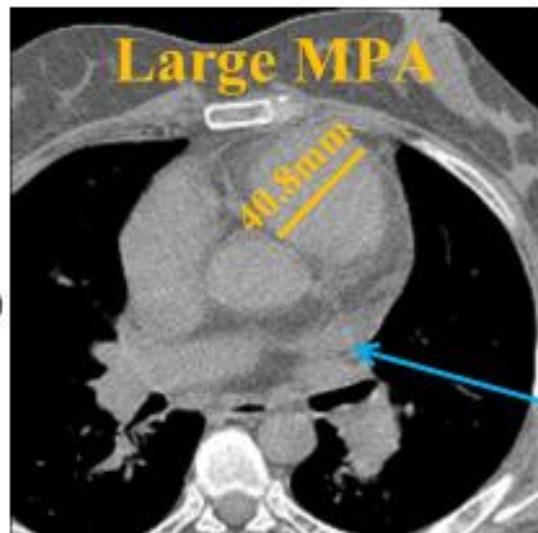
MT	MD	PPL	HR (95% CI) ↑	C-index (%) ↑	Prior
			1.50 (1.14-1.98)	60.38	-
		✓	1.97 (1.39-2.79)	65.24	Hidden & explicit
	✓		1.65 (1.17-2.33)	64.85	-
✓	✓		1.64 (1.16-2.31)	64.93	Hidden
	✓	✓	<b>2.14 (1.52-3.02)</b>	<b>70.19</b>	Hidden & explicit

**消融实验：** DeepSurv仅使用记忆漂移，可获得64.85的C-index，加上多任务学习，由于先验特征的引入，c-index增加到64.93。再加入我们的先验提示学习策略，c-index进一步提高到70.19.

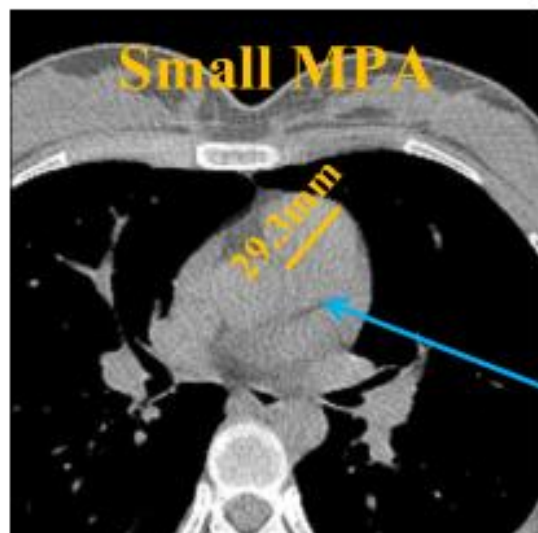


# 模型分析

High risk

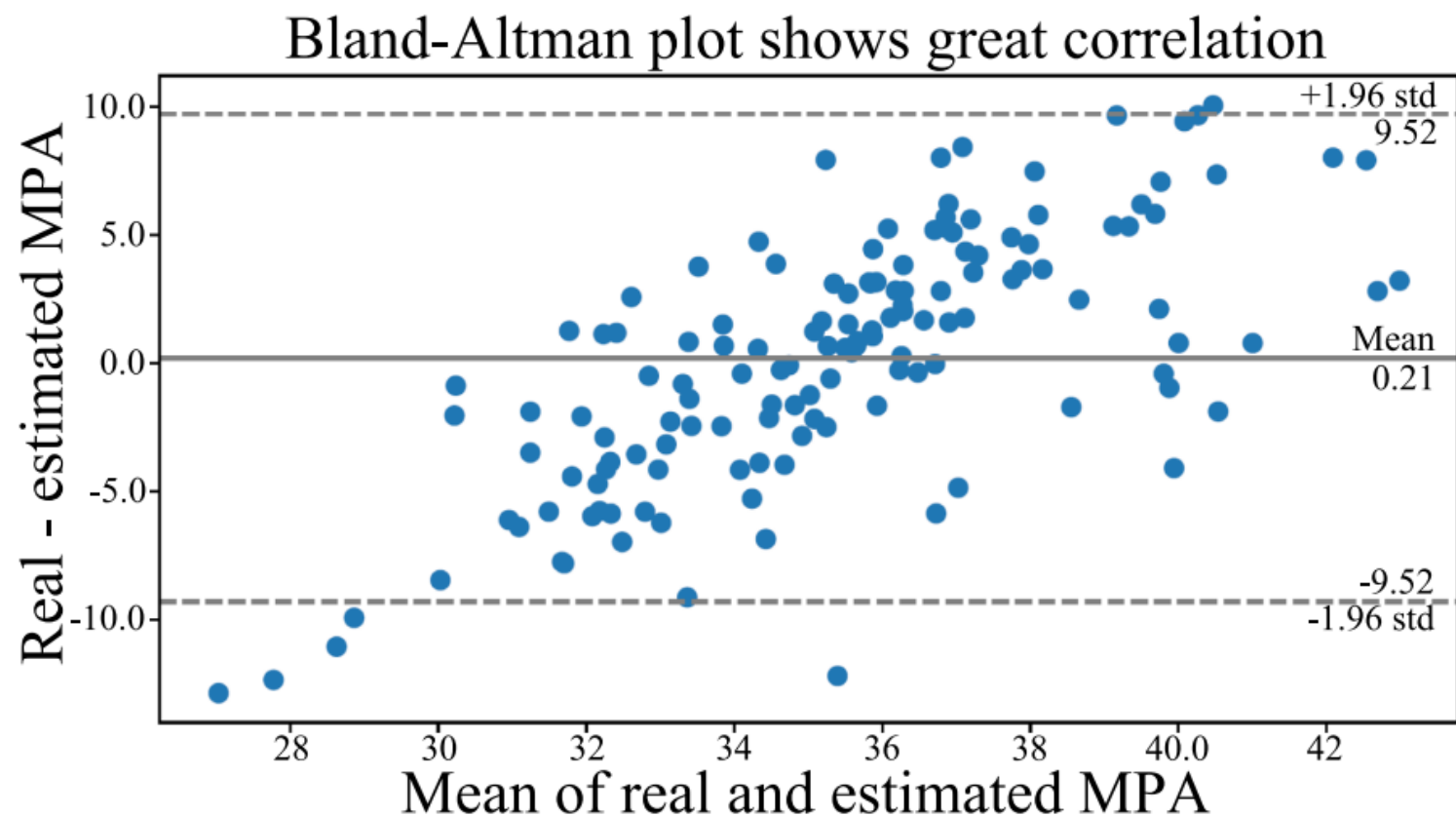


Low risk



**分组分析：**利用我们模型输出的预后因素可视化各组别对象。高风险普遍表现出更大的的心脏体积和更大的肺动脉直径。这是因为分动脉高压累积在肺动脉和心室，使得心脏膨大，肺动脉变粗。低风险则心脏更小，肺动脉更细。这证明我们模型预后结果的准确性。

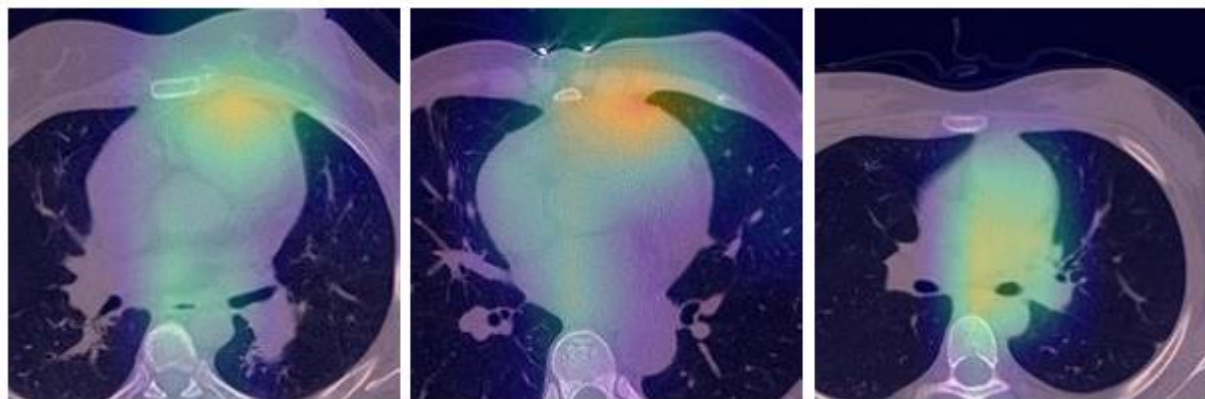
# 模型分析



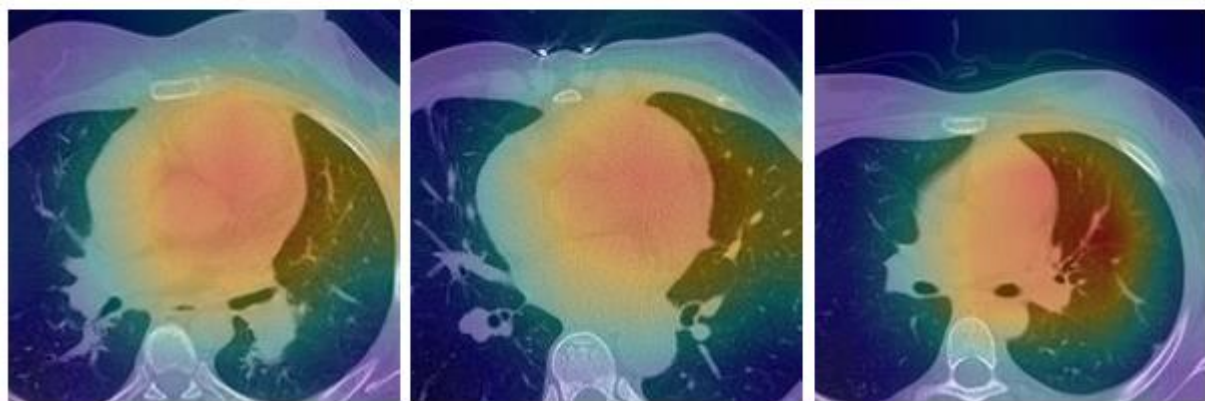
**MPA估计准确性分析：**MPA估计具有临床参考意义，能够为模型的临床使用提供一定的参考。利用Bland-Altman分析，模型估计MPA和医生测量MPA普遍在95%LoA之内，在临床上可接受。

# 模型分析

**CAM from Priori Network**



**CAM from Prognosis Network**



**各分支关注区域分析：**利用CAM图分析模型关注区域，先验网络拥有更聚焦的关注区域，能够为主网络提供好的偏好。

感谢大家，欢迎提问