SOM2LM: Self-Organized Multi-Modal Longitudinal Maps

Jiahong Ouyang[1], Qingyu Zhao[2], Ehsan Adeli[1], Greg Zaharchuk[1⋆], and Kilian M. Pohl[1⋆⋆]

[1] Stanford University, Stanford CA 94305, USA
[2] Cornell University, Ithaca NY 14850, USA

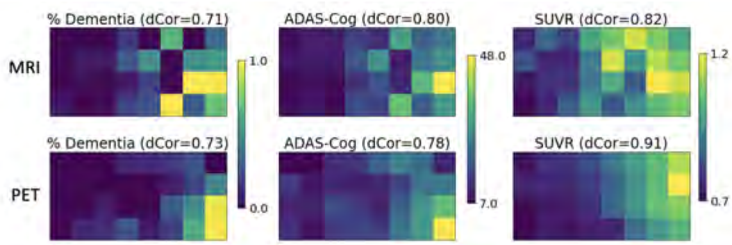

Fig. S1: The color at each SOM representation encodes the average value of (a) % of dementia, (b) ADAS-Cog score across the training samples of that cluster, and (c) amyloid summary SUVR. The learned SOM grids are stratified by disease abnormality with high correlations with those three disease abnormality-related markers.

⋆ co-founder, equity Subtle Medical
⋆⋆ corresponding author

	MRI	PET
Input	303 ROI features, normalized by z-score	160 ROI SUVRs to the composite region
Encoder $F_{(\circ)}$	FC: 256 ReLU FC: 256 ReLU FC: 64 Normalized to unit vector	FC: 128 ReLU FC: 128 ReLU FC: 64 Normalized to unit vector
Decoder $H_{(\circ)}$	FC: 256 ReLU FC: 256 ReLU FC: 303 -	FC: 128 ReLU FC: 128 ReLU FC: 160 SoftPlus
Predictor P	FC: 64 ReLU FC: 64 Sigmoid	

Table S1: Network architectures. FC denotes fully connected layers.

	MRI (m)	PET (p)
$\lambda_{commit,(\circ)}$	50	200
$\lambda_{prox,(\circ)}$	50	200
$\lambda_{long,(\circ)}$	5	0.001
λ_{multi}	0.001	
$\alpha_{(\circ)}$	0.1	0.1
$\alpha_{(m,p)}$	0.01	
τ_{min}, τ_{max}	0.1, 1.0	
N_r, N_c	4, 8	

Table S2: Weighing parameters and other thresholds in loss terms.

	Amyloid Status	MRI converter
Pre-training	Setting: SGD, lr=0.01, momentum=0.9, weight decay=10^{-5}, bs=64	
	20 epochs, MRI-specific SOM: $min_{F_{(m)},H_{(m)},\mathcal{G}_{(m)}}(L_{recon,(m)} + \lambda_{commit,(m)} \cdot L_{commit,(m)} + \lambda_{prox,(m)} \cdot L_{prox,(m)} + \lambda_{long,(m)} \cdot L_{long,(m)})$ 20 epochs, PET-specific SOM: $min_{F_{(p)},H_{(p)},\mathcal{G}_{(p)}}(L_{recon,(p)} + \lambda_{commit,(p)} \cdot L_{commit,(p)} + \lambda_{prox,(p)} \cdot L_{prox,(p)} + \lambda_{long,(p)} \cdot L_{long,(p)})$	
	10 epochs, $min_{F_{(m)},H_{(m)},\mathcal{G}_{(m)}}(L_{recon,(m)} + \lambda_{commit,(m)} \cdot L_{commit,(m)} + \lambda_{prox,(m)} \cdot L_{prox,(m)} + \lambda_{long,(m)} \cdot L_{long,(m)} + \lambda_{multi} \cdot L_{multi})$	10 epochs, $min_{F_{(m)},F_{(p)},H_{(m)},H_{(p)},\mathcal{G}_{(m)},\mathcal{G}_{(p)}} \sum_{\circ \in \{u,v\}}(L_{recon,(\circ)} + \lambda_{commit,(\circ)} \cdot L_{commit,(\circ)} + \lambda_{prox,(\circ)} \cdot L_{prox,(\circ)} + \lambda_{long,(\circ)} \cdot L_{long,(\circ)} + \lambda_{multi} \cdot L_{multi})$
Frozen	Setting: SGD, lr=0.001, momentum=0.9, weight decay=10^{-5}, bs=64	
	regress out age from $z_{(m)}$, $min_P BCE(P(z_{(m)}), y_{(m)})$	$min_P BCE(P([z_{(m)}, z_{(p)}]), y_{(m)})$
Fine-tuned	Setting: SGD, lr=0.001, momentum=0.9, weight decay=10^{-5}, bs=64	
	$min_{F_{(m)},P} BCE(P(z_{(m)}), y_{(m)})$	$min_{F_{(m)},F_{(p)},P} BCE(P([z_{(m)}, z_{(p)}]), y_{(m)})$

Table S3: Training strategy and hyperparameter setup. P denotes the predictor, and BCE stands for binary cross entropy loss.

Methods	Amyloid Status				MCI converter			
	Frozen		Fine-tuned		Frozen		Fine-tuned	
	BACC	AUC	BACC	AUC	BACC	AUC	BACC	AUC
w/o L_{multi}	0.65	0.69	0.68	0.75	0.63	0.69	0.64	0.74
with L_{multi}	**0.66**	**0.75**†	**0.74**	**0.80**†	**0.67**	**0.74**†	**0.67**	**0.75**

Table S4: Ablation study of removing L_{multi}. Building the cross-modal relationship via L_{multi} achieved significantly superior performance than only using modality-specific SOMs. ($\dagger : p < 0.05$, Delong's test).