SepCLR: Separating common from salient patterns with Contrastive Representation Learning

Contributions :

- Based on the **InfoMax Principle**, we introduce **SepCLR**, a novel theoretical framework for **Contrastive Analysis (CA)**.
- To estimate and maximize the **Mutual Information terms**, we use **Contrastive Representation Learning (CLR)**.
- We introduce **k-JEM**, a strategy to reduce information leakage between two blocks of the latent space.

Caution: do not confond **Contrastive Learning** and **Contrastive Analysis** ! These are different !

The InfoMax principle for Contrastive Analysis

$$\arg\max_{\theta} \underbrace{\lambda_C(I(x;c)+I(y;c))}_{\text{Common InfoMax}} + \underbrace{\lambda_S I(y;s)}_{\text{Salient InfoMax}} \quad \text{s.t.} \quad \underbrace{D_{KL}(s_x||\delta(s'))=0}_{\text{Information-less hyp.}} \text{ and } \underbrace{I(c,s)=0}_{\text{Independence hyp.}} \qquad (1)$$

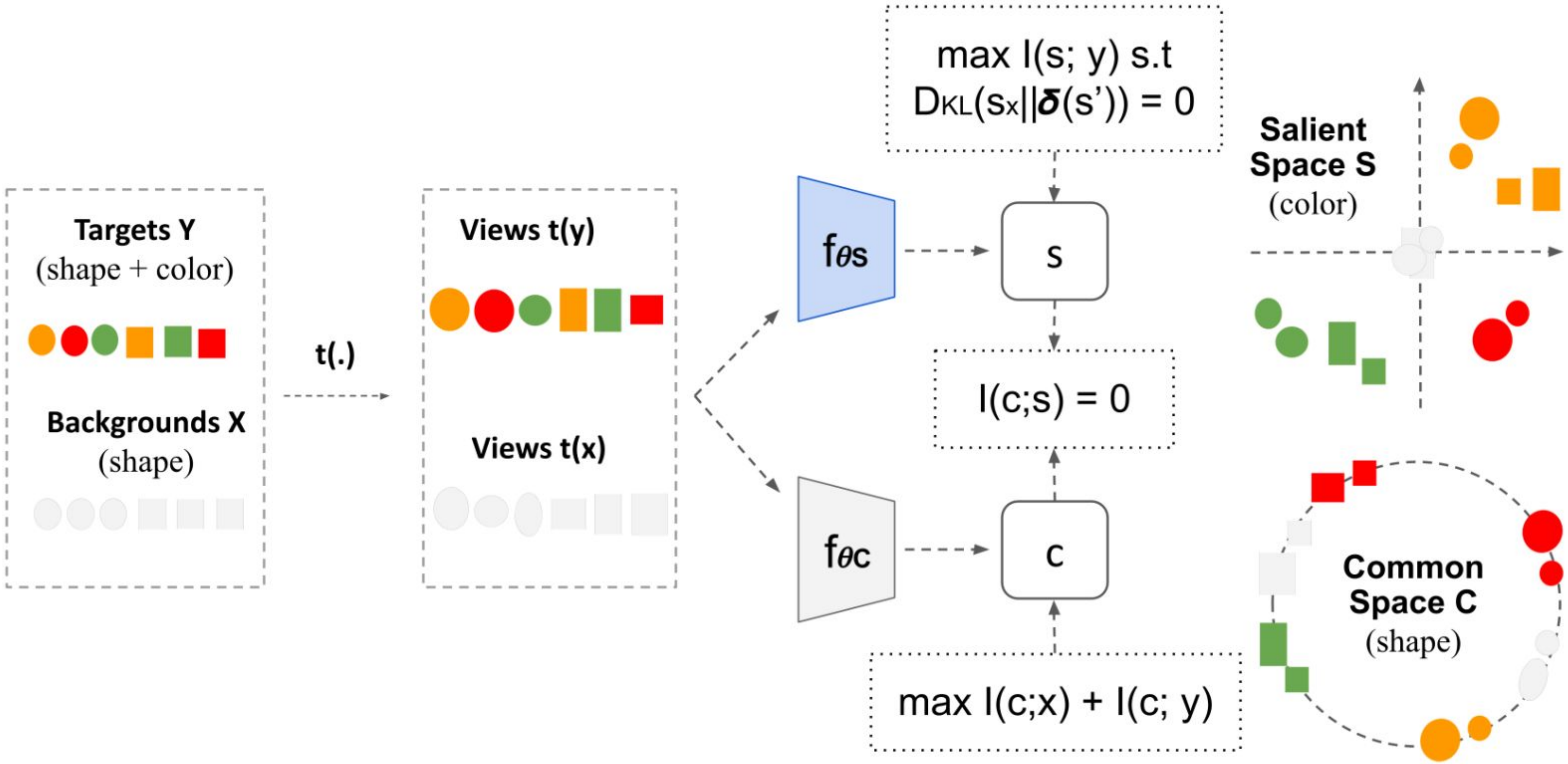

Fig. 3: Scheme of SepCLR. SepCLR is trained to identify and separate the salient patterns (color variations) of the target dataset Y from the common patterns (shape) shared between background X and target dataset Y.

The Common InfoMax principle terms

Objective: **Estimate the common InfoMax term:** I(x; c) + I(y; c).

$$I(x;c) = \underbrace{-\mathbf{E}_{x\sim p_x} H(c|x)}_{\text{Alignment}} + \underbrace{H(c)}_{\text{Entropy}} \quad (2)$$

Entropy term: as in Wang et al. 2021, it can be estimated with a Kernel Density Estimation (KDE) technique.

- H(c) $\lesssim$

$$\mathcal{L}_{\text{uniform}} = \log \frac{1}{N_X+N_Y} \sum_{i=1}^{N_X+N_Y} \frac{1}{N_X+N_Y} \sum_{j=1}^{N_X+N_Y} \exp \frac{-||f_{\theta_C}(v_i) - f_{\theta_C}(v_j)||_2^2}{2\tau} + \underbrace{\log(\sqrt{2\pi\tau})}_{\text{Constant term}} \quad (3)$$

Alignment term: In our paper, we derive a Multi-View alignment term with a KDE, linking the usual Alignment term with Mutual Information:

- $E_{x\sim p_x}$ H(c | x) =

$$\mathcal{L}_{\text{align}} = -\frac{1}{N_X+N_Y} \sum_{i=1}^{N_X+N_Y} \log \frac{1}{K} \sum_{k=1}^{K} \exp \frac{-||f_{\theta_C}(v_i) - f_{\theta_C}(v_i^k)||_2^2}{2\tau} + \underbrace{\log(\sqrt{2\pi\tau})}_{\text{Constant term}} \quad (4)$$

In the case where … en et al, 2020 (SimCLR).

Wang et al., Understanding Contrastive Learning with alignment and uniformity, ICLR 2021.
Chen et al., A Simpe Framework for Contrastive Learning of Visual Representations, PMLR 2021.

The Common InfoMax principle terms

Objective: **Maximize the common InfoMax term:** $I(x; c) + I(y; c)$**.**

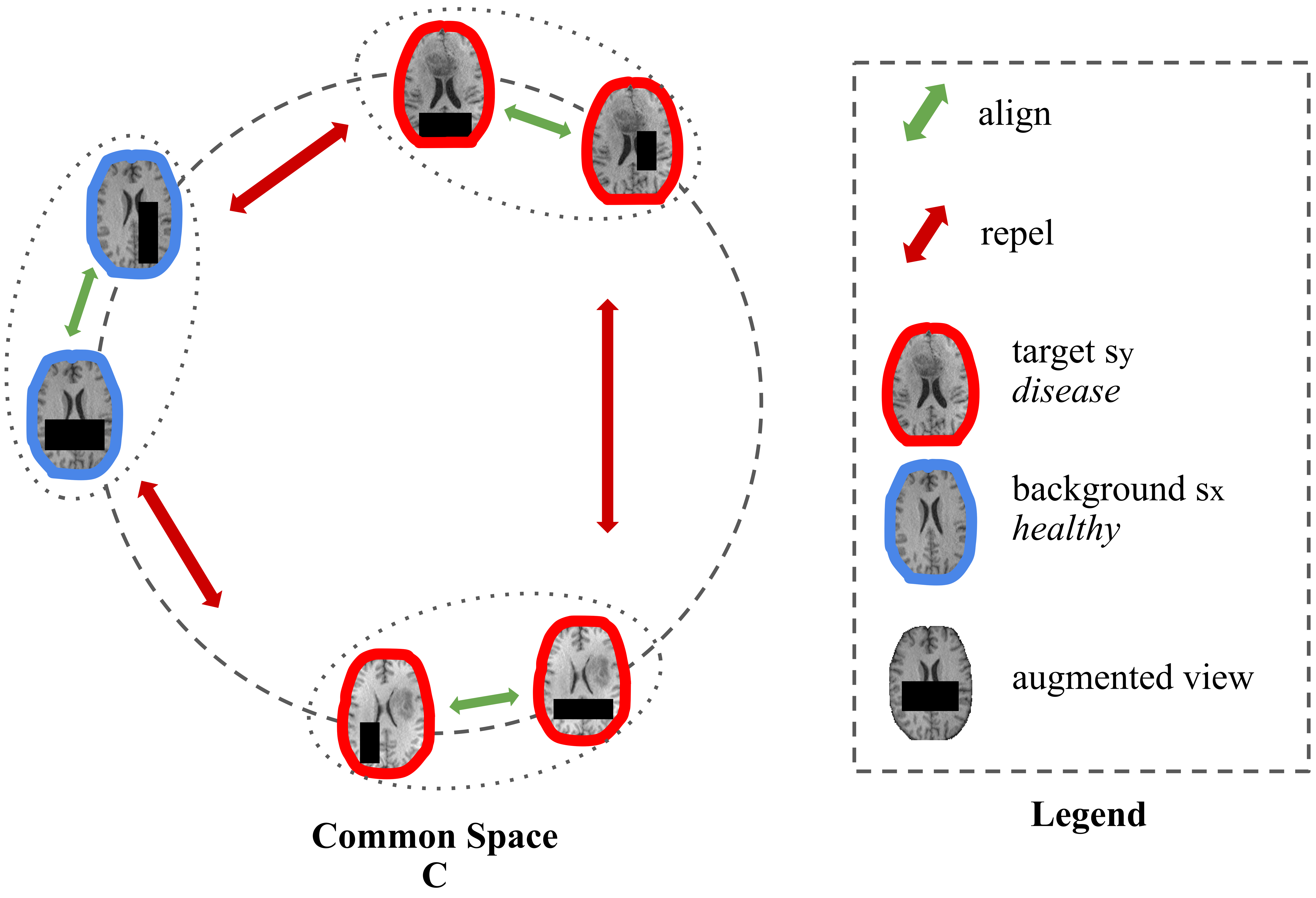

Fig. 4: Effect of the Common InfoMax on the common latent space.

The constrained Salient InfoMax term

Objective: **Estimate the salient InfoMax term constrained by the information-less hyp:** max $I(y; s)$ s.t $D_{KL}(s_x; \boldsymbol{\delta}(s'))=0$.

$$\arg\max I(s;y) = -\underbrace{\mathbf{E}_{y\sim p_y} H(s|y)}_{\text{Target-only Alignment}} + \underbrace{H(s)}_{s'\text{-Entropy}} \quad \text{s.t. } \underbrace{D_{KL}(s_x||\delta(s')) = 0}_{\text{Information-less hyp.}} \tag{5}$$

Target-only Alignment term:

- $E_{y\sim p_y} H(s \mid y)$ =

$$\mathcal{L}_{\text{y-align}} = -\frac{1}{N_Y}\sum_{i=1}^{N_Y} \log \frac{1}{K}\sum_{k=1}^{K} \exp \frac{-||f_{\theta_C}(v_i) - f_{\theta_C}(v_i^k)||_2^2}{2\tau} + \underbrace{\log(\sqrt{2\pi\tau})}_{\text{Constant term}} \tag{6}$$

s'- Entropy term:

- $H(s) \lesssim$

$$\mathcal{L}_{s'\text{-unif}} = \log \frac{1}{N_Y}\sum_{i=1}^{N_Y}\left(\exp\frac{-||f_{\theta_s}(t(y_i)) - s'||_2^2}{\tau} + \frac{1}{N_Y}\sum_{j=1}^{N_Y}\exp\frac{-||f_{\theta_s}(t(y_i)) - f_{\theta_s}(t(y_j))||_2^2}{2\tau}\right) \tag{7}$$

Re-writin

$$\mathcal{F}(\theta_S, \beta; x, y, s) = -\lambda_S \mathcal{L}_{y\text{-align}} - \lambda_S \mathcal{L}_{s'\text{-unif}} - \beta\frac{1}{N_X}\sum_{i=1}^{N_X}\frac{||f_{\theta_s}(t(x_i)) - s'||_2^2}{2\tau} \tag{8}$$

The constrained Salient InfoMax term

Objective: **Maximize the salient InfoMax term constrained by the information-less hyp:** $I(y; s)$ s.t $D_{KL}(s_x; \boldsymbol{\delta}(s')) = 0$.

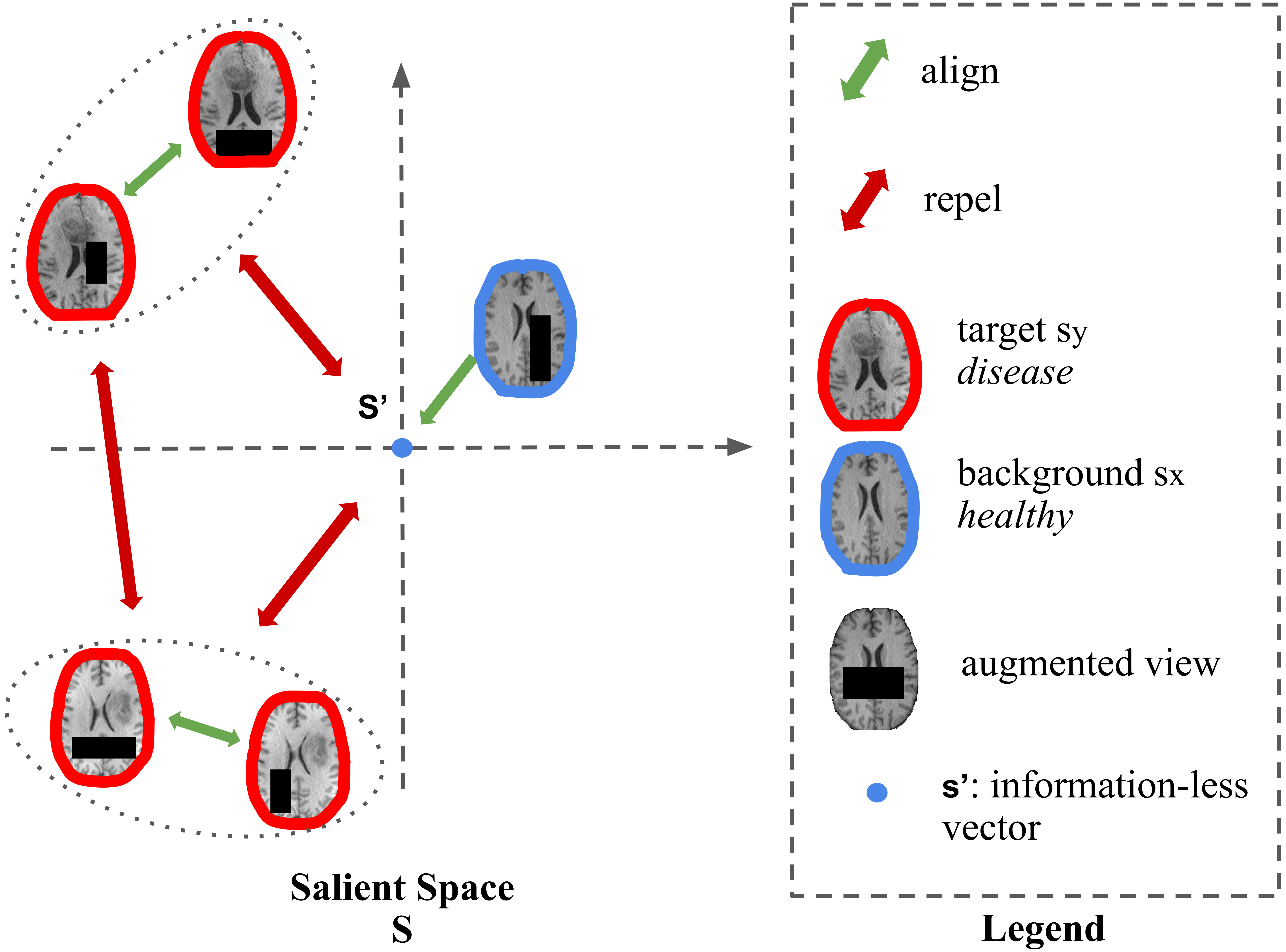

Fig. 5: Effect of the Salient InfoMax on the latents of target and background samples.

The independence term between common and salient spaces.

Objective: **Estimate and minimize the Mutual Information term:** I(c; s).

$$\boxed{\text{By def., } I(c;s) = -H(c,s) + H(c) + H(s)} \geq 0$$

Entropy terms: H(c) + H(s).
The other losses should absolutely not minimize these terms !

k-JEM - Kernel estimation of the Joint Entropy term: - H(c; s).

Interestingly, the Mutual Information is null whenever H(c; s) = H(c) + H(s).
Therefore, to nullify the mutual info, we maximize H(c; s) until it is equal to H(c) + H(s).

$$\boxed{-H(c,s) = \frac{1}{N_X + N_Y} \sum_{i=1}^{N_X+N_Y} \log \frac{1}{N_X + N_Y} \sum_{j=1}^{N} \exp \frac{-||c_i - c_j||_2^2}{2\tau} \exp \frac{-||s_i - s_j||_2^2}{2\tau} \qquad (9)}$$

Experiments and results on simple datasets

Objective: **Design simple and controlled datasets to estimate the separation of the patterns.**

Figure 6: the Superimposed MNIST digits on CIFAR background dataset. Target images are CIFAR-10 images overlaid with an MNIST digit. Background images are CIFAR-10 images.

Experiments and results on simple datasets

Objective: **Results on MNIST digits watermarked on CIFAR-10 objects.**

Table 1: Digits on CIFAR-10 (B-ACC). Details in Sec. F.7.

	Digits		Objects		$\delta_{\text{TOT}} \downarrow$
	S ↑	C ↓	S ↓	C ↑	
cVAE	90.6		11.2		
ConVAE	86.2		10.6		
MM-cVAE	88.8		12.2		
SepVAE	90.6		10.6		
SepCLR-vCLUB sym	94.4		8.0		
SepCLR-vCLUB C → S	95.2		9.2		
SepCLR-vCLUB S → C	95.2		8.8		
SepCLR-vL1o sym	95.0		8.4		
SepCLR-vL1o C → S	94.0		10.0		
SepCLR-vL1o S → C	95.4		9.2		
SepCLR-vUB sym	94.6		8.2		
SepCLR-vUB C → S	92.8		7.8		
SepCLR-vUB S → C	96.6		8.6		
SepCLR-TC	95.2		10.2		
SepCLR-MMD	94.6		9.0		
SepCLR-no k-JEM	95.6		9.0		
SepCLR-k-MI	96.2		8.0		
SepCLR-k-JEM	$\boxed{96.2}$		$\boxed{10.4}$		
Best expected	100.0	10.0	10.0	100.0	0.0

Experiments and results on simple datasets

Objective: **Results on MNIST digits watermarked on CIFAR-10 objects.**

Table 1: Digits on CIFAR-10 (B-ACC). Details in Sec. F.7.

	Digits		Objects		$\delta_{TOT} \downarrow$
	S ↑	C ↓	S ↓	C ↑	
cVAE	90.6	23.0	11.2	33.4	90.2
ConVAE	86.2	21.0	10.6	35.6	89.8
MM-cVAE	88.8	19.6	12.2	32.0	93.6
SepVAE	90.6	17.8	10.6	36.6	81.2
SepCLR-vCLUB sym	94.4	18.0	8.0	14.6	97.0
SepCLR-vCLUB C → S	95.2	39.4	9.2	27.2	106.2
SepCLR-vCLUB S → C	95.2	57.0	8.8	31.8	118.8
SepCLR-vL1o sym	95.0	18.4	8.4	15.4	96.4
SepCLR-vL1o C → S	94.0	23.0	10.0	31.8	87.2
SepCLR-vL1o S → C	95.4	41.0	9.2	28.8	106.0
SepCLR-vUB sym	94.6	42.0	8.2	29.0	106.6
SepCLR-vUB C → S	92.8	23.4	7.8	22.6	95.8
SepCLR-vUB S → C	96.6	41.8	8.6	28.6	105.2
SepCLR-TC	95.2	68.6	10.2	24.2	139.4
SepCLR-MMD	94.6	21.2	9.0	62.2	53.4
SepCLR-no k-JEM	95.6	94.4	9.0	42.0	145.8
SepCLR-k-MI	96.2	19.8	8.0	65.8	45.8
SepCLR-k-JEM	96.2	11.0	10.4	73.2	**32.0**
Best expected	100.0	10.0	10.0	100.0	0.0

Experiments and results on a neuroimaging task

Objective: **Experiment on a neuroimaging application - Schizophrenia Disorders vs Healthy Controls setup.**

Datasets: **Neuroanatomical T1w brain MRIs** with (128x128x128) voxels pre-processed with Voxel Based Morphometry.

Common space : should capture **age and sex, or acquisition site.**

Table 3: Separate healthy from pathological variability in Schizophrenia disorder. Best in **bold**.

	Age MAE		Sex B-ACC		Site B-ACC	
	C $\downarrow$	S $\uparrow$	C $\uparrow$	S $\downarrow$	C $\uparrow$	S $\downarrow$
cVAE	6.43±0.18	7.27±0.25	75.06±3.48	74.99±2.15	65.12±4.06	59.62±5.42
ConVAE	6.40±0.26	7.46±0.18	74.45±1.80	72.72±1.32	60.42±3.67	54.46±2.46
MM-cVAE	6.55±0.18	7.10±0.34	72.80±3.95	72.15±2.47	63.24±1.41	56.69±9.84
SepVAE	**6.40±0.13**	**7.98±0.25**	74.19±1.81	72.61±2.19	63.89±2.16	44.10±5.78
SepCLR-k-JEM	6.64±0.21	7.72±0.45	**76.5±1.98**	**70.85±1.89**	**66.94±5.06**	**42.40±4.91**

Experiments and results on a neuroimaging task

Objective: **Experiment on a neuroimaging application - Schizophrenia Disorders vs Healthy Controls setup.**

Datasets: **Neuroanatomical T1w brain MRIs** with (128x128x128) voxels pre-processed with Voxel Based Morphometry.

Common space : should capture **age and sex, or acquisition site.**
Salient space : should capture **pathology-specific patterns** such as **SANS** (Negative Symptoms)
SAPS (Positive Symptoms).

Table 3: Separate healthy from pathological variability in Schizophrenia disorder. Best in **bold**.

	AGE MAE		SEX B-ACC		SITE B-ACC	
	C ↓	S ↑	C ↑	S ↓	C ↑	S ↓
CVAE	6.43±0.18	7.27±0.25	75.06±3.48	74.99±2.15	65.12±4.06	59.62±5.42
CONVAE	6.40±0.26	7.46±0.18	74.45±1.80	72.72±1.32	60.42±3.67	54.46±2.46
MM-CVAE	6.55±0.18	7.10±0.34	72.80±3.95	72.15±2.47	63.24±1.41	56.69±9.84
SEPVAE	**6.40±0.13**	**7.98±0.25**	74.19±1.81	72.61±2.19	63.89±2.16	44.10±5.78
SEPCLR-K-JEM	6.64±0.21	7.72±0.45	**76.5±1.98**	**70.85±1.89**	**66.94±5.06**	**42.40±4.91**
	SANS MAE		**SAPS MAE**		**DIAGNOSIS**	
	C ↑	S ↓	C ↑	S ↓	C ↓	S ↑
CVAE	5.89±0.67	4.35±0.26	4.65±0.34	2.98±0.18	60.66±2.63	68.24±5.42
CONVAE	6.17±0.45	3.95±0.28	4.50±0.37	2.76±0.18	61.85±2.60	58.53±4.87
MM-CVAE	6.78±0.54	4.92±0.58	4.52±0.33	3.16±0.05	64.25±2.98	70.94±4.08
SEPVAE	7.05±0.67	4.14±0.39	4.79±0.67	2.60±0.27	60.90±1.75	79.15±3.39
SEPCLR-K-JEM	**9.17±2.49**	**3.74±0.12**	**5.54±0.70**	**2.52±0.16**	**60.16±1.19**	**79.90±1.57**

Conclusion

Perspectives:

Develop a **Contrastive Analysis method** with **Diffusion Models**.

Investigate the use of **k-JEM** in other tasks (such as **multi modality content separation, debiasing, domain adaptation**).

Contact: robin.louiset@gmail.com

Currently ***looking for a post-doctoral position*** *in* ***France or in Switzerland****.*

Thanks to:

- *Pietro Gori:* *pietro.gori@telecom-paris.fr*
- *Edouard Duchesnay:* *edouard.duchesnay@cea.fr*
- *Antoine Grigis:* *antoine.grigis@cea.fr*

Thanks for listening!

Contact: robin.louiset@gmail.com

*Currently **looking for a post-doctoral position** in **France or in Switzerland**.*

Thanks to:

- *Pietro Gori: pietro.gori@telecom-paris.fr*
- *Edouard Duchesnay: edouard.duchesnay@cea.fr*
- *Antoine Grigis: antoine.grigis@cea.fr*