Sequential Disentanglement by Extracting Static Information From A Single Sequence Element

Nimrod Berman*, Ilan Naiman*, Idan Arbiv*, Gal Fadlon*, Omri Azencot (* Equal Contribution)

Problem Formulation

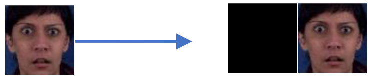
Given a dataset, $D = \{x^j_{1:T}\}_{j=1}^N$ compute a single static factor s and multiple dynamic components $d_{1:T}$.

Dynamics (Expression)

Static (Appearance)

Challenges

Mode Collapse

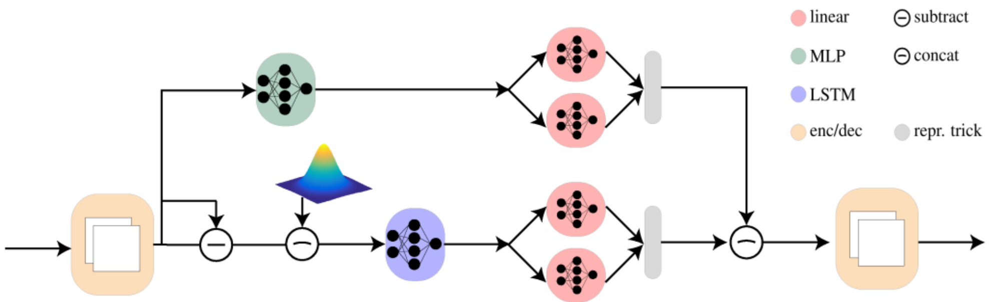


x_j s d_j

Hyper-parameter Tuning

Domain-Dependent Data Augmentation

Architecture and Objective



Consider the following distribution:

$$P = [p(s) \prod_{t=1}^{T} p(d_t | \; d_{<t}; \psi)] \cdot \prod_{t=1}^{T} p(x_t | \; s, d_t; \theta)$$

The posterior is approximated via:

$$q(s, d_{1:T} | x_{1:T}; \phi) = \; q_{\phi_s} \cdot q_{\phi_s} =: q(s | x_i; \phi_s) \cdot \prod_{t=1}^{T} q(d_t | \; d_{<t}, x_{\leq t}; \phi_d)$$

Loss Function:

$$\mathcal{L}_{\text{recon}} = \mathbb{E}_{s \sim q_{\phi_s}} [\mathbb{E}_{s \sim q_{\phi_s}} \log p(x_{2:T} | s, d_{2:T}; \theta)] + \; \alpha \mathbb{E}_{s \sim q_{\phi_s}} \log p(x_1 | s, d_1; \theta)]$$

$$\mathcal{L}_{\text{reg}} = \beta \; KL[q(s | x_1; \phi_s) \parallel p(s)] + \beta \; KL[q(d_{1:T} | x_{1:T}; \phi_s) \parallel p(d_{1:T}; \psi)]$$

$$\mathcal{L} \; = \max_{\theta, \phi, \psi} \mathbb{E}_{p_D} (\mathcal{L}_{\text{recon}} - \mathcal{L}_{\text{reg}})$$

Contributions

- A **novel** sequential disentanglement model for extracting static and dynamic factors.
- Our approach supports complex dynamics, **mitigates information leakage**, and simplifies training.
- **SOTA** quantitative and qualitative results on **challenging benchmarks** in comparison to recent strong baselines.

Results

Standard Benchmarks

Method	MUG Acc↑	IS↑	$H(y\|x)$↓	$H(y)$↑
MoCoGAN	63.12%	4.332	0.183	1.721
DSVAE	54.29%	3.608	0.374	1.657
R-WAE	71.25%	5.149	0.131	1.771
S3VAE	70.51%	5.136	0.135	1.760
SKD	77.45%	5.569	0.052	1.769
C-DSVAE	81.16%	5.341	0.092	1.775
SPYL	85.71%	5.548	0.066	1.779
Ours	**86.90%**	**5.598**	**0.041**	**1.782**

Method	PhysioNet ↑	Air Quality ↑
VAE	34.71 ± 0.23	27.17 ± 0.03
GP-VAE	42.47 ± 2.02	36.73 ± 1.40
C-DSVAE	32.54 ± 0.00	47.07 ± 1.20
GLR	38.93 ± 2.48	50.32 ± 3.87
SPYL	46.98 ± 3.04	57.93 ± 3.53
Ours w.o. loss	42.16 ± 0.104	50.14 ± 0.013
Ours w.o. sub	46.15 ± 0.014	56.11 ± 0.021
Ours w.o. both	41.42 ± 0.019	48.32 ± 0.102
Ours	$\mathbf{56.87 \pm 0.34}$	$\mathbf{65.87 \pm 0.01}$
RF	62.00 ± 2.10	62.43 ± 0.54

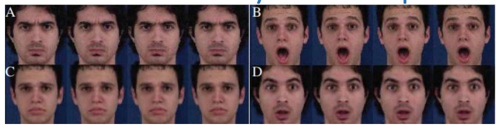
Method	static EER↓	dynamic EER ↑	Disentanglement Gap ↑
FHVAE	5.06%	22.77%	17.71%
DSVAE	5.64%	19.20%	13.56%
R-WAE	4.73%	23.41%	18.68%
S3VAE	5.02%	25.51%	20.49%
SKD	4.46%	26.78%	22.32%
C-DSVAE	4.03%	31.81%	27.78%
SPYL	**3.41%**	33.22%	29.81%
Ours	3.50%	**34.62%**	**31.11%**

Information Leakage

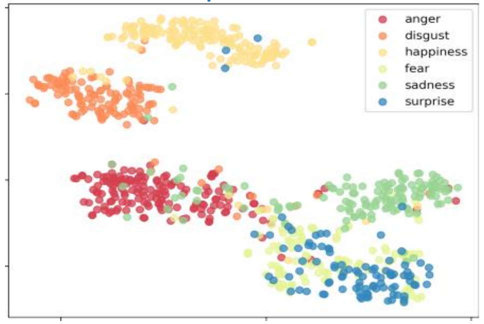
Classifier	Method	Static Features: Static Acc ↑	Dynamic Acc ↓	Leakage Gap ↑	Dynamic Features: Static Acc ↓	Dynamic Acc ↑	Leakage Gap ↑
Generation	random	-	16.66%	-	1.92%	-	-
	C-DSVAE	99.12%	29.9%	69.22%	3.75%	81.16%	77.41%
	SPYL	**99.45%**	27.65%	71.8%	3.63%	85.71%	82.08%
	Ours	99.42%	**20.85%**	**78.57%**	**2.89%**	**86.90%**	**84.01%**
	Ours w.o. loss	52.97%	18.93%	34.04%	2.58%	66.28%	63.70%
	Ours w.o. sub	34.41%	19.22%	15.19%	39.82%	83.31%	43.18%
	Ours w.o. both	10.94%	17.86%	6.92%	15.76%	71.95%	56.19%
Latent	random	-	16.66%	-	1.92%	-	-
	C-DSVAE	98.75%	76.25%	22.25%	26.25%	82.50%	56.25%
	SPYL	98.12%	68.75%	29.37%	**10.00%**	**85.62%**	**75.62%**
	Ours	**99.35%**	**45.06%**	**54.29%**	11.36%	85.51%	74.15%

Dynamic Generation

Static and Dynamic Swap



MUG Dynamic t-SNE



Air Quality Static t-SNE