Probabilistic Graphical Models for Image Analysis - Lecture 11

Stefan Bauer

30th November 2018

Max Planck ETH Center for Learning Systems

Overview

1. Evaluating Deep Representation Learning
2. Disentanglement
3. Robustness of Representations
4. Outlook

Evaluating Deep Representation Learning

The Problem - Metrics

How do we evaluate generative models?

- For tractable likelihood models: Evaluate generalization by reporting likelihoods on test data
- Proxy to likelihood might be available e.g. ELBO for VAEs.
- Visual Evaluation or e.g. using some metrics like Inception Scores, Frechet Inception Distance, Kernel Inception Distance based on heuristics like diversity, sharpness, similarities in feature representation ....

Difficult Problem (recall guest lecture) ....

Are GANs Created Equal? A Large-Scale Study

Mario Lucic* **Karol Kurach*** **Marcin Michalski** **Olivier Bousquet** **Sylvain Gelly**
Google Brain

How to evaluate latent representation

If available, we can evaluate the latent representation using the metric from a downstream task e.g. accuracy for semi-supervised learning.

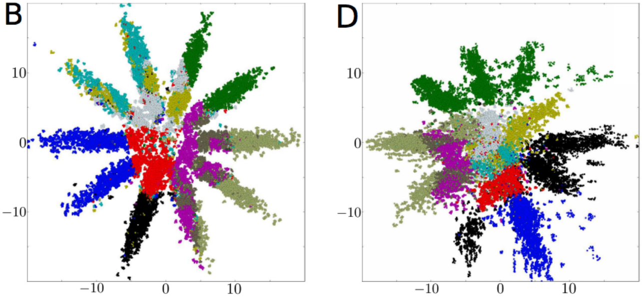
For unsupervised evaluations can be based on:

- Clustering using some attribute labels available e.g. color in MNIST.
- Compression
- ’Disentanglement’
- ....

Problem: At least partially a renaming of the problem i.e shifts the focus from how to evaluate representations to the evaluation of clusterings ...

Clustering*

Example: Cluster based on learned 2D representation, color denotes true labels.



Problem: How do we validate clustering? (Spring 2019: Prof. J. Buhmann *Statistical Learning Theory*)

*Makhzani et.al Adversarial Autoencoders 2016
https://arxiv.org/pdf/1511.05644.pdf

Disentanglement

Recall: Images to Torques

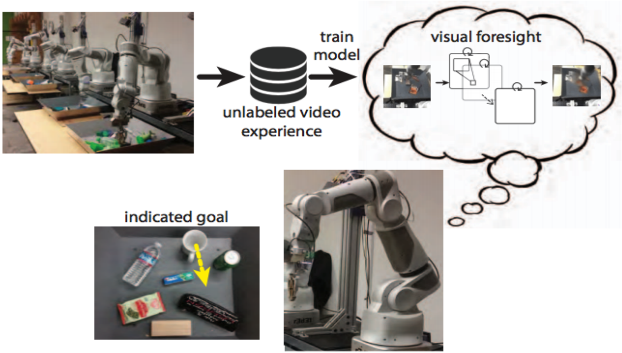
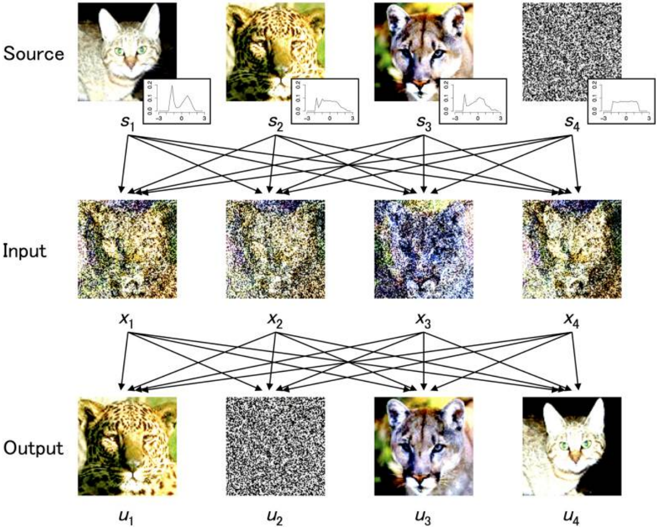


Fig. 1. Using our approach, a robot uses a learned predictive model of images, i.e. a visual imagination, to push objects to desired locations.

Finn and Levine, Deep Visual Foresight for Planning Robot Motion, ICRA 2017

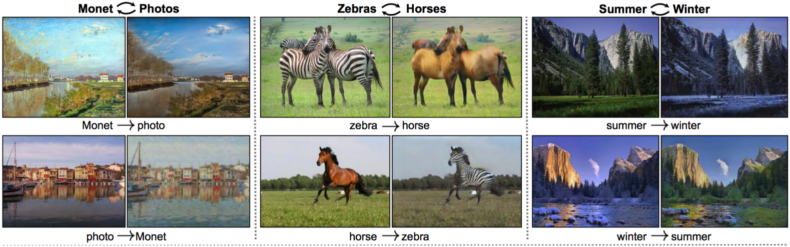
ICA*



*Isomura and Toyoizumi, A Local Learning Rule for Independent Component Analysis, Sci. Reports 2016

The Problem*

Example of State of the art:



Problem: Impressive results but would like to change factors on more granular level e.g. shape of mountain, color of rooftop, trees in background, ...

*Zhu, Park et.al Cycle GAN `https://arxiv.org/pdf/1703.10593.pdf`

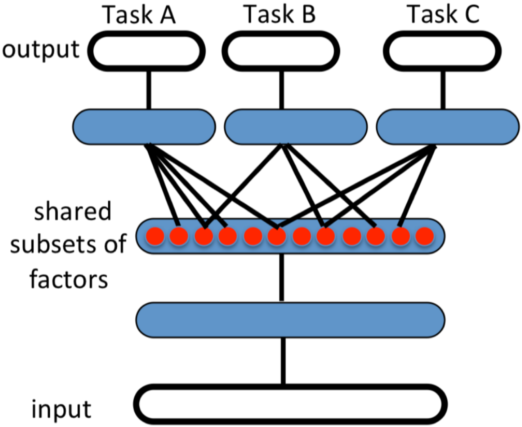
The Problem*

Representation Learning: A Review and New Perspectives

Yoshua Bengio[†], Aaron Courville, and Pascal Vincent[†]
Department of computer science and operations research, U. Montreal
† also, Canadian Institute for Advanced Research (CIFAR)

*Bengio et.al Representation Learning: A Review and New Perspectives
`https://arxiv.org/pdf/1206.5538.pdf`

The Problem*



Source: Bengio (2013)

*slides courtesy of Raphael Suter

Disentanglement

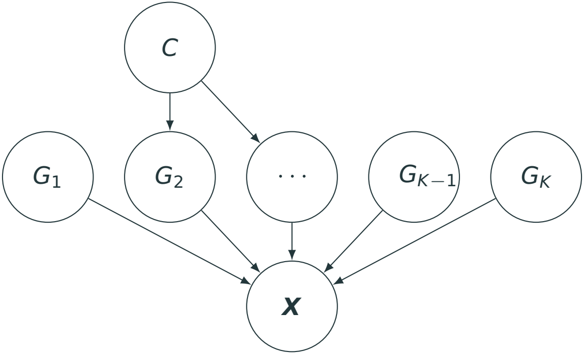
arbitrariness: $\hat{\boldsymbol{x}} = D(E(\boldsymbol{x})) = D(f(f^{-1}(E(\boldsymbol{x})))) = \tilde{D}(\tilde{E}(\boldsymbol{x}))$

Disentanglement $\iff$ splitting sources of variation

- Supervised: split known factors from unknowns
- Unsupervised: independence regularization, e.g.:
  - β-VAE: $D_{KL}\left(q_\phi(\boldsymbol{z}|\boldsymbol{x}) \| p(\boldsymbol{z})\right)$
  - FactorVAE, β-TCVAE: $\mathrm{TC}(\boldsymbol{z}) = D_{KL}\left(q(\boldsymbol{z}) \| \prod_i q(z_i)\right)$
  - DIP-VAE: factorize $q_\phi(\boldsymbol{z}) = \int q_\phi(\boldsymbol{z}|\boldsymbol{x}) p(\boldsymbol{x})\, d\boldsymbol{x}$

Problems: How to enforce ’disentanglement’ during training? How to find a trade-off between terms e.g. give up on reconstruction to what extent? (Recall - *Are all GANs created equal?*)

Causal Perspective



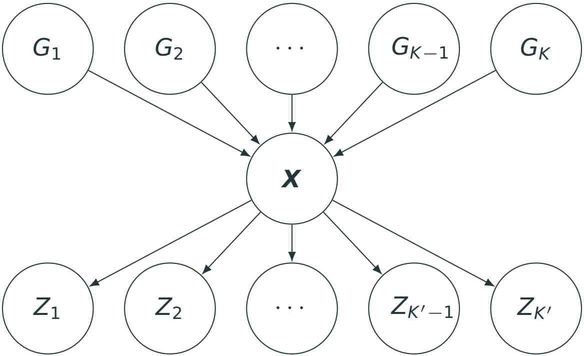
Disentangled Generative Factors

$$\boldsymbol{C} \leftarrow \boldsymbol{N}_c$$

$$G_i \leftarrow f_i(\boldsymbol{PA}_i^C, N_i), \quad \boldsymbol{PA}_i^C \subset \{C_1, \ldots, C_L\}, \quad i = 1, \ldots, K$$

$$\boldsymbol{X} \leftarrow g(\boldsymbol{G}, N_x)$$

Generative Factors

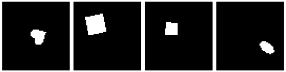


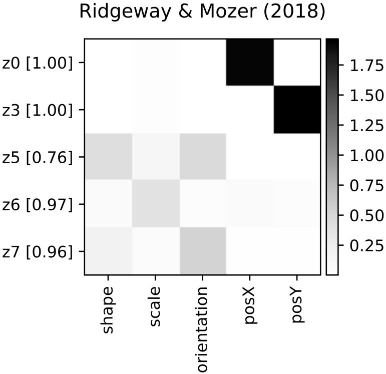
Feature Representation

Robustness of Representations

Information Based Validation

- Ground truth $\boldsymbol{G}$
- Mutual information $I(Z_i, G_j)$
- (or feature importance)
- Demand sparse rows





Interventional Robustness

Post Interventional Disagreement

$$PIDA(L|\boldsymbol{g}_I, \boldsymbol{g}_J^{\triangle}) := d\left(\mathbb{E}[\boldsymbol{Z}_L|\mathrm{do}(\boldsymbol{G}_I \leftarrow \boldsymbol{g}_I)], \mathbb{E}[\boldsymbol{Z}_L|\mathrm{do}(\boldsymbol{G}_I \leftarrow \boldsymbol{g}_I, \boldsymbol{G}_J \leftarrow \boldsymbol{g}_J^{\triangle})]\right)$$

Expected Maximum PIDA

$$EMPIDA(L|I,J) := \mathbb{E}_{\boldsymbol{g}_I}\left[\sup_{\boldsymbol{g}_J^{\triangle}} PIDA(L|\boldsymbol{g}_I, \boldsymbol{g}_J^{\triangle})\right]$$

Interventional Robustness Score

$$IRS(L|I,J) := 1 - \frac{EMPIDA(L|I,J)}{EMPIDA(L|\emptyset,\{1,\dots,K\})} \quad \in [0,1]$$

Interventional Robustness

Post Interventional Disagreement

$$PIDA(L|\boldsymbol{g}_I, \boldsymbol{g}_J^{\triangle}) := d\left(\mathbb{E}[\boldsymbol{Z}_L|\mathrm{do}(\boldsymbol{G}_I \leftarrow \boldsymbol{g}_I)], \mathbb{E}[\boldsymbol{Z}_L|\mathrm{do}(\boldsymbol{G}_I \leftarrow \boldsymbol{g}_I, \boldsymbol{G}_J \leftarrow \boldsymbol{g}_J^{\triangle})]\right)$$

Expected Maximum PIDA

$$EMPIDA(L|I,J) := \mathbb{E}_{\boldsymbol{g}_I}\left[\sup_{\boldsymbol{g}_J^{\triangle}} PIDA(L|\boldsymbol{g}_I, \boldsymbol{g}_J^{\triangle})\right]$$

Interventional Robustness Score

$$IRS(L|I,J) := 1 - \frac{EMPIDA(L|I,J)}{EMPIDA(L|\emptyset,\{1,\ldots,K\})} \quad \in [0,1]$$

Interventional Robustness

Post Interventional Disagreement

$$PIDA(L|\boldsymbol{g}_I, \boldsymbol{g}_J^{\triangle}) := d\left(\mathbb{E}[\boldsymbol{Z}_L|\mathrm{do}(\boldsymbol{G}_I \leftarrow \boldsymbol{g}_I)], \mathbb{E}[\boldsymbol{Z}_L|\mathrm{do}(\boldsymbol{G}_I \leftarrow \boldsymbol{g}_I, \boldsymbol{G}_J \leftarrow \boldsymbol{g}_J^{\triangle})]\right)$$

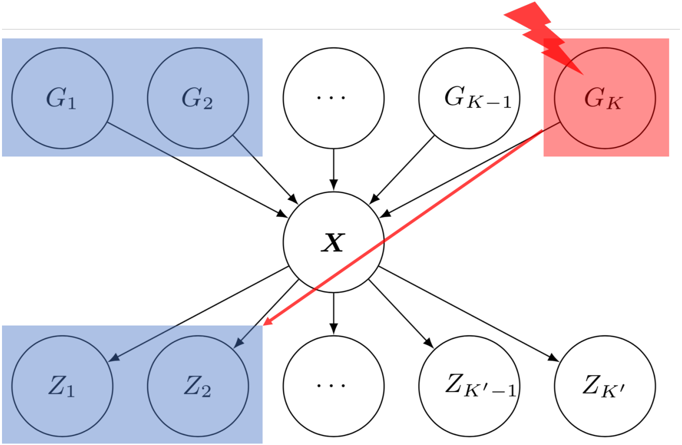
Expected Maximum PIDA

$$EMPIDA(L|I,J) := \mathbb{E}_{\boldsymbol{g}_I}\left[\sup_{\boldsymbol{g}_J^{\triangle}} PIDA(L|\boldsymbol{g}_I, \boldsymbol{g}_J^{\triangle})\right]$$

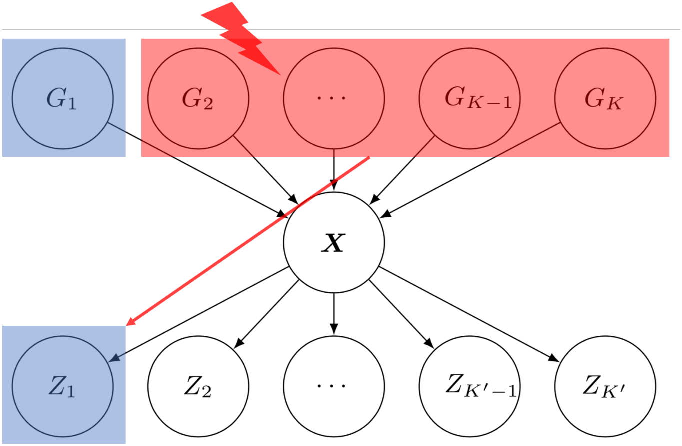
Interventional Robustness Score

$$IRS(L|I,J) := 1 - \frac{EMPIDA(L|I,J)}{EMPIDA(L|\emptyset,\{1,\ldots,K\})} \quad \in [0,1]$$

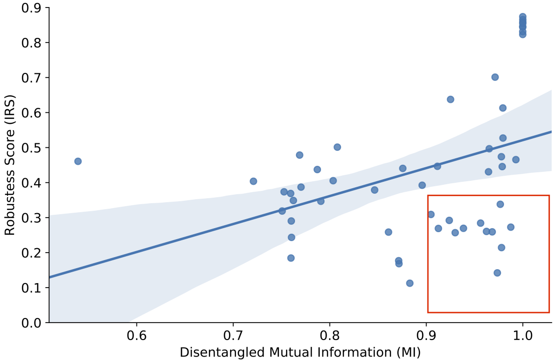
Interventional Robustness



Special Case: Disentanglement



Robustness as Complementary Viewpoint

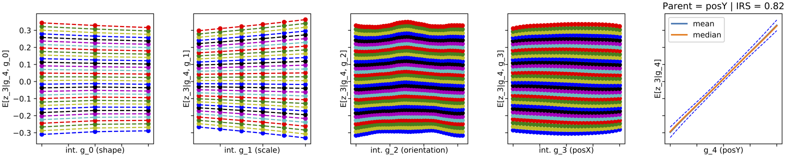


- Rare events
- Cumulative effects

Visualizing Robustness

Plot: $\mathbb{E}[Z_l | g_{i*}, \mathrm{do}(G_j \leftarrow g_j)]$

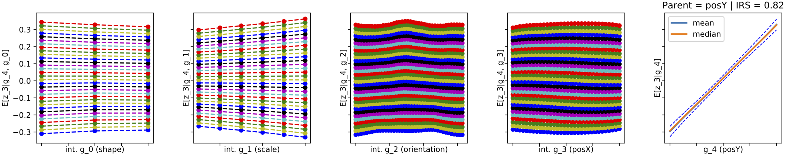
Robust Feature



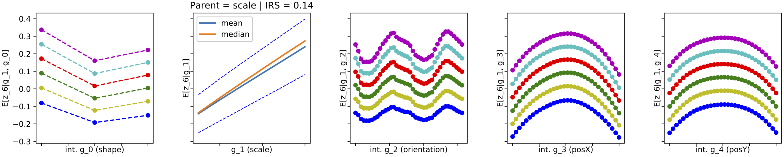
Visualizing Robustness

$$\text{Plot: } \mathbb{E}[Z_l | g_{i*}, \mathrm{do}(G_j \leftarrow g_j)]$$

Robust Feature



disentangled but not robust



Summary

Recall *Are all GANs created equal*:

Model	IRS	FI	MI	Info
VAE	0.33 (5)	0.23 (4)	0.90 (3)	0.82 (1)
Annealed β-VAE	0.57 (2)	0.35 (2)	0.86 (5)	0.79 (4)
DIP-VAE	0.43 (4)	0.39 (1)	0.89 (4)	0.82 (1)
FactorVAE	0.51 (3)	0.31 (3)	0.92 (1)	0.79 (4)
β-TCVAE	0.72 (1)	0.16 (5)	0.92 (1)	0.74 (5)

For more Details: Raphael Suter et.al. Interventional Robustness of Deep Latent Variable Models, `https://arxiv.org/pdf/1811.00007.pdf`

Problem

Challenging Common Assumptions in the Unsupervised Learning of Disentangled Representations:

In recent years, the interest in *unsupervised* learning of *disentangled* representations has significantly increased. The key assumption is that real-world data is generated by a few explanatory factors of variation and that these factors can be recovered by unsupervised learning algorithms. A large number of unsupervised learning approaches based on *auto-encoding* and quantitative evaluation metrics of disentanglement have been proposed; yet, the efficacy of the proposed approaches and utility of proposed notions of disentanglement has not been challenged in prior work. In this paper, we provide a sober look on recent progress in the field and challenge some common assumptions.

We first theoretically show that the unsupervised learning of disentangled representations is fundamentally impossible without inductive biases on both the models and the data. Then, we train more than 12 000 models covering the six most prominent methods, and evaluate them across six disentanglement metrics in a reproducible large-scale experimental study on seven different data sets. On the positive side, we observe that different methods successfully enforce properties "encouraged" by the corresponding losses. On the negative side, we observe that in our study (1) "good" hyperparameters seemingly cannot be identified without access to ground-truth labels, (2) good hyperparameters neither transfer across data sets nor across disentanglement metrics, and (3) that increased disentanglement does not seem to lead to a decreased sample complexity of learning for downstream tasks.

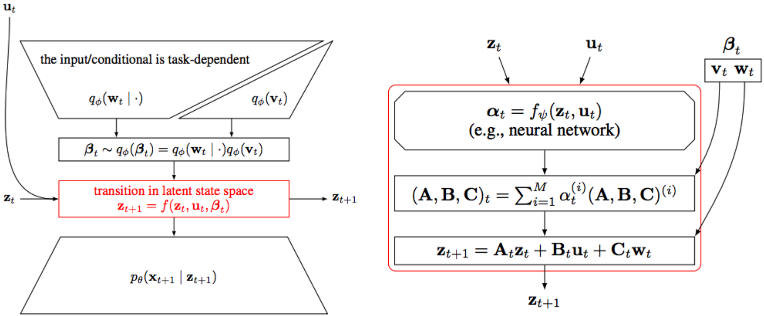
These results suggest that future work on disentanglement learning should be explicit about the role of inductive biases and (implicit) supervision, investigate concrete benefits of enforcing disentanglement of the learned representations, and consider a reproducible experimental setup covering several data sets.

For more Details: Locatello et.al.

`https://arxiv.org/pdf/1811.12359.pdf`

Outlook

Deep Bayes

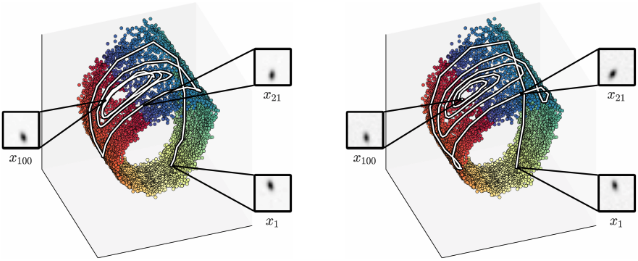


(a) General scheme for arbitrary transitions.

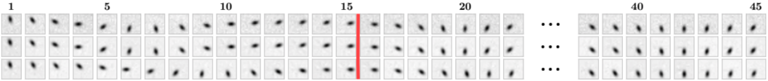
(b) One particular example of a latent transition: local linearity.

Karl et. al. Deep Variational Bayes Filters: Unsupervised Learning of State Space Models from Raw Data, ICLR 2017

Deep Bayes



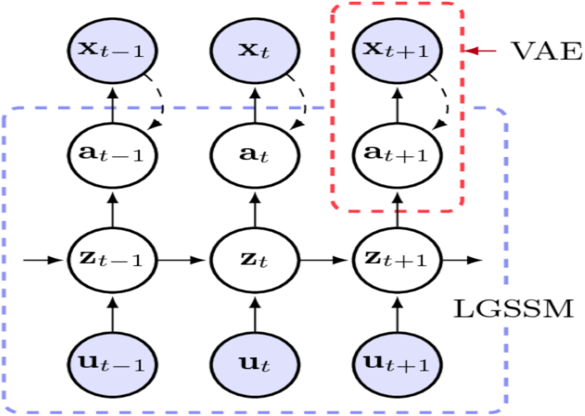
(a) Generative latent walk. (b) Reconstructive latent walk.



(c) Ground truth (top), reconstructions (middle), generative samples (bottom) from identical initial latent state.

Karl et. al. Deep Variational Bayes Filters: Unsupervised Learning of State Space Models from Raw Data, ICLR 2017

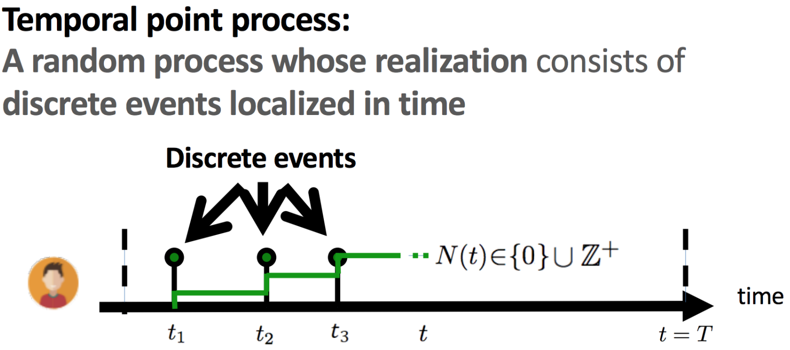
Kalman Variational Autoencoders



Fraccaro et. al. A Disentangled Recognition and Nonlinear Dynamics Model for Unsupervised Learning, NIPS 2017

Announcement - Guest Lecture

Isabel Valera (14th of December)



Plan for December

- **Next week** NO LECTURE (NIPS)
- Temporal Point Processes (14th December, guest lecture Isabel Valera)

Questions?