Deep Generative Models

Petra Poklukar

30% Seminar

Seminar outline

- Deep Generative Models
  - Definition & motivation
  - Shortcomings

- Current work
  - Work #1: Evaluation of likelihood estimates produced by Deep Generative Models
  - Work #2: Latent Space Roadmap for Visual Action Planning

- Future work
  - Short term projects
  - Long term vision

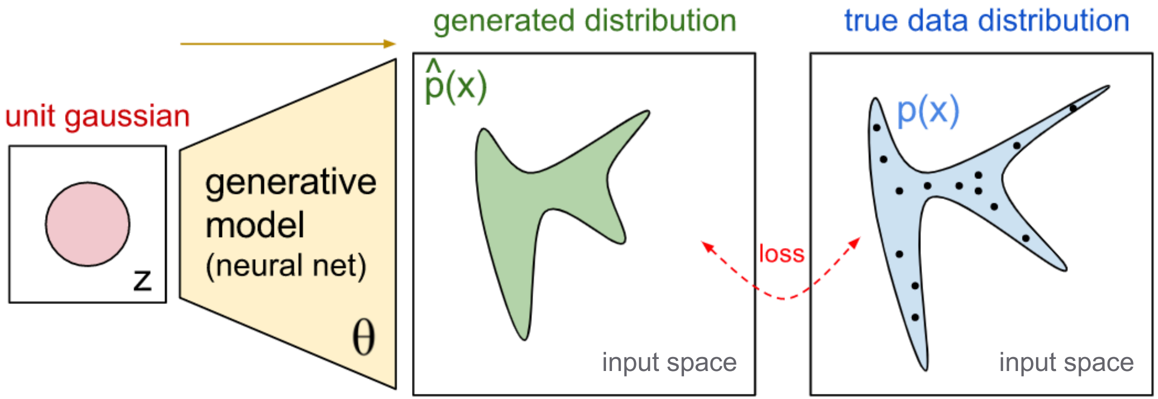
Deep Generative Models (DGMs)

- Definition & motivation
- Shortcomings

DGMs: definition

For an observable random variable X a **generative model** M captures its probability distribution $p(X)$.

When M is learned with a neural network we say that M is a **deep generative model.**



https://openai.com/blog/generative-models/

DGMs: advantages

1. **Density estimation**
   – $p_{model}(x_{test})$ for a given test point x_{test}

Anomaly detection

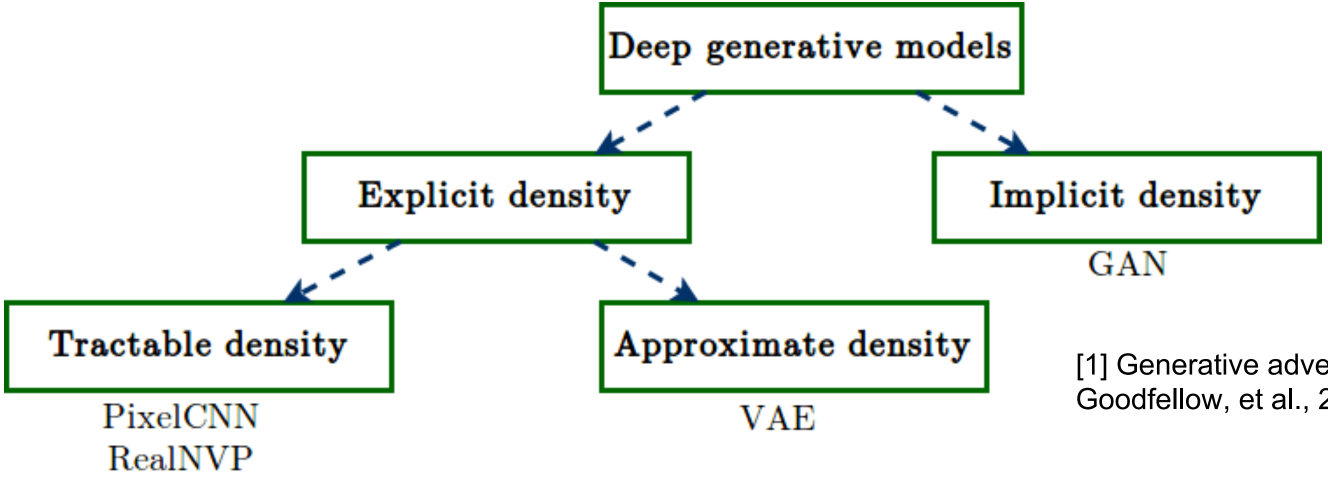
2. **Sampling**
   – generate $x_{new} \sim p_{model}(X)$

Planning, interpolation

3. **Unsupervised representation learning**
   – learn meaningful feature representation of an input point x

Dimensionality reduction

DGMs: taxonomy



[1] Generative adversarial nets, Goodfellow, et al., 2014.

[2] Auto-encoding variational bayes, Kingma, et al, 2013.

[3[Stochastic backpropagation and variational inference in deep latent gaussian models, Rezende, et al, 2014.

[4] Conditional image generation with pixelcnn decoders, Van den Oord, et al., 2016.

[5] Density estimation using realNVP, Dinh, et al, 2017

DGMs: the big picture

Density estimation	Sampling	Representation learning
Quality of $p_{model}(x)$	Sample quality	Latent space structure
Flows	**GANs**	**VAEs**

[6] Adaptive Density Estimation for Generative Models, Lucas et al. NeurIPS 2019

[7] Glow: Generative flow with invertible 1x1 convolutions, Kingma et al. NeurIPS 2018.

[8] Manifold-Valued Image Generation with Wasserstein Generative Adversarial Nets, Huang et al. *AAAI* 2019.

[9] Neural discrete representation learning, van den Oord et al. NIPS 2017

[10] Improved precision and recall metric for assessing generative models, Kynkäänniemi et al. NeurIPS 2019.

[11] Continuous Hierarchical Representations with Poincaré Variational Auto-Encoders, Emile, et al. NeurIPS 2019.

[12] Residual flows for invertible generative modelling, Chen et al. NeurIPS 2019

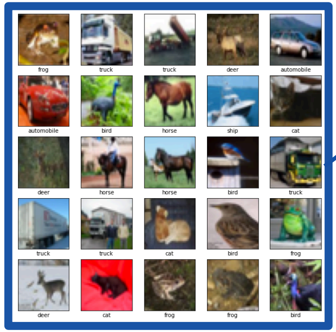
[13] Explicitly disentangling image content from translation and rotation with spatial-VAE, Bepler et al. NeurIPS 2019.

Current work

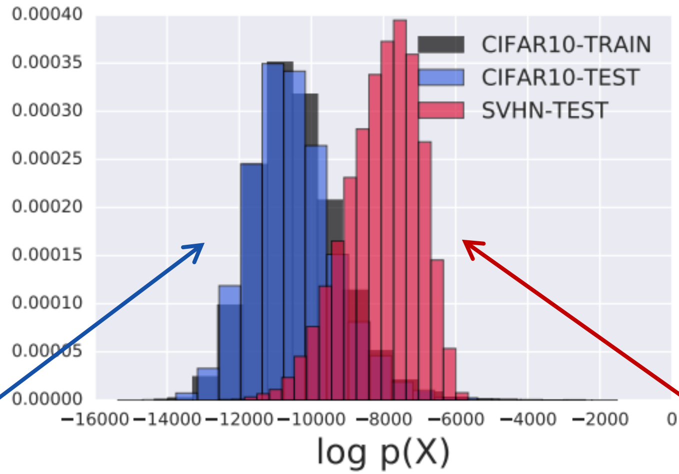
Work #1: Detecting unknown unknowns with DGMs (with Judith)

Quality of $p_{model}(x)$

In-Distribution (ID)



DGM likelihood estimates



(d) **VAE: CIFAR-10 vs SVHN**

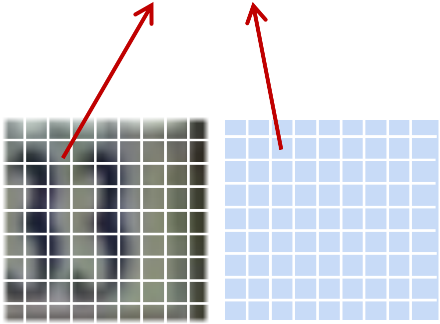
Out-of-Distribution (OOD)

[14] Do deep generative models know what they don't know, Nalisnick et al. (2019)

Why are DGMs overconfident?

Hypothesis (simplified): Pixel iid assumption is too *local:*

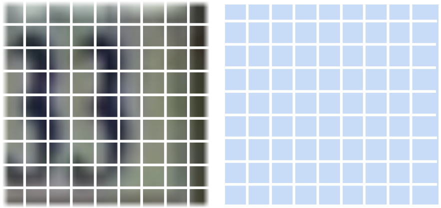
$$p_{model}(x_{test} | \theta_{model}) = \prod_{d=1}^{n_{pixels}} p_{model}\left(x_{test}^{d} \middle| \theta_{model}^{d}\right)$$

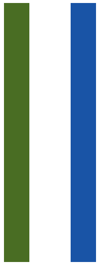


Idea: compare to something *global.*

Pixel level vs image level evaluation during test time

$$p_{model}(x_{test}|\theta_{model}) = \overbrace{\left[\prod_{d=1}^{n_{pixels}} p_{model}(x_{test}^{d}|\theta_{model}^{d})\right]}^{\text{pixel level}} \overbrace{p(f_{x_{test}}|f_{\theta})}^{\text{feature level}}$$





Two workshop papers:

1. Modelling assumptions and evaluation schemes: On the assessment of deep latent variable models, J. Bütepage, P. Poklukar, D. Kragic, *CVPR & ICML workshop 2018*
2. Seeing the whole picture instead of a single point: Self-supervised likelihood learning, P. Poklukar, J. Bütepage, D. Kragic, *AABI 2019*

Work #2

Work #2: Planning in latent space (with Michael & Martina)

Cornell University

We
the Simons

arXiv.org > cs > arXiv:2003.08974

Search...

Help | Advanced

Computer Science > Robotics

Latent Space Roadmap for Visual Action Planning of Deformable and Rigid Object Manipulation

Martina Lippi, Petra Poklukar, Michael C. Welle, Anastasiia Varava, Hang Yin, Alessandro Marino, Danica Kragic

(Submitted on 19 Mar 2020)

We present a framework for visual action planning of complex manipulation tasks with high-dimensional state spaces such as manipulation of deformable objects. Planning is performed in a low-dimensional latent state space that embeds images. We define and implement a Latent Space Roadmap (LSR) which is a graph-based structure that globally captures the latent system dynamics. Our framework consists of two main components: a Visual Foresight Module (VFM) that generates a visual plan as a sequence of images, and an Action Proposal Network (APN) that predicts the actions between them. We show the effectiveness of the method on a simulated box stacking task as well as a T-shirt folding task performed with a real robot.

Comments:	Project website: this https URL
Subjects:	**Robotics (cs.RO)**; Machine Learning (cs.LG)
Cite as:	arXiv:2003.08974 [cs.RO]
	(or arXiv:2003.08974v1 [cs.RO] for this version)

Bibliographic data

[Enable Bibex (What is Bibex?)]

Submitted to IROS2020

Sample quality

Latent space structure

Latent Space Roadmap (LSR): problem setting

Input: a system with high dimensional states

Goal: visual action planning

start state

goal state

Latent Space Roadmap (LSR): problem setting

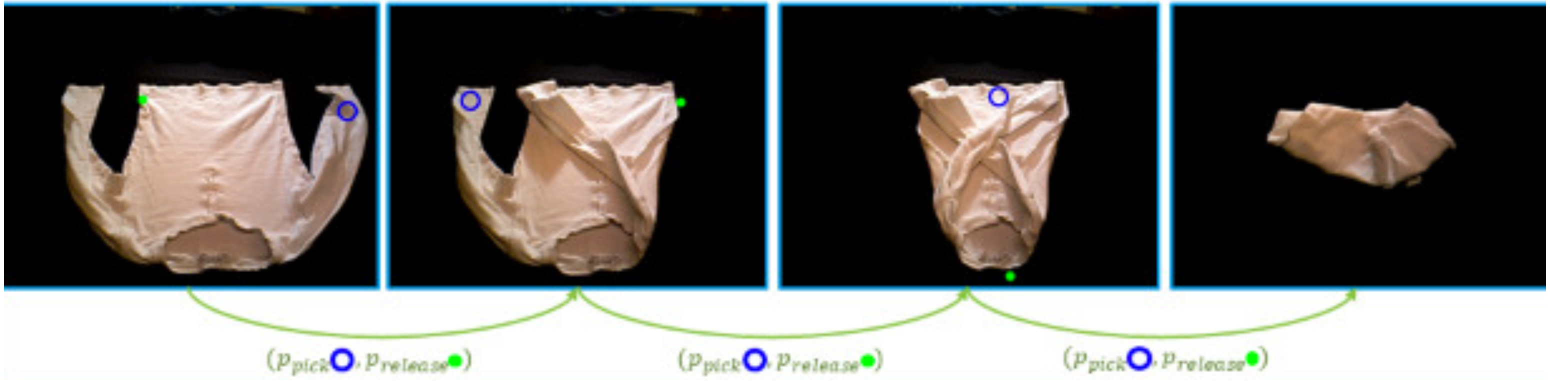
Input: a system with high dimensional states

Goal: visual action planning

How to achieve it: learn a low-dimensional representation of its structure and dynamics.

start state

goal state



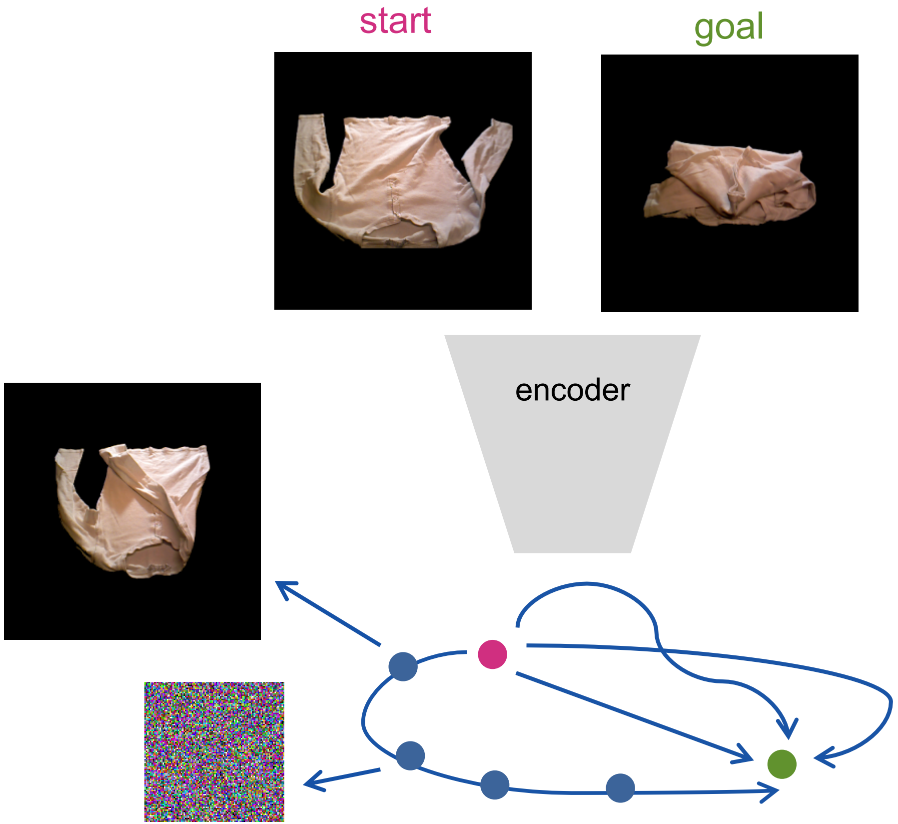
LSR: challenges

Goal: low-dimensional representation of a system used for visual action planning.

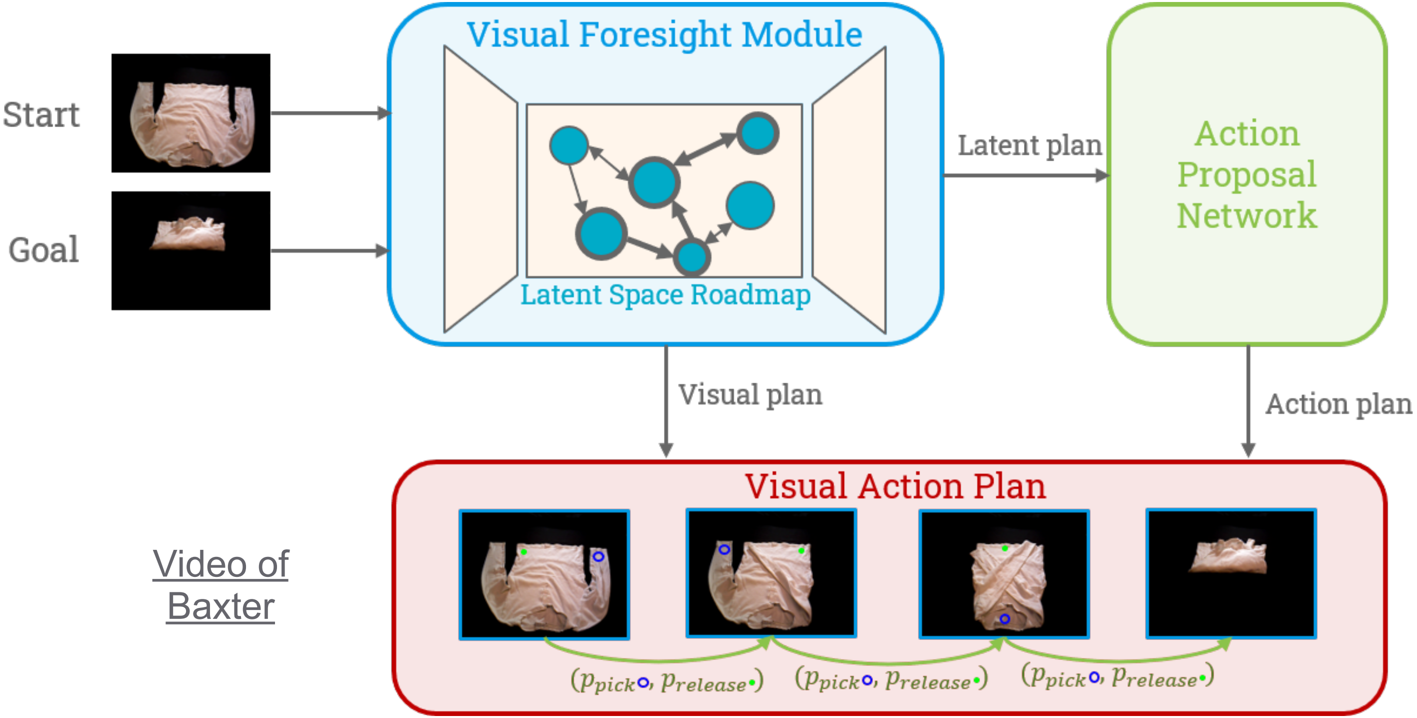
Idea: latent variable model

One of the problems:

- How to find a path consisting of meaningful states?



LSR: overview



Video of Baxter

Future work

Future projects and long-term goal

Integrate domain knowledge in DGMs

Distant future

Quality of $p_{model}(x)$

Sample quality

Latent space structure

Immediate future

Geometric approach for OODs detection (with Anastasiia)

LSR continuation (with Michael & Martina)

RL policy training using DGMs (with Ali)

Topologically constrained generation process in DGMs (with Zehang)

Thanks

References

[1] Goodfellow, Ian, et al. "Generative adversarial nets." *Advances in neural information processing systems*. 2014.

[2] Kingma, Diederik P., and Max Welling. "Auto-encoding variational bayes." *arXiv preprint arXiv:1312.6114* (2013).

[3[Rezende, Danilo Jimenez, Shakir Mohamed, and Daan Wierstra. "Stochastic backpropagation and variational inference in deep latent gaussian models." *International Conference on Machine Learning*. Vol. 2. 2014.

[4] Van den Oord, Aaron, et al. "Conditional image generation with pixelcnn decoders." *Advances in neural information processing systems*. 2016.

[5] Dinh, Laurent, Jascha Sohl-Dickstein, and Samy Bengio. "Density estimation using real nvp." *ICLR* (2017).

[6] Lucas, Thomas, et al. "Adaptive Density Estimation for Generative Models." *Advances in Neural Information Processing Systems*. 2019.

[7] Kingma, Durk P., and Prafulla Dhariwal. "Glow: Generative flow with invertible 1x1 convolutions." *Advances in Neural Information Processing Systems*. 2018.

[8] Huang, Zhiwu, Jiqing Wu, and Luc Van Gool. "Manifold-Valued Image Generation with Wasserstein Generative Adversarial Nets." *Proceedings of the AAAI Conference on Artificial Intelligence*. Vol. 33. 2019.

[9] van den Oord, Aaron, and Oriol Vinyals. "Neural discrete representation learning." *Advances in Neural Information Processing Systems,* 2017.

References

[10] Kynkäänniemi, Tuomas, et al. "Improved precision and recall metric for assessing generative models." *Advances in Neural Information Processing Systems*. 2019.

[11] Mathieu, Emile, et al. "Continuous Hierarchical Representations with Poincaré Variational Auto-Encoders." *Advances in neural information processing systems*. 2019.

[12] Chen, Tian Qi, et al. "Residual flows for invertible generative modeling." *Advances in Neural Information Processing Systems*. 2019

[13] Bepler, Tristan, et al. "Explicitly disentangling image content from translation and rotation with spatial-VAE." *Advances in Neural Information Processing Systems*. 2019.

[14] Nalisnick, Eric, et al. "Do deep generative models know what they don't know?." *ICLR* (2019).