ETH
Eidgenössische Technische Hochschule Zürich
Swiss Federal Institute of Technology Zurich

Distributed Computing

Prof. R. Wattenhofer

Quantifying E ects of Design Decisions in VAE-based Unsupervised Disentanglement Learning

Disentanglement Learning is at the forefront of unsupervised learning, as disentangled representations of data are thought to improve generalization, interpretability, and performance in downstream tasks. It works on the assumption of the existence of low-dimensional data generating factors for high-dimensional data and tries to recover these factors in an unsupervised fashion. Unfortunately, this is not possible without the use of any inductive biases [2]. Di erent examples of such biases can be found in successful architectures for disentanglement learning. The simplest example is the use of a higher KL-Divergence weight in the loss function of a -VAE [1].

In this thesis, we are interested in quantifying the e ects some of the inductive biases have on the learned representations. We want to rigorously evaluate whether some common assumptions in the eld hold true and analyze the consequences for the learning of disentangled representations with variational auto-encoders.

Requirements: Strong motivation, programming skills, and basic knowledge of machine and deep learning.

Interested? Please contact us for more details!

z1 ---------- z10

Contact

- Benjamin Estermann: besterma@ethz.ch, ETZ G60.1
- Peter Belcak: pbelcak@ethz.ch, ETZ G61.3

References

[1] Irina Higgins et al. \beta-vae: Learning basic visual concepts with a constrained variational framework". In: (2016).

[2] Francesco Locatello et al. \Challenging common assumptions in the unsupervised learning of disentangled representations". In: *international conference on machine learning*. PMLR. 2019, pp. 4114{4124.