Quantifying and Learning Linear Symmetry-Based Disentanglement

*Loek Tonnaer, *Luis A. Pérez Rey, Vlado Menkovski, Mike Holenderski, Jacobus W. Portegies

* Equal Contribution

Motivation

- Disentanglement is **important**, but there is no agreed-upon definition
- Higgins et al. (2018) provide a formal definition:

  Linear Symmetry-Based Disentanglement (LSBD)
- Based on modelling **symmetries** of the real world, using group theory

Irina Higgins, et al. (2018) *"Towards a Definition of Disentangled Representations"*.

Motivation

- Disentanglement is **important**, but there is no agreed-upon definition
- Higgins et al. (2018) provide a formal definition:

  Linear Symmetry-Based Disentanglement (LSBD)
- Based on modelling **symmetries** of the real world, using group theory
- Observations from real world -> data

Irina Higgins, et al. (2018) *“Towards a Definition of Disentangled Representations”*.

Motivation

- Disentanglement is **important**, but there is no agreed-upon definition
- Higgins et al. (2018) provide a formal definition:

  Linear Symmetry-Based Disentanglement (LSBD)
- Based on modelling **symmetries** of the real world, using group theory
- Observations from real world -> data
- Transformations in real world -> variation in data

Irina Higgins, et al. (2018) *"Towards a Definition of Disentangled Representations"*.

Challenges

1. There is no **metric** to quantify the level of LSBD in arbitrary representations

   - Crucial for evaluation of LSBD methods, benchmarks

Challenges

1. There is no **metric** to quantify the level of LSBD in arbitrary representations

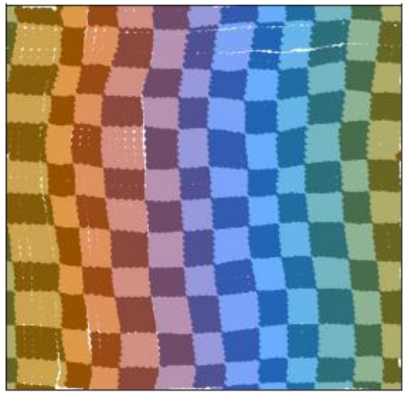
    - Crucial for evaluation of LSBD methods, benchmarks

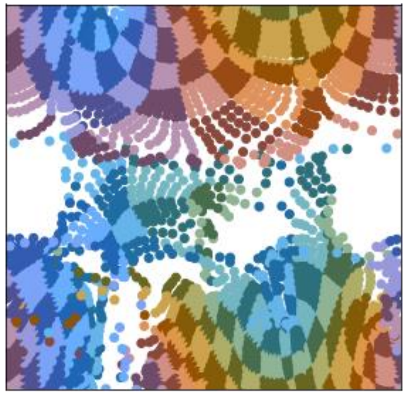
2. The LSBD definition defines no **method** to obtain LSBD representations

    - Some methods exist, but lack formal evaluation due to missing metric

Our contributions

$\mathcal{D}_{LSBD} = 0.013$ $\mathcal{D}_{LSBD} = 1.179$



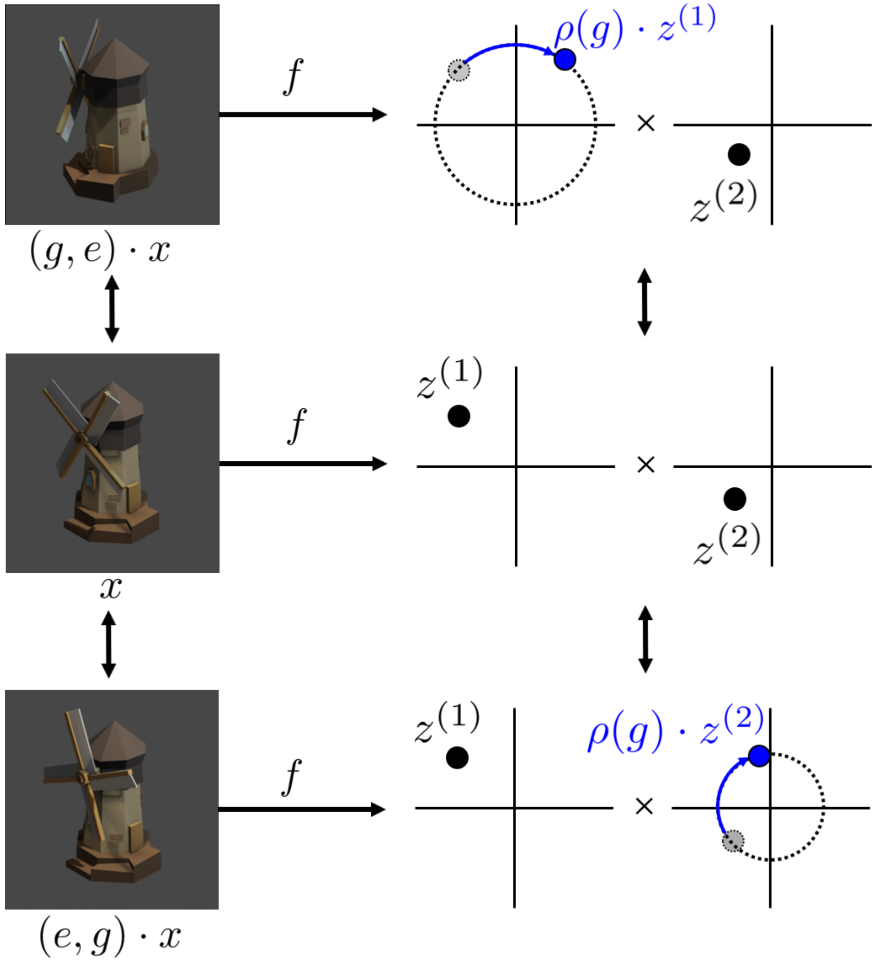


1. $\boldsymbol{\mathcal{D}_{LSBD}}$: a general metric that quantifies LSBD
   - Follows the formal LSBD definition
   - Works for any encoding and group structure (in theory)
   - Practical implementation for common subgroup: SO(2)

Our contributions

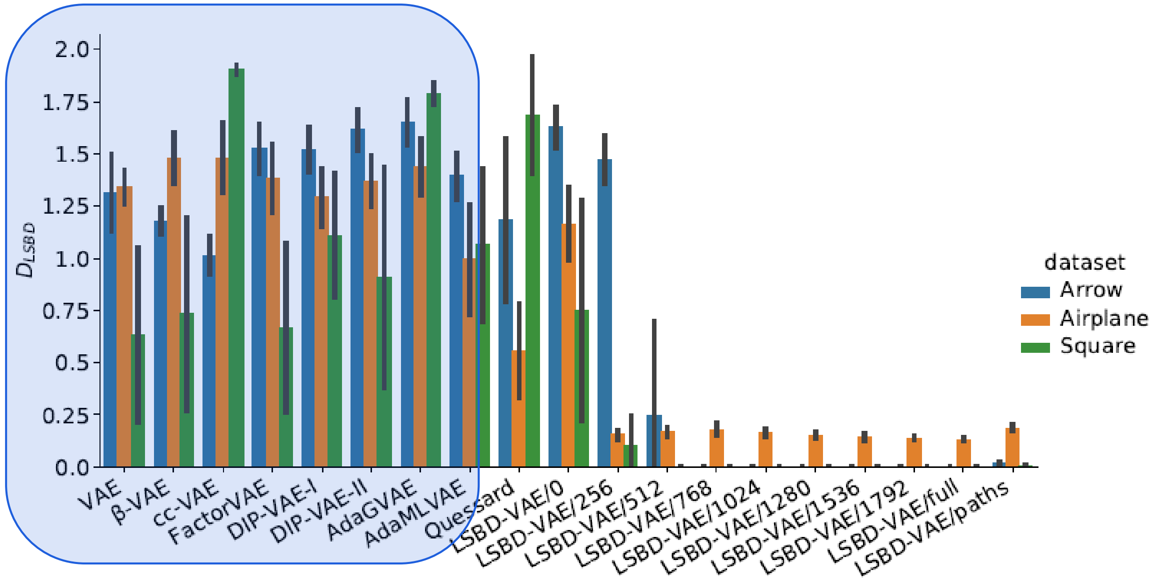
1. $\boldsymbol{\mathcal{D}_{LSBD}}$**:** a general metric that quantifies LSBD
- Follows the formal LSBD definition
- Works for any encoding and group structure (in theory)
- Practical implementation for common subgroup: SO(2)

2. LSBD-VAE: a method to learn LSBD representations
- Idea: use $\mathcal{D}_{LSBD}$ as additional loss component in ΔVAE,
  - (ΔVAE = VAE with suitable latent topology)
- Requires some assumptions (expert knowledge) to ensure computability

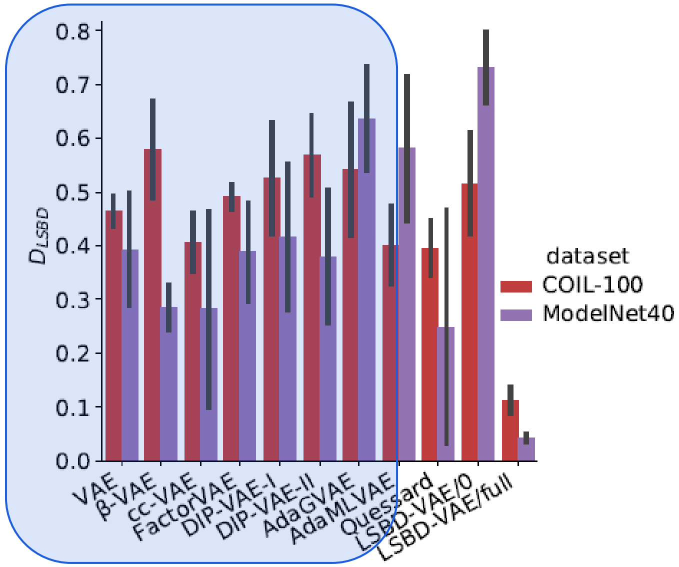


Results

1. Traditional disentanglement methods don’t learn LSBD representations



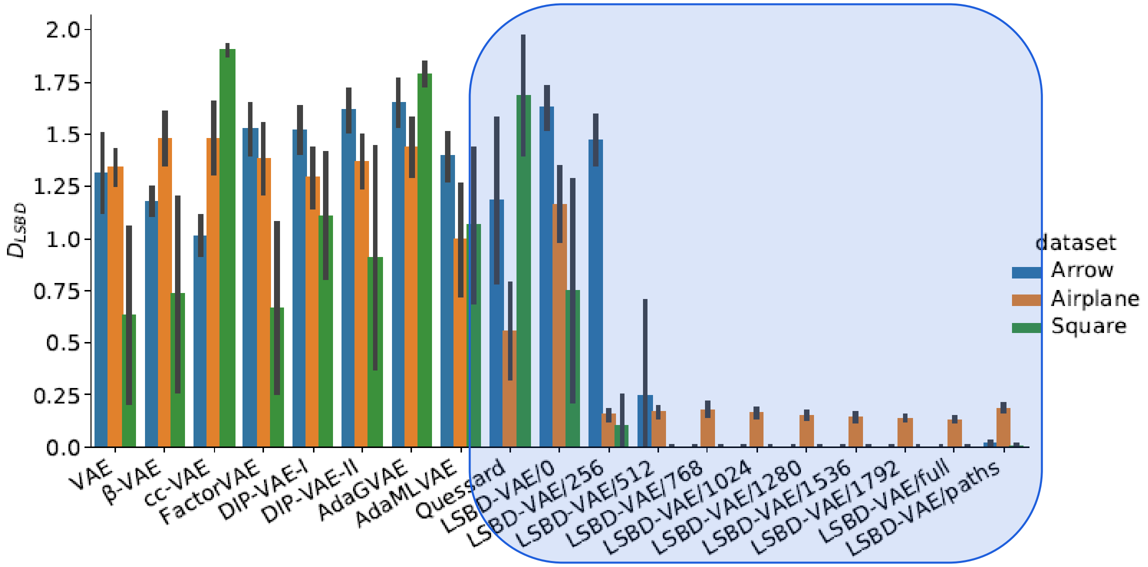
(a) Datasets with $\mathrm{SO}(2) \times \mathrm{SO}(2)$ symmetries



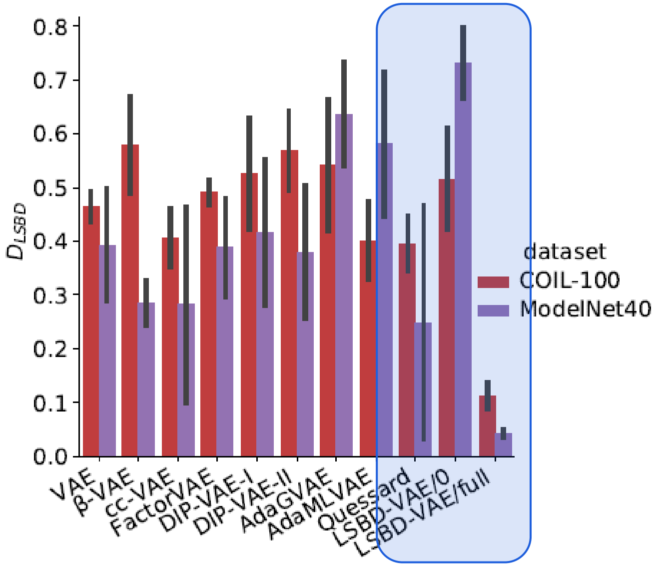
(b) Datasets with $\mathrm{SO}(2)$ and non-symmetric variation

Results

1. Traditional disentanglement methods don't learn LSBD representations
2. Our method and other LSBD methods *can* learn LSBD representations, with limited supervision on transformations



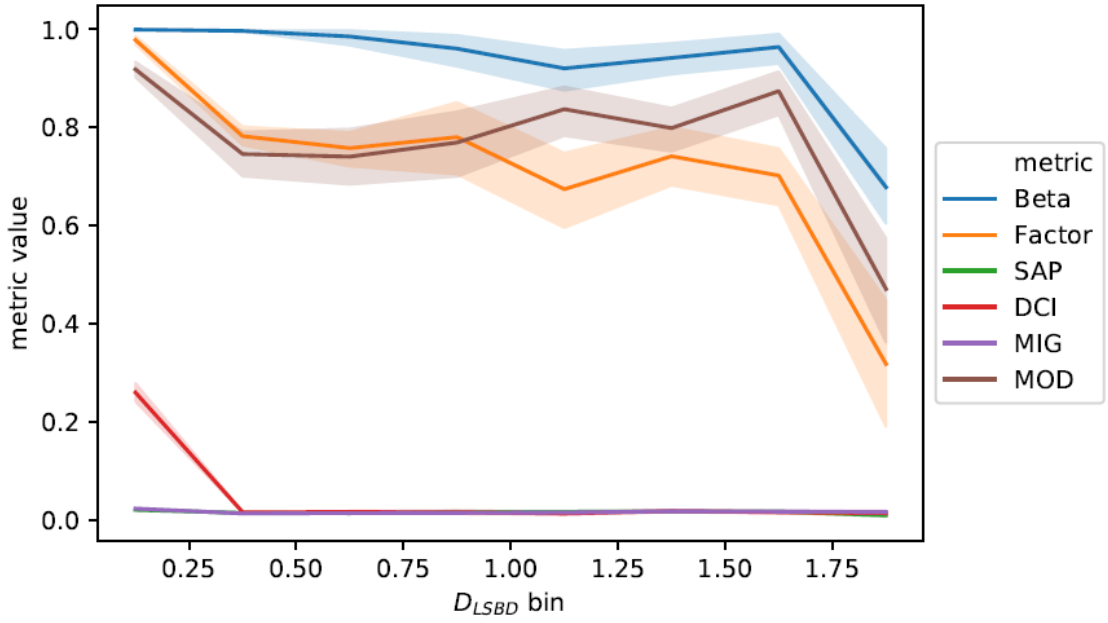
(a) Datasets with $\mathrm{SO}(2) \times \mathrm{SO}(2)$ symmetries



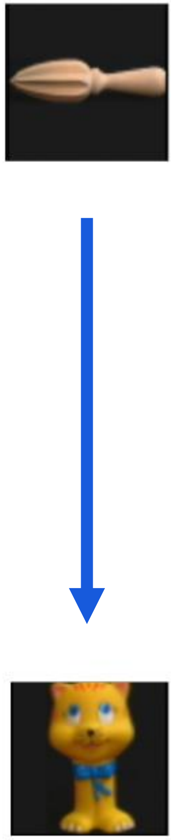
(b) Datasets with $\mathrm{SO}(2)$ and non-symmetric variation

Results

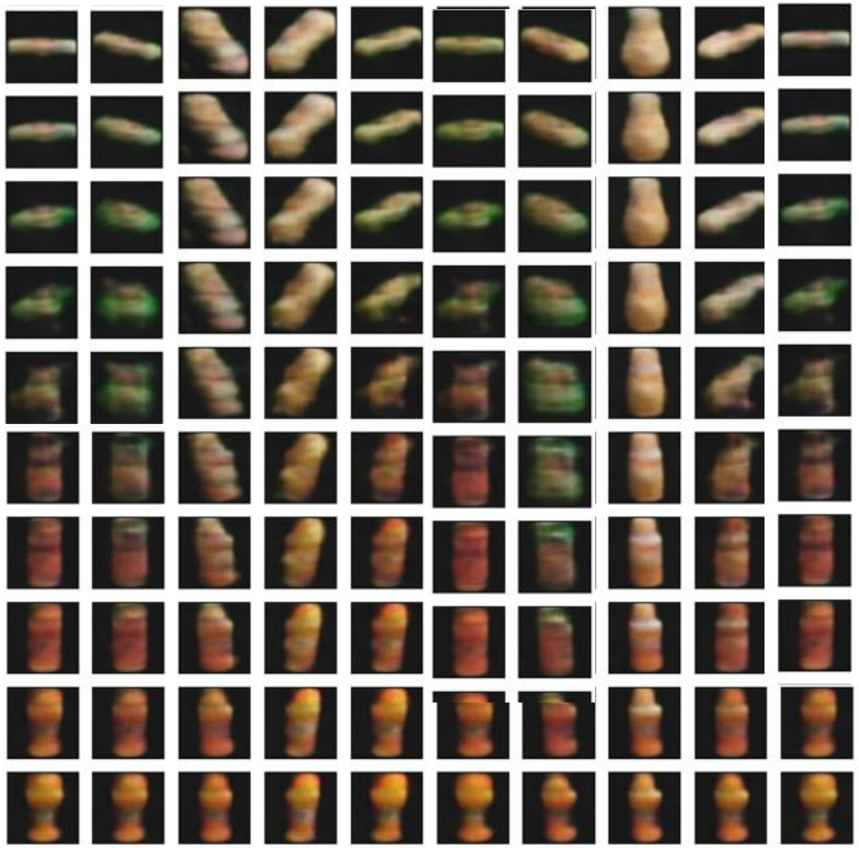
1. Traditional disentanglement methods don't learn LSBD representations
2. LSBD-VAE and other LSBD methods *can* learn LSBD representations, with limited supervision on transformations
3. LSBD representations also satisfy previous disentanglement notions (but not vice versa)



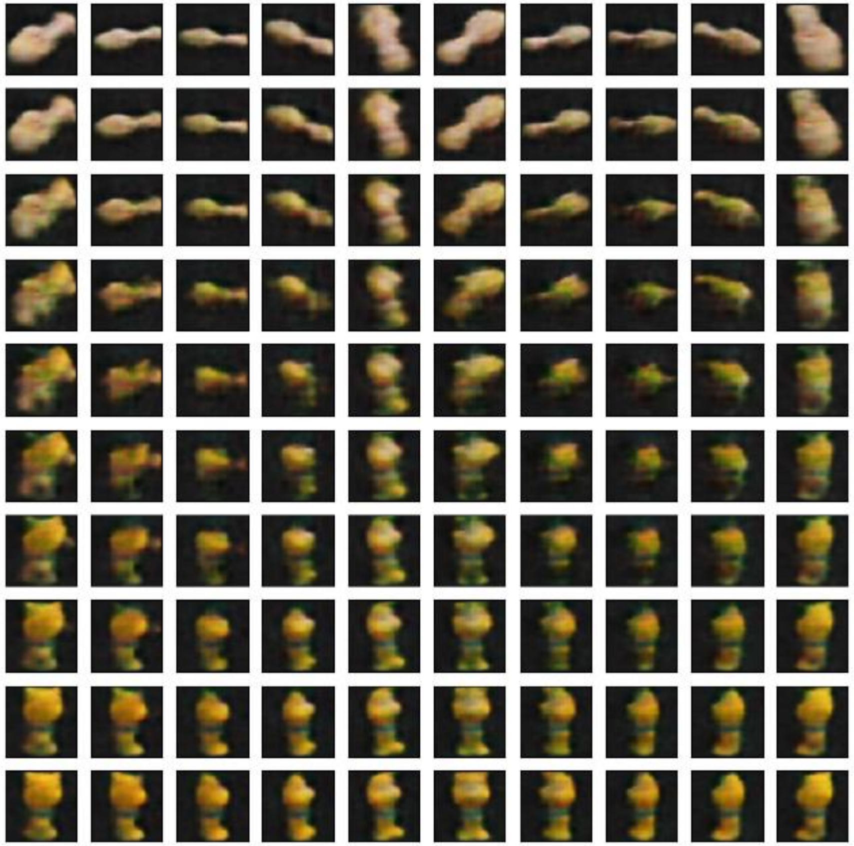
Results



β –VAE

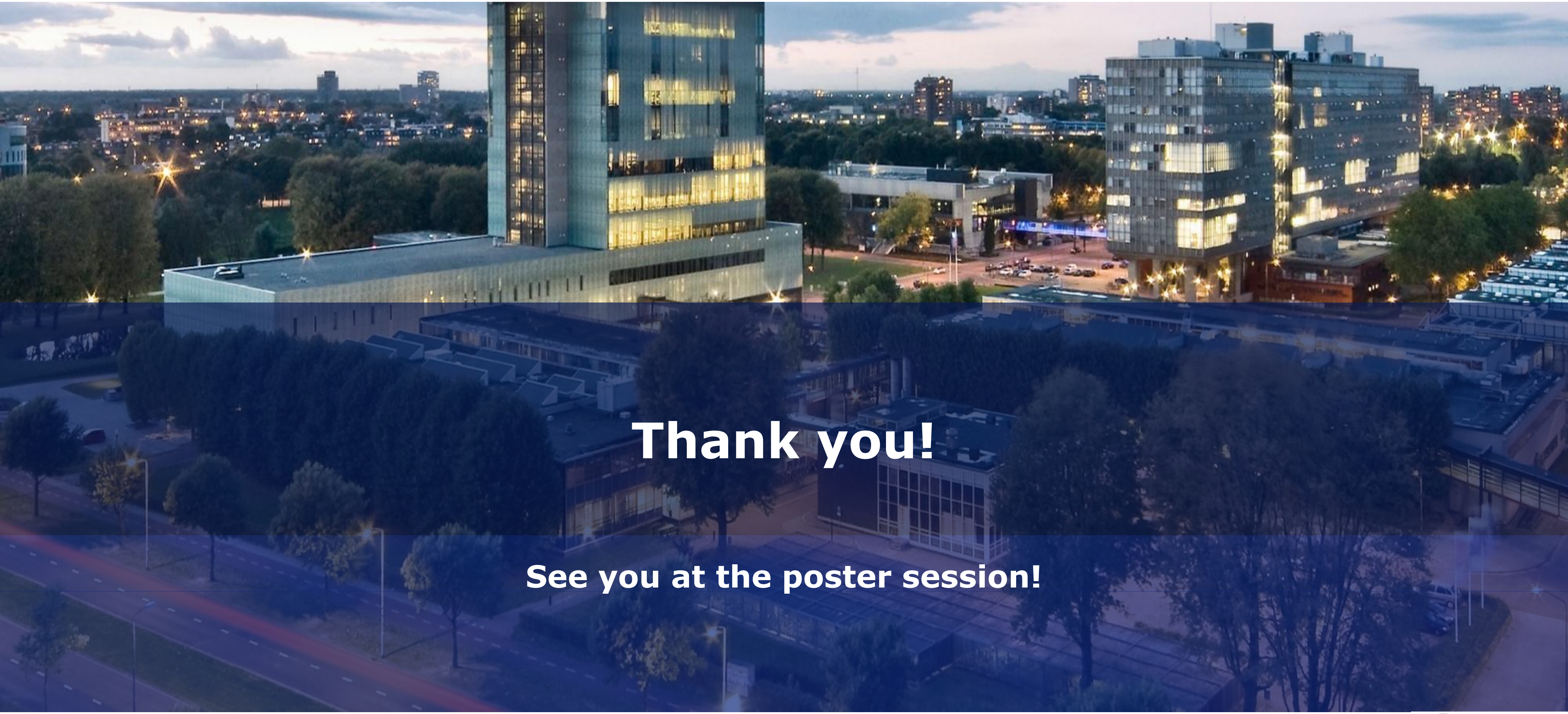


Ours



Conclusion

- We have devised a metric ($\mathcal{D}_{LSBD}$) and method (LSBD-VAE)
- We have shown that our method can learn LSBD representations with little supervision
- We have confirmed that traditional disentanglement models do not optimize for LSBD
- We demonstrate that LSBD satisfies previous notions of disentanglement, but also captures other desirable properties.



Thank you!

See you at the poster session!