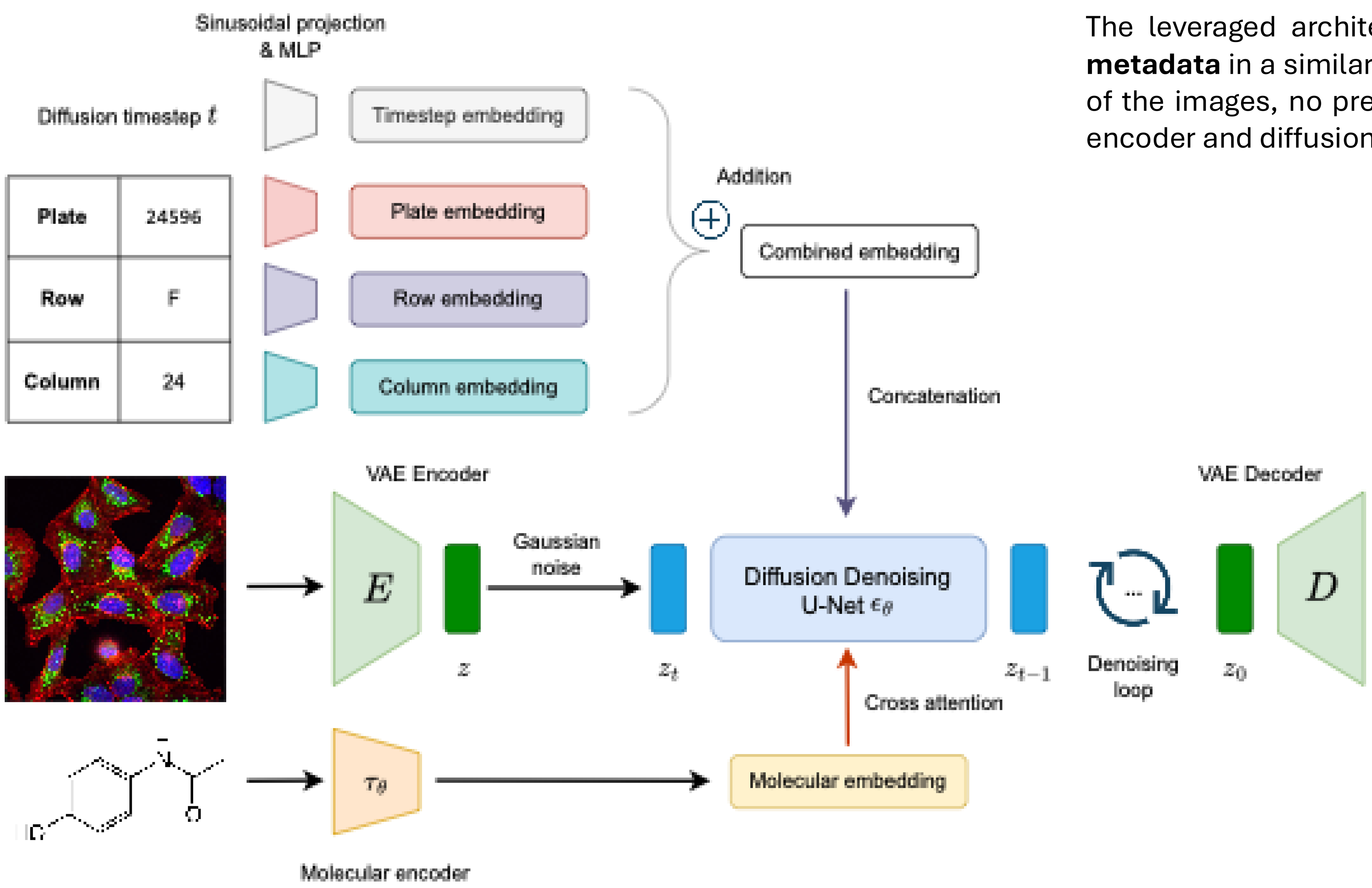
Diffusion models for predicting cell morphology

Télio Cropsal*, Rocío Mercado
AI Laboratory for Molecular Engineering
Computer Science And Engineering Department
Chalmers University of Technology, University of Gothenburg
***{telio@chalmers.se}**

Motivation & Research Goals

Predicting cell morphology under different perturbations can reduce the number of costly experiments and improve our understanding of the interaction between cells, molecules and genes. With the rise of **image-based high-throughput screening**, large-scale datasets have been created, enabling the use of deep learning methods and generative models. However, current approaches are either conditioned on the perturbations or the experimental settings but not the two of them at the same time. Moreover, they provide limited control to the user and the problem of **disentanglement** between biological effects and experimental effects is rarely addressed. The goal of this study is to showcase the utility of **latent diffusion models** to alleviate these issues.

Methods



The leveraged architecture is a **latent diffusion model** [1] **handling metadata** in a similar fashion as DiffusionSAT [2]. Because of the nature of the images, no pre-trained models are used for the variational auto-encoder and diffusion denoising U-Net.

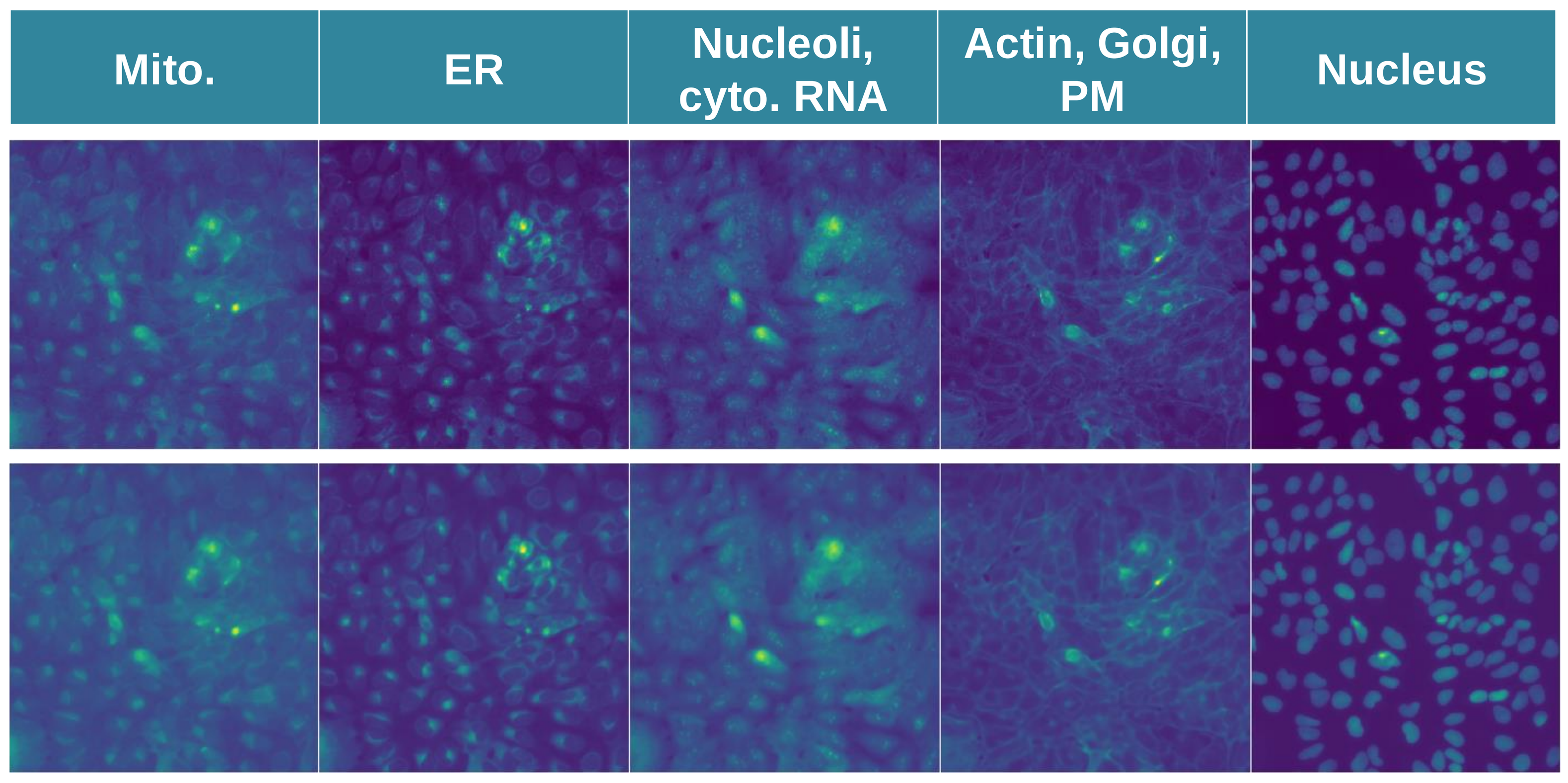
Different molecular encoders will be evaluated.

The model need not only be trained on perturbations [3][4] or experimental effects [5] and it can be easily extended for a neg-con to pos-con [4] task or Brightfield-2-CP [6] with ControlNet [7].

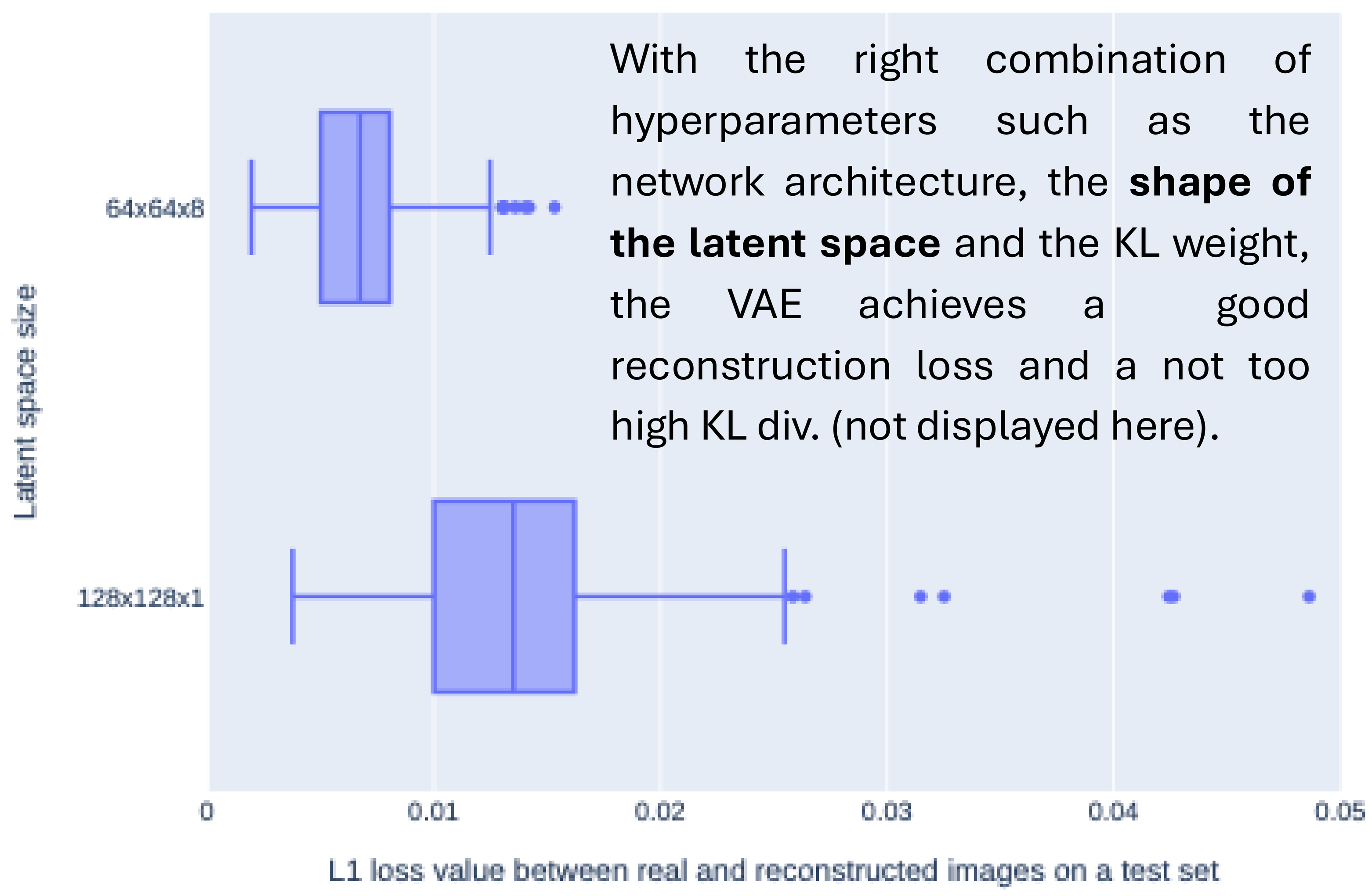
It is possible to evaluate the impact of biological and experimental effects over generated images, hence it is feasible **at inference to increase disentanglement** [8].

This approach will be compared to Mol2Image [3] with the same dataset, splits and benchmark. The datasets consists of **284K 5-channels 512x512 images** collected **from 10.5K molecular interventions**.

Evaluation of the VAE



With the right combination of hyperparameters such as the network architecture, the **shape of the latent space** and the KL weight, the VAE achieves a good reconstruction loss and a not too high KL div. (not displayed here).



References

[1] Rombach, Robin et al. “High-Resolution Image Synthesis with Latent Diffusion Models.”
[2] Khanna, Samar et al. “DiffusionSat: A Generative Foundation Model for Satellite Imagery.”
[3] Yang, Karren D. et al. “Mol2Image: Improved Conditional Flow Models for Molecule to Image Synthesis.”
[4] Palma, Alessandro De et al. “Predicting cell morphological responses to perturbations using generative modeling.”
[5] Pernice, Wolfgang M. et al. “Out of Distribution Generalization via Interventional Style Transfer in Single-Cell Microscopy.”
[6] Cross-Zamirski, Jan Oscar et al. “Class-Guided Image-to-Image Diffusion: Cell Painting from Brightfield Images with Class Labels.”
[7] Zhang, Lvmin et al. “Adding Conditional Control to Text-to-Image Diffusion Models.”
[8] Scimeca, Luca et al. “Leveraging Diffusion Disentangled Representations to Mitigate Shortcuts in Underspecified Visual Tasks.”

This work is generously funded by the Wallenberg Autonomous Systems and Software Program (WASP) and Chalmers University of Technology,.

WALLENBERG AI, AUTONOMOUS SYSTEMS AND SOFTWARE PROGRAM