HELDENKOMBINAT
MACHINE LEARNING

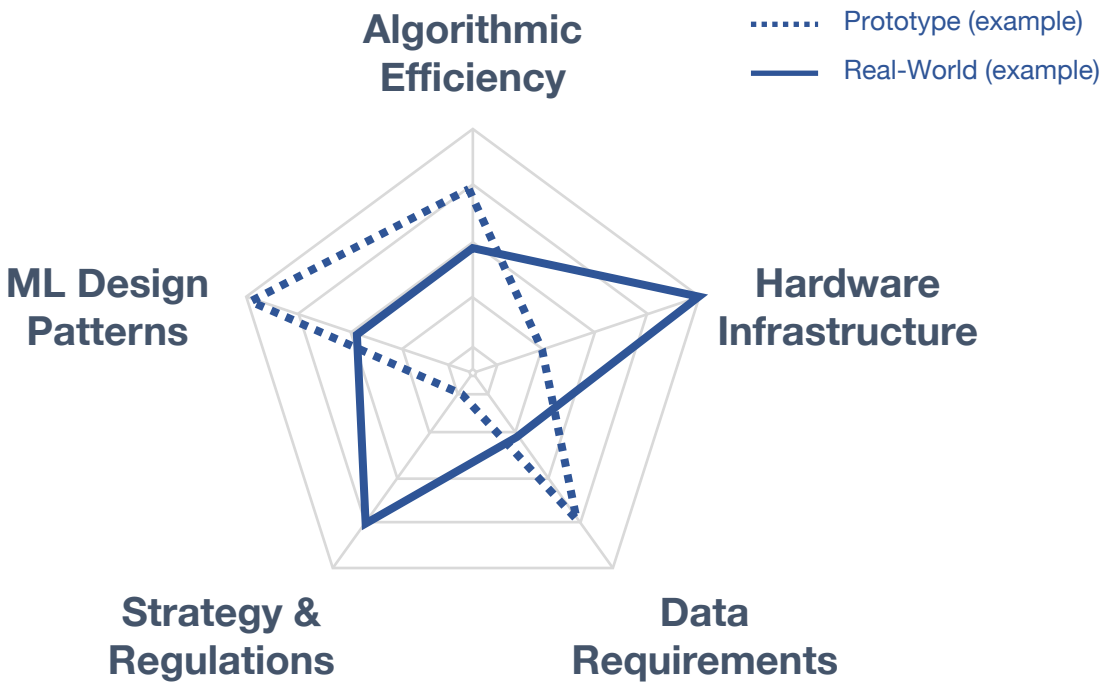
Generative models in the arts:
Exploring creative possibilities with underspecified objectives

Jorge Davila-Chacon, PhD

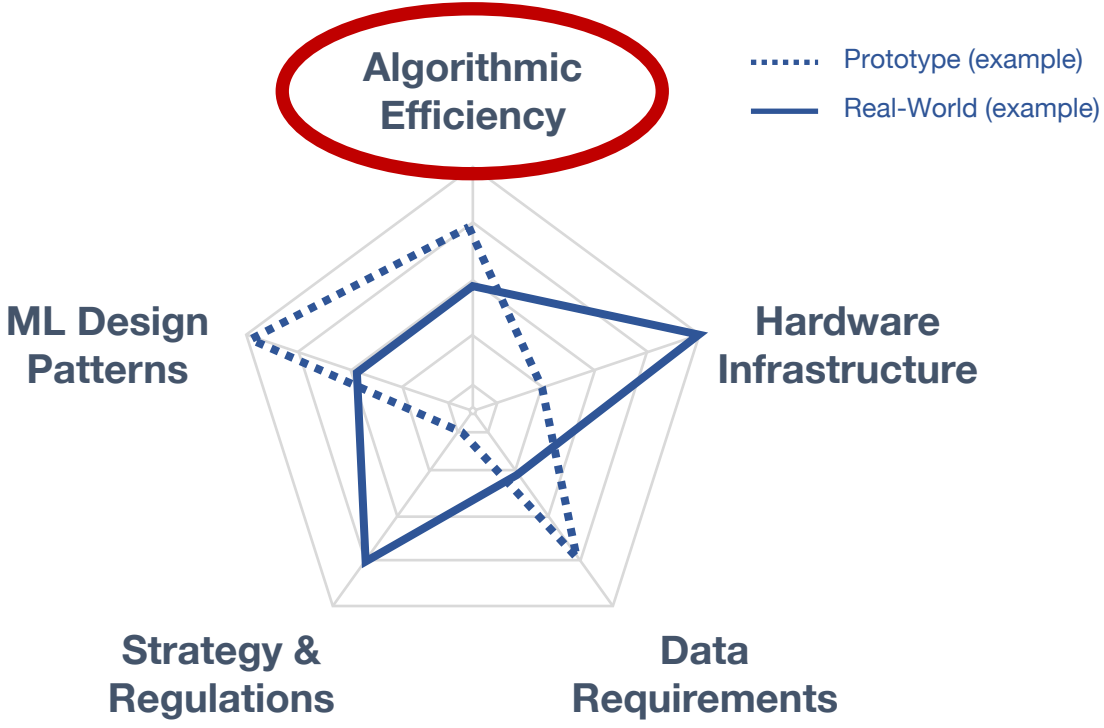
Co-founder and CTO
Heldenkombinat Technologies GmbH

ARIC Brown-Bag Sessions
19.01.2021

5 Dimensions of ML Projects



5 Dimensions of ML Projects



Instrumental Convergence

Starting point

VAE

Constraints

- Increase resolution as much as possible.
- Allow the reconstruction of video input.
- Images should be aesthetically intersting, the focus is not on the accuracy of the image reconstruction.
- Discover ways for exploring the latent space and control image features.
- Training data and features are not fixed in advance.

Architecture

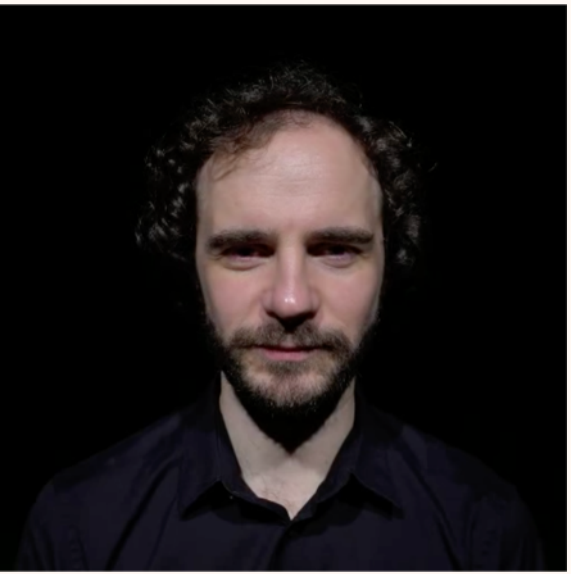
- VAEs [1] produce a smoother, more continuous, latent space than the good old AEs [2] and can take images as input.

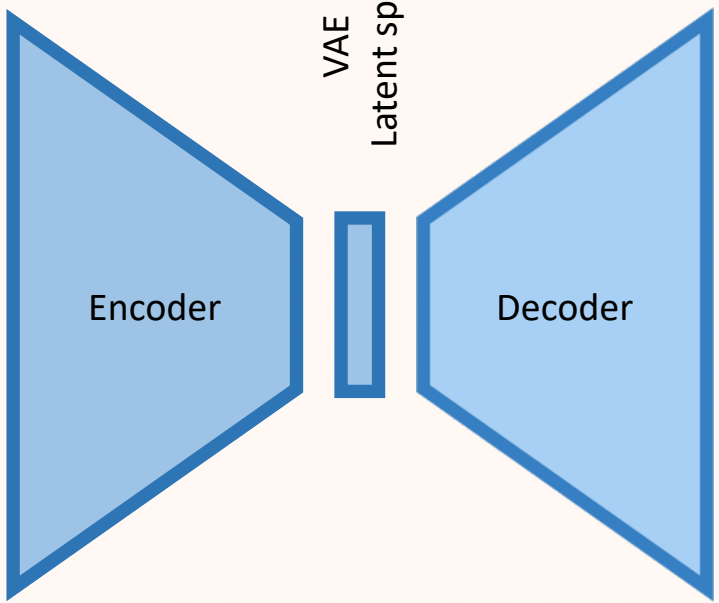
Which GAN variant should we choose?

- DCGANs [3] (the first GAN using CNNs): Capture more high-frequency components in the images. They are unstable and data-hungry.
- WGAN [4]: iImproves stability during learning (the algorithm converges more often to a good optimum).
- WGAN-GP [5]: reduces the amount of data needed to train the network.
- Pix2Pix [6]: Can produce interpolations and domain-transfer. More data-hungry.
- CycleGANs [7]: Perform style-transfer without the need for lots of paired translations between both domains.

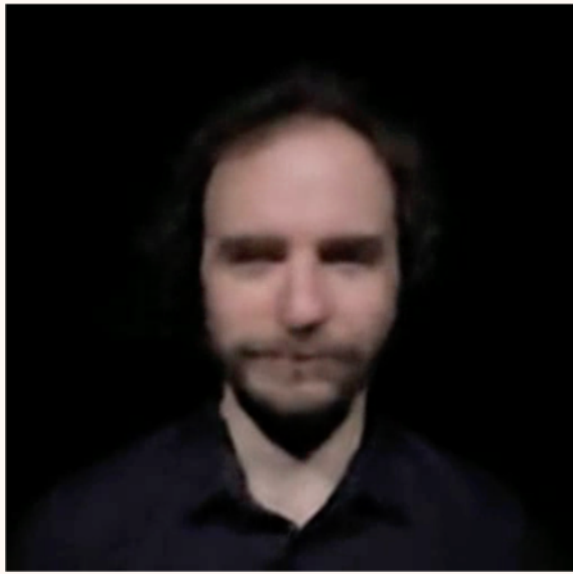
Variational Autoencoder (VAE)

Input:
Real image



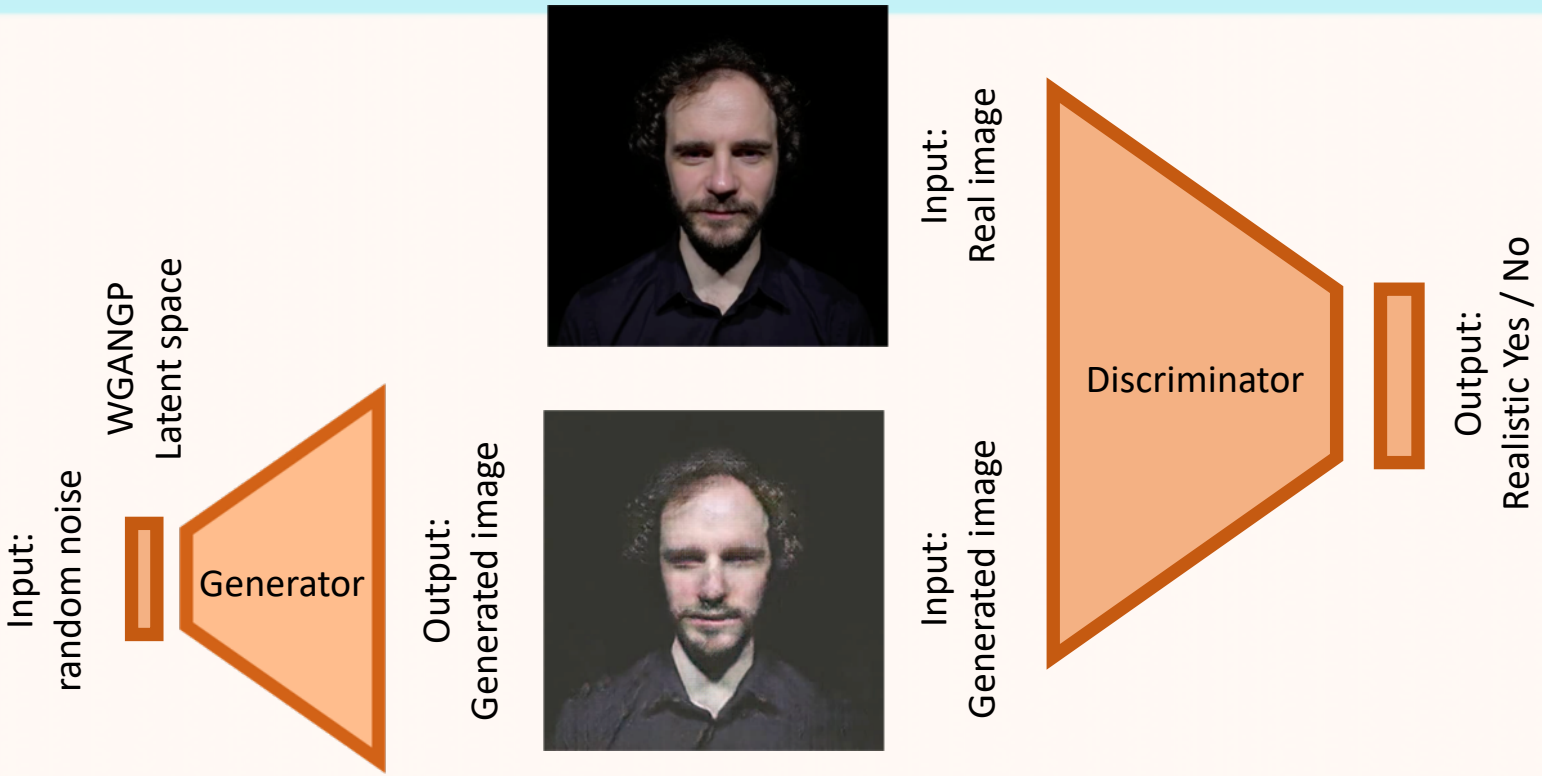


Output:
Generated image



Kingma, D. P., & Welling, M. (2013). Auto-encoding variational bayes. *arXiv preprint arXiv:1312.6114*.

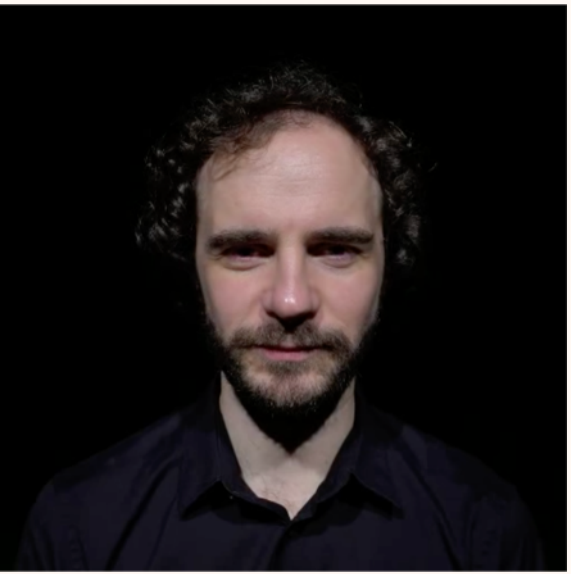
Wasserstein GAN with Gradient Penalty (WGAN-GP)

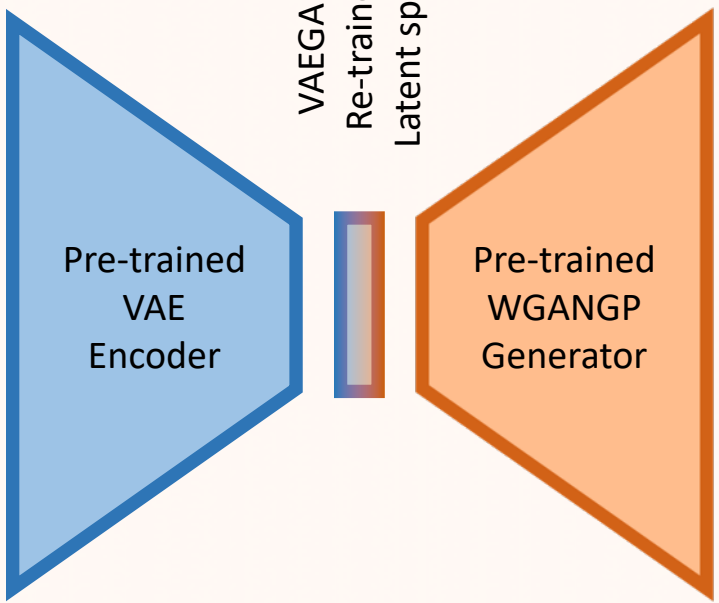


Gulrajani, I., Ahmed, F., Arjovsky, M., Dumoulin, V., & Courville, A. C. (2017). Improved training of wasserstein gans. In *Advances in neural information processing systems* (pp. 5767-5777).

Custom architecture: VAE + WGAN-GP (VAEGAN)

Input:
Real image





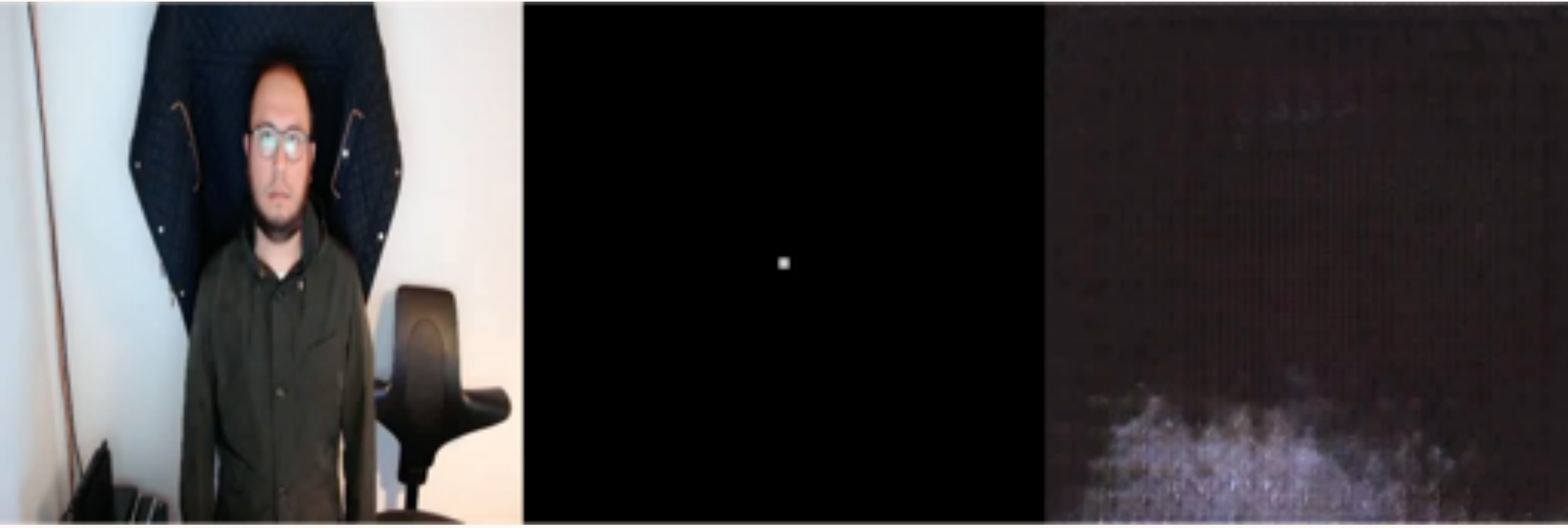
Output:
Generated image

First results

Original VAE GAN VAE-GAN

Interactive output

Face tracking



GUI

Composition: **Alexander Schubert** (www.alexanderschubert.net)
Performance: **Ensemble Resonanz** (www.ensembleresonanz.com)

Composition

Festival PODIUM Esslingen #BeBeethoven

Composition: **Alexander Schubert** (www.alexanderschubert.net)
Performance: **Ensemble Resonanz** (www.ensembleresonanz.com)

References

[1] VAE: https://arxiv.org/pdf/1312.6114.pdf
[2] AE: https://web.stanford.edu/class/psych209a/ReadingsByDate/02_06/PDPVolIChapter8.pdf
[3] DCGAN: https://arxiv.org/pdf/1511.06434.pdf
[4] WGAN: https://arxiv.org/pdf/1701.07875.pdf
[5] WGAN-GP: https://arxiv.org/pdf/1704.00028.pdf
[6] Pix2Pix: https://arxiv.org/pdf/1611.07004.pdf
[7] CycleGAN: https://arxiv.org/pdf/1703.10593.pdf

Let’s connect in LinkedIn!

Jorge Davila-Chacon, PhD
Co-founder and CTO

Focus
AI | Reinforcement Learning
Data Science
Robotics

jorge @ heldenkombinat.com

We deliver advice, skills and business opportunities in all machine learning playing fields.