Implementation and Evaluation of Variational Autoencoders for High Resolution Medical Imaging

Stefan Dünhuber

Munich, 10 January 2020

Department of Orthopaedics and Sportsorthopaedics
Prof. Dr. R. von Eisenhart-Rothe

Agenda

- Motivation, Goal
- Deep Learning
  - Autoencoder (AE)
  - Variational Autoencoder (VAE)
  - Generative Adversarial Net (GAN)
  - Advanced Generative Models (SVAE, VQ-VAE, IntroVAE)
- Experiments, Results
- Conclusion

Vision

Improve medical imaging analysis

Support radiologists

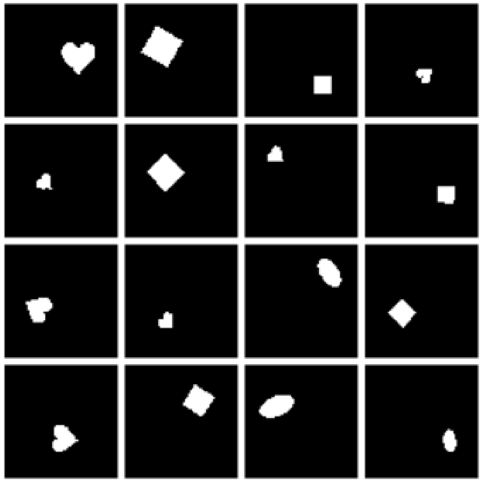
Improve diagnosis and treatments for patients

More Concrete

⇨ Disentangling, analyzing and manipulating the latent space of X-Ray knee data in an unsupervised manner

What does it mean?

Latent space, a simple example:



Ground truth of the data:

- Geometric shape
- X-position
- Y-position
- Rotation
- Size

Disentangling: independence of each latent variable to the others

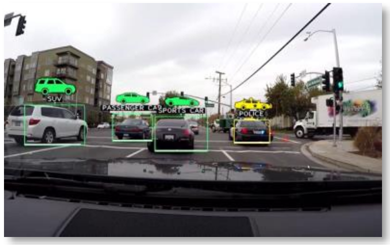
Using VAE Models

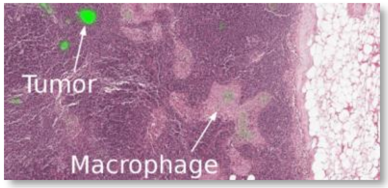
⇨ Disentangling, analyzing and manipulating the latent space of X-Ray knee data in an unsupervised manner

But how?

⇨ Using Deep Learning, due to great success in many different areas

- Natural language processing (NLP) [1]
- Autonomous driving [2]
- Predicting earthquakes [3]
- Predicting the stock market [4]
- Predicting Alzheimer's [5]
- Cancer detection [6]
- Surgical assistance [7]
- And many more…

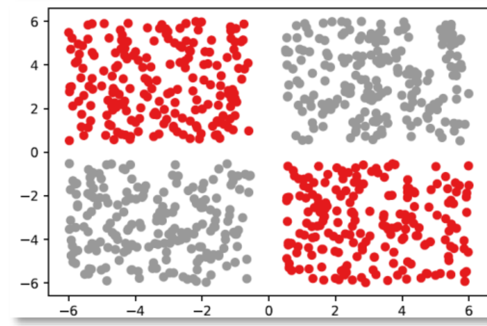




Do it unsupervised...

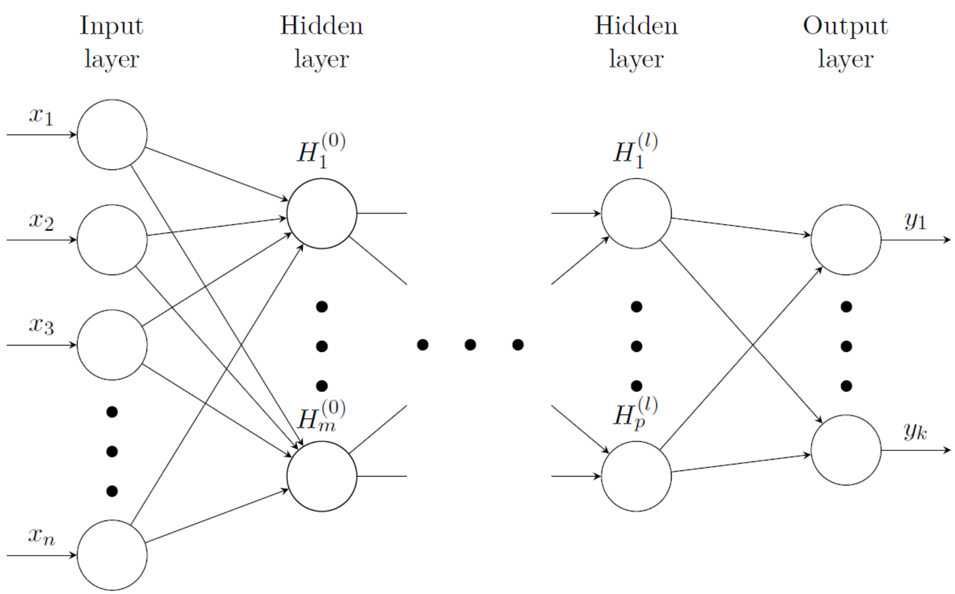
⇨ Using Variational Autoencoder (VAE) Models proposed in [8]

Deep Learning aka MLP aka ANN



Can handle non-linearity in data

- Series of composed non-linear functions in a chain
- Composed functions map the input to a given output
- Each function represents a layer
- A layer contains untis → artificial neurons

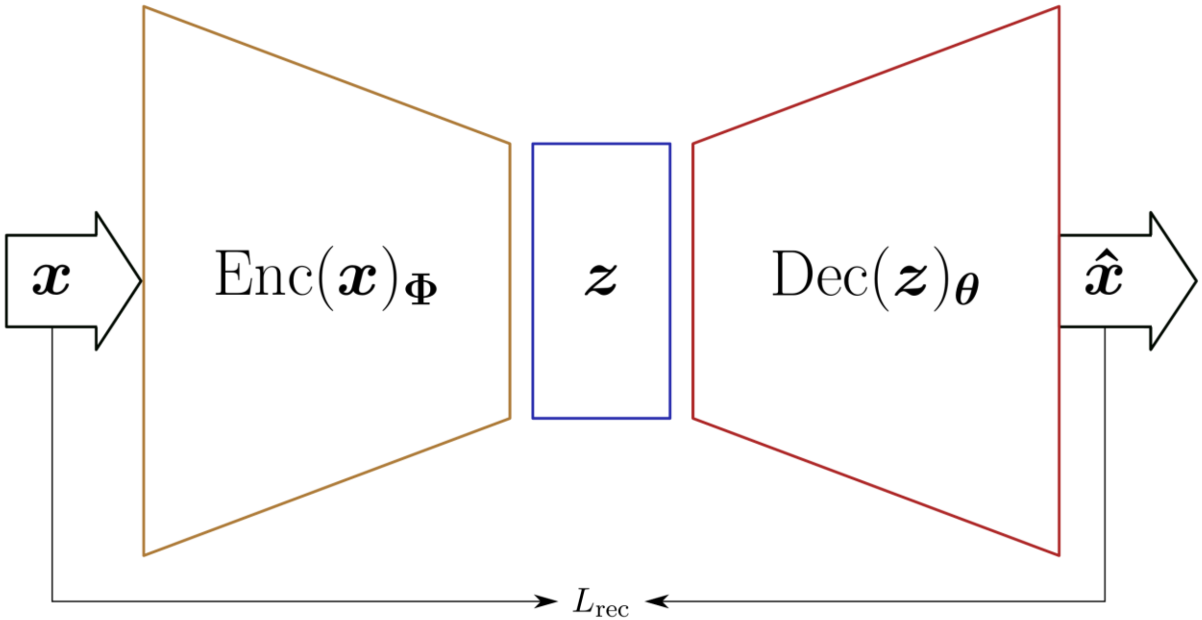


Autoencoder (AE)

Unsupervised neural network that is trained to predict the input itself

1. Encoder: maps the input $\boldsymbol{x}$ to a lower dimensional space $\boldsymbol{z}$
2. Decoder: tries to reproduce the input from $\boldsymbol{z}$

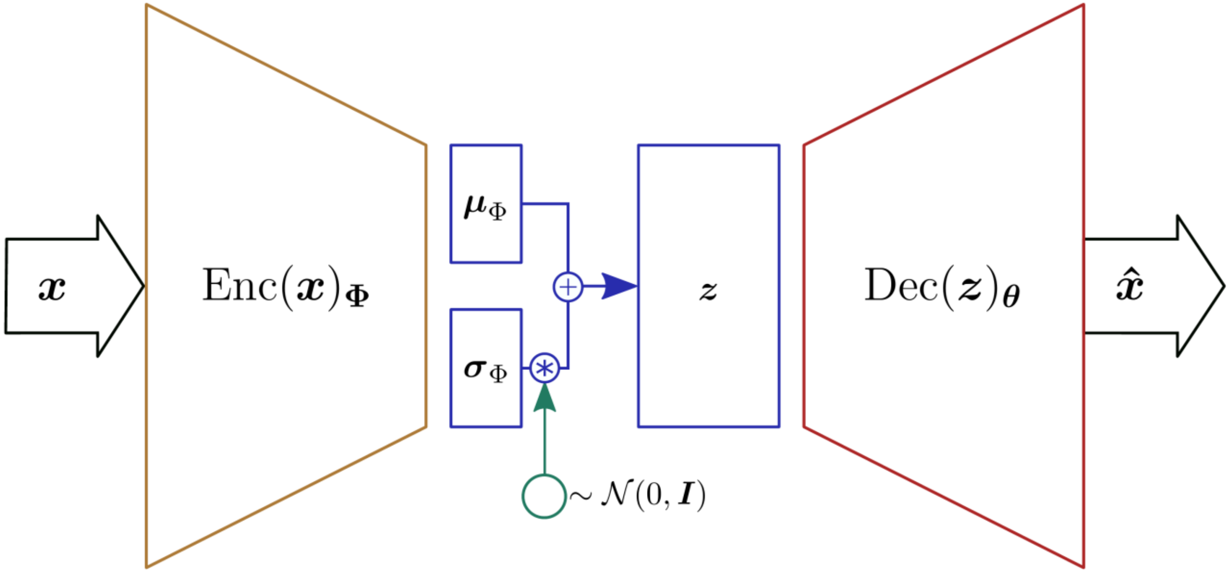
Loss: $$L_{\text{rec}}(\boldsymbol{x}) = \frac{1}{2}||\boldsymbol{x} - \hat{\boldsymbol{x}}||^2 = \frac{1}{2}||\boldsymbol{x} - \text{Dec}_{\boldsymbol{\theta}}(\text{Enc}_{\boldsymbol{\Phi}}(\boldsymbol{x}))||^2$$



Variational Autoencoder (VAE)

Force the output of the encoder to be normal distributed

Loss: $L_{\mathrm{VAE}}(\boldsymbol{\theta}, \boldsymbol{\Phi}) = \underbrace{-\mathbb{E}_{z \sim q_{\boldsymbol{\Phi}}(\boldsymbol{z}|\boldsymbol{x})}[\log p_{\boldsymbol{\theta}}(\boldsymbol{x}|\boldsymbol{z})]}_{L_{\mathrm{rec}}} + \underbrace{\mathbb{KL}(q_{\boldsymbol{\Phi}}(\boldsymbol{z}|\boldsymbol{x})||p(\boldsymbol{z}))}_{L_{\mathrm{KL}}}$

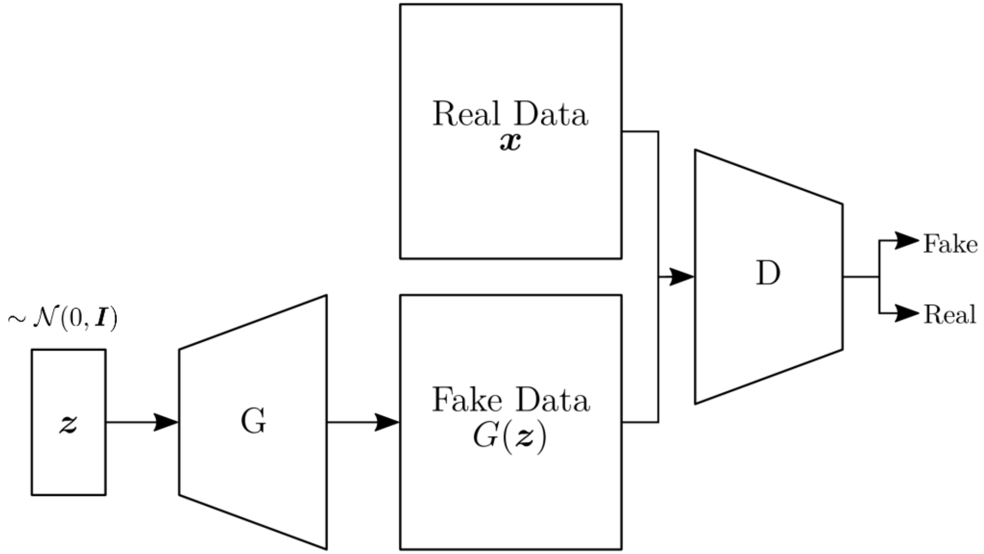


- Generative model, which allows disentangling of the latent space
- Tends to produce blurry reconstruction images

Generative Adversarial Net (GAN)

Generative Model consisting of an adversarial process

➡ Min-max game between a generator G and a discriminator D



➡ Generative model generating sharper images as VAE

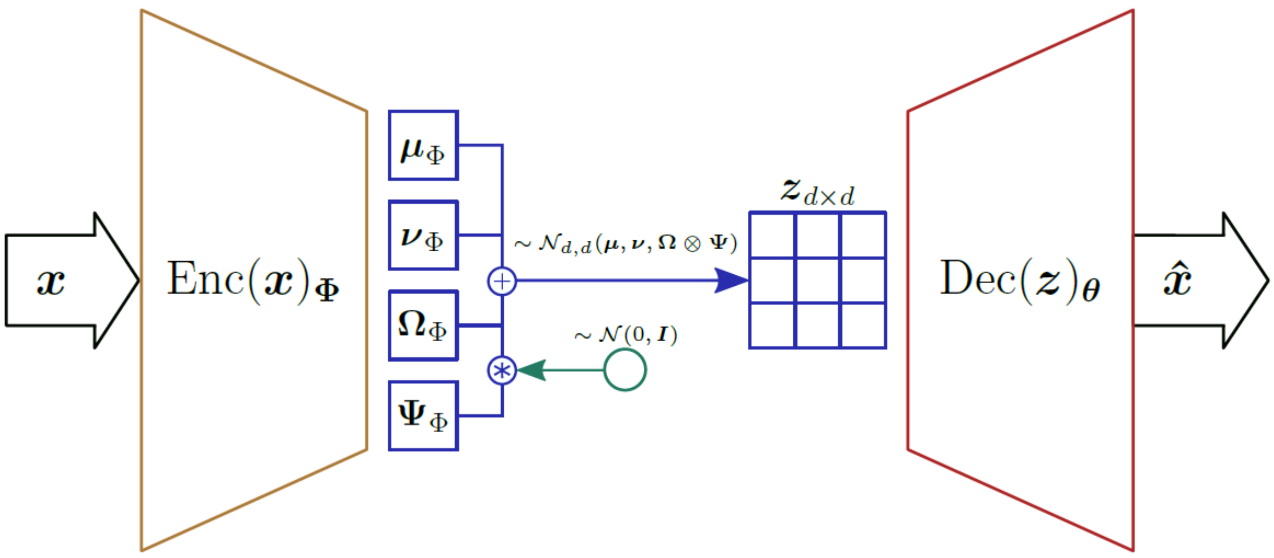
➡ No analyzation of the latent space for a given input

Spatial Variational Autoencoder (SVAE)

VAE tends to produce blurry reconstruction images

How to tackle this problem?

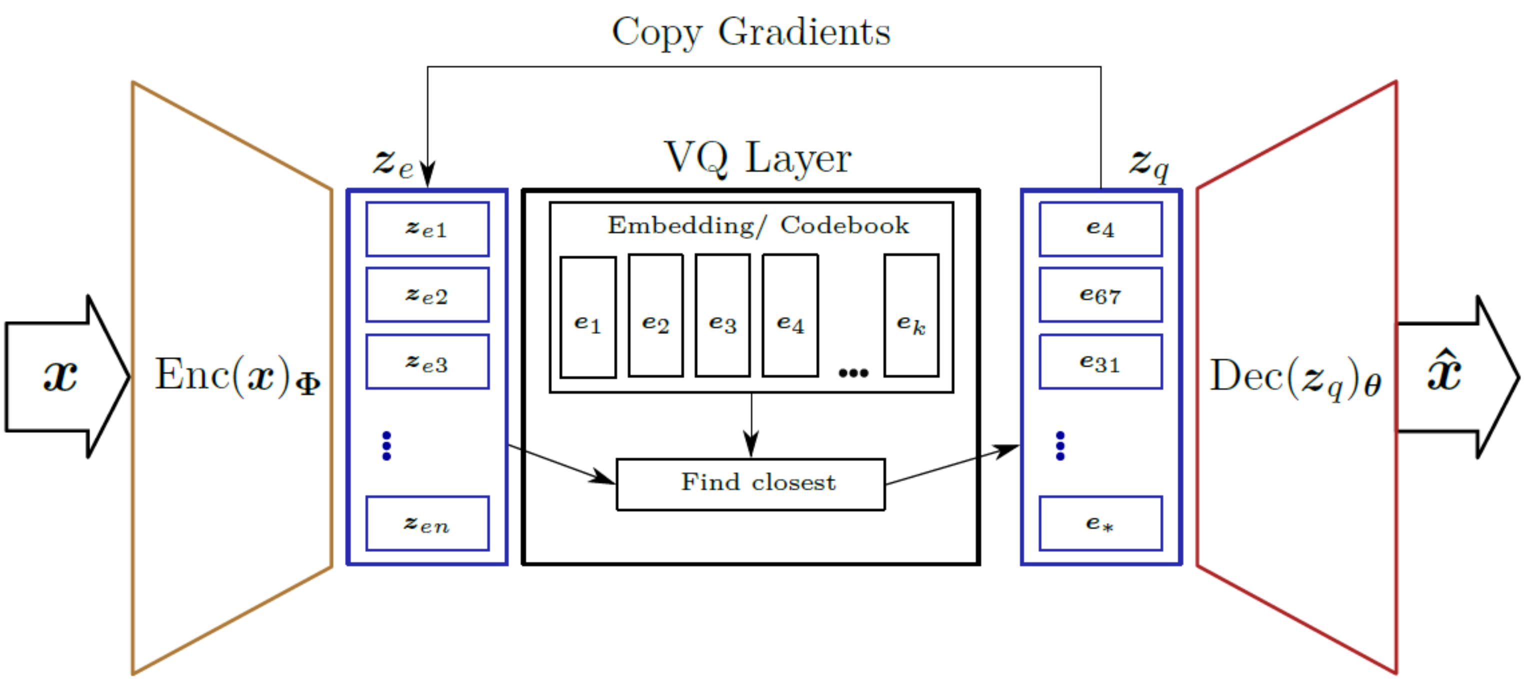
Trivial solution: multidimensional latent space → Spatial Variational Autoencoder



Hard to analyze a multidimensional latent space

Vector Quantized Variational Autoencoder (VQ-VAE)

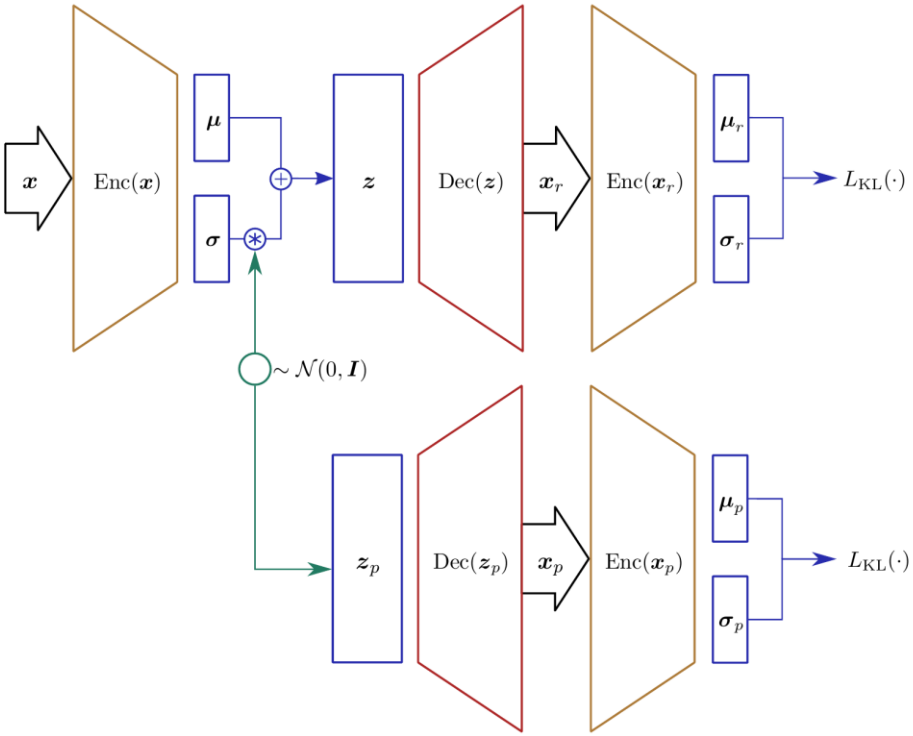
Discretize the latent space → Vector Quantized Variational Autoencoder (VQ-VAE)



Introspective Variational Autoencoder (IntroVAE)

Combines VAE and GAN into one model → IntroVAE

Uses a min max game approach applied to the VAE model

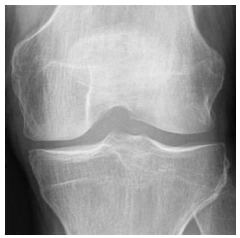


Experiments

Knee Osteoarthritis Severity Grading dataset

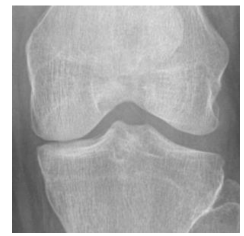
KL- 0

KL- 1

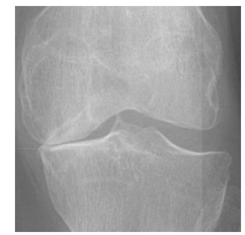


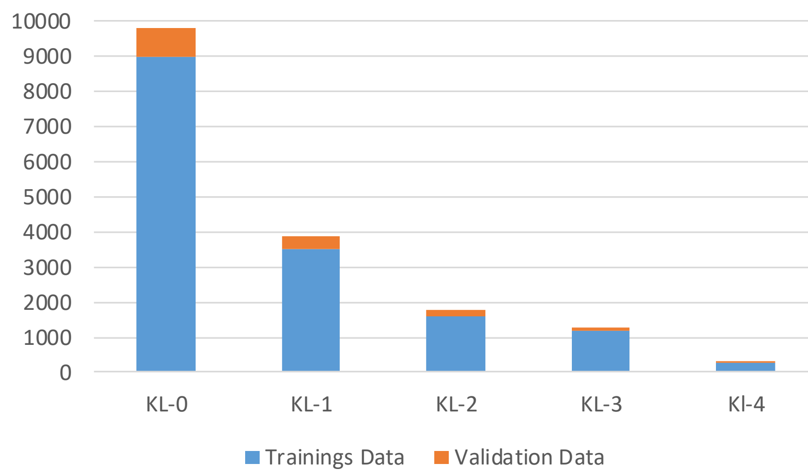
KL- 2

KL- 3



KL- 4

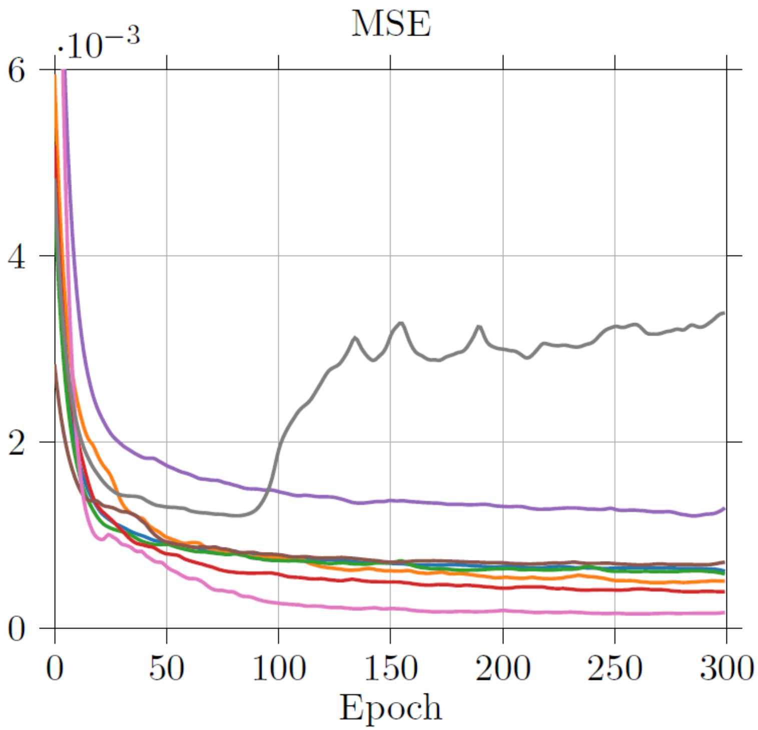




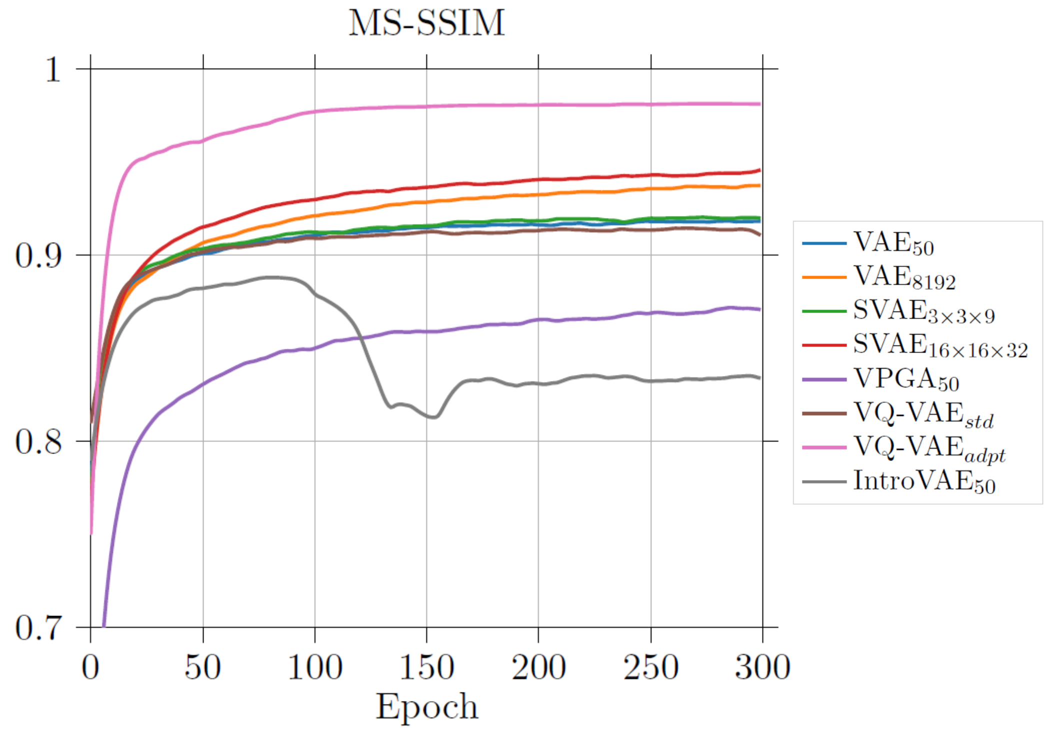
Use almost same architecture for the Encoder and Decoder

Quantitative Results

Mean Square Error

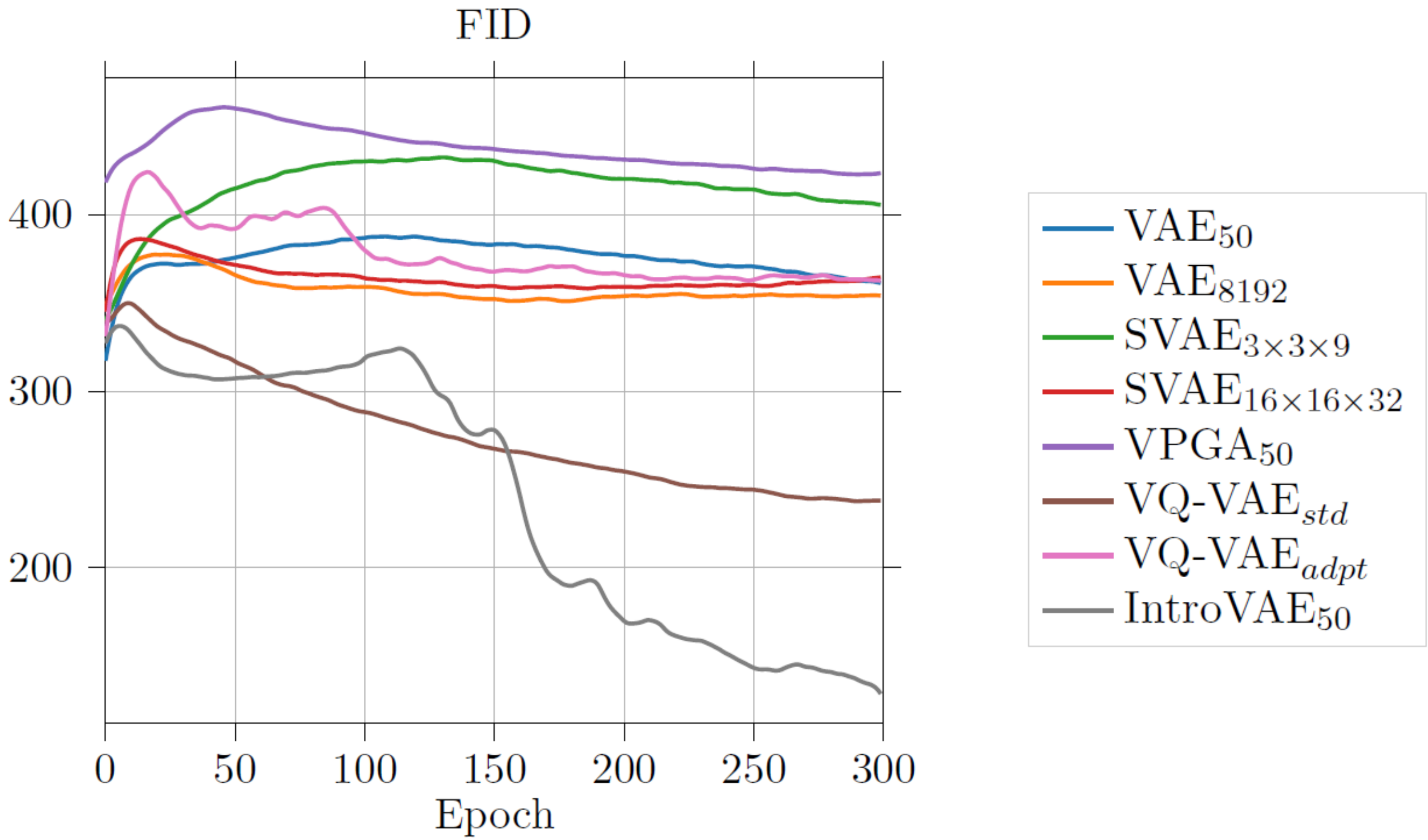


Multi-Scale Structural Similarity Method

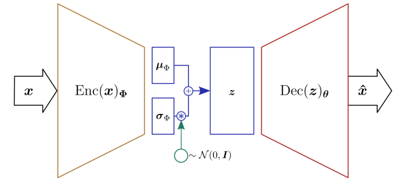


Quantitative Results

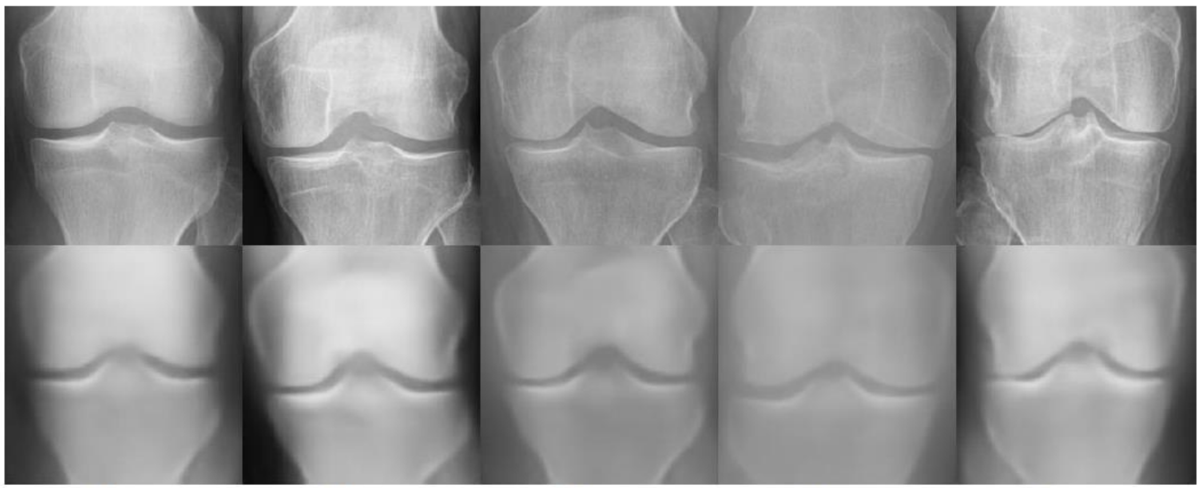
Fréchet Inception Distance



Qualitative Results - VAE

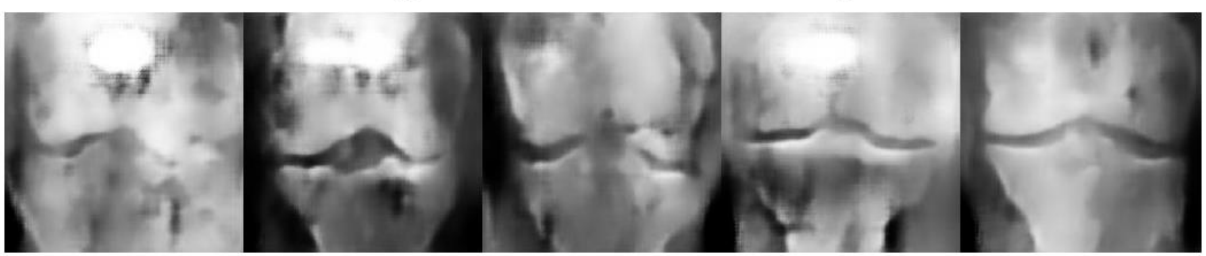


VAE$_{50}$ - Reconstructions of Test Data



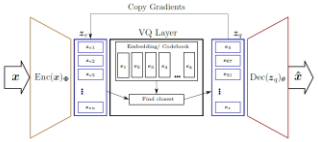
KL-0 KL-1 KL-2 KL-3 KL-4

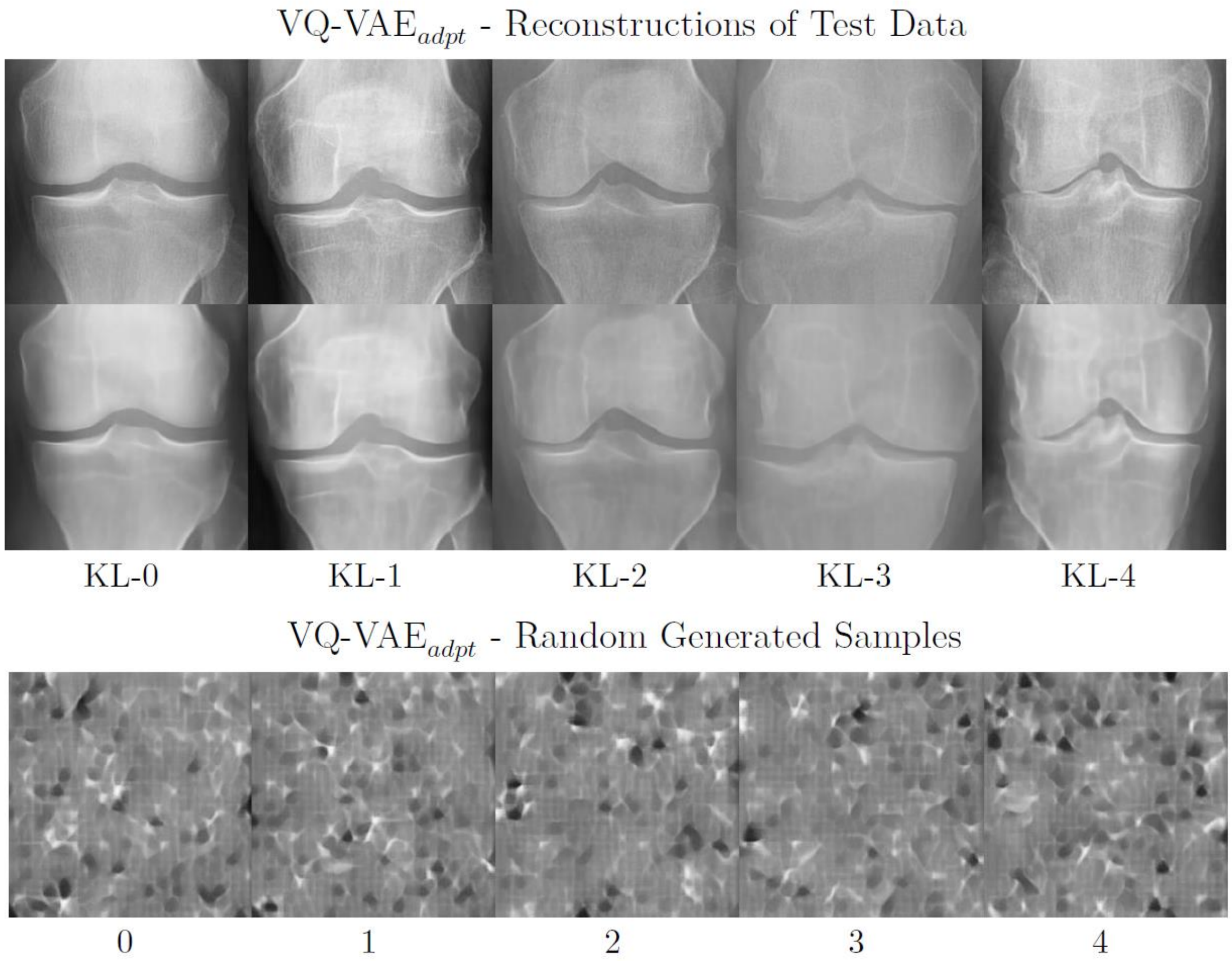
VAE$_{50}$ - Random Generated Samples

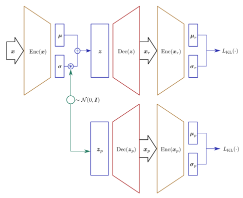


0 1 2 3 4

Qualitative Results – VQ-VAE

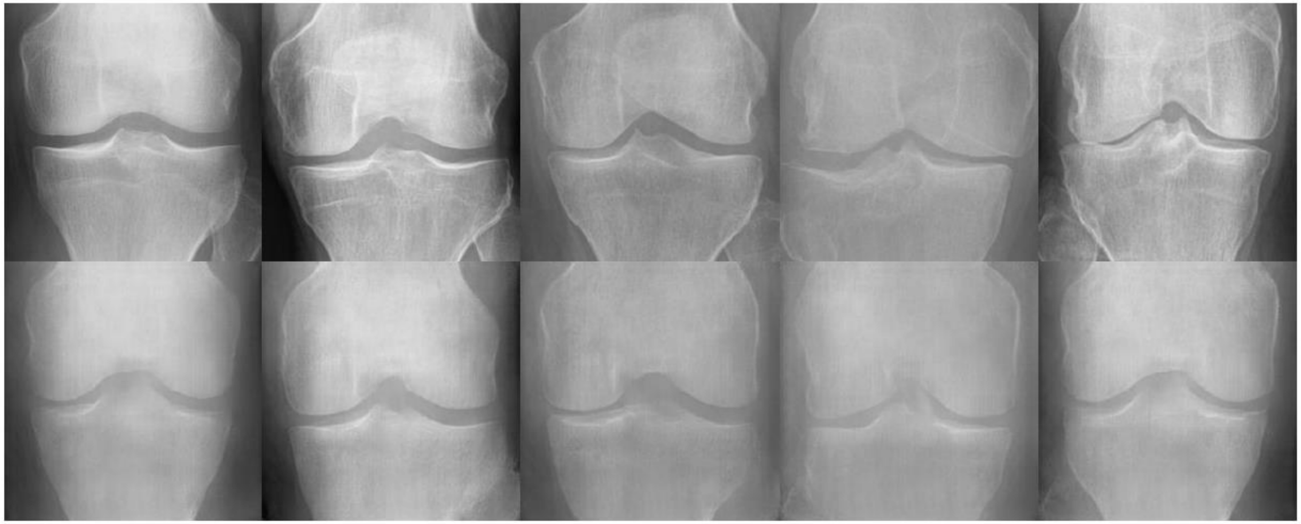






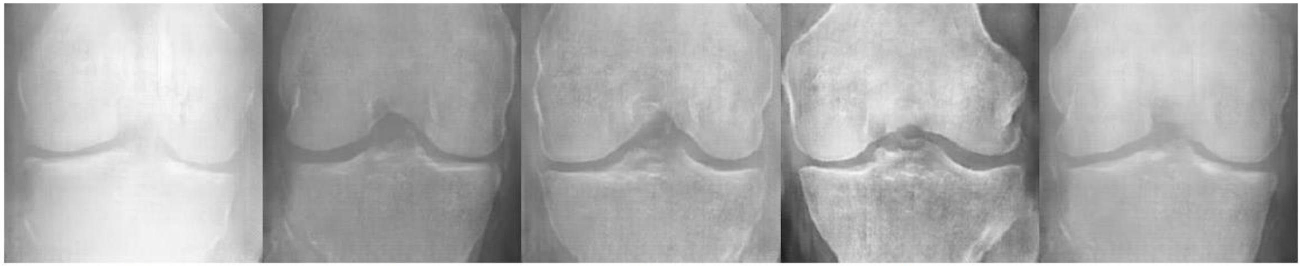
Qualitative Results – IntroVAE

IntroVAE$_{50}$ - Reconstructions of Test Data



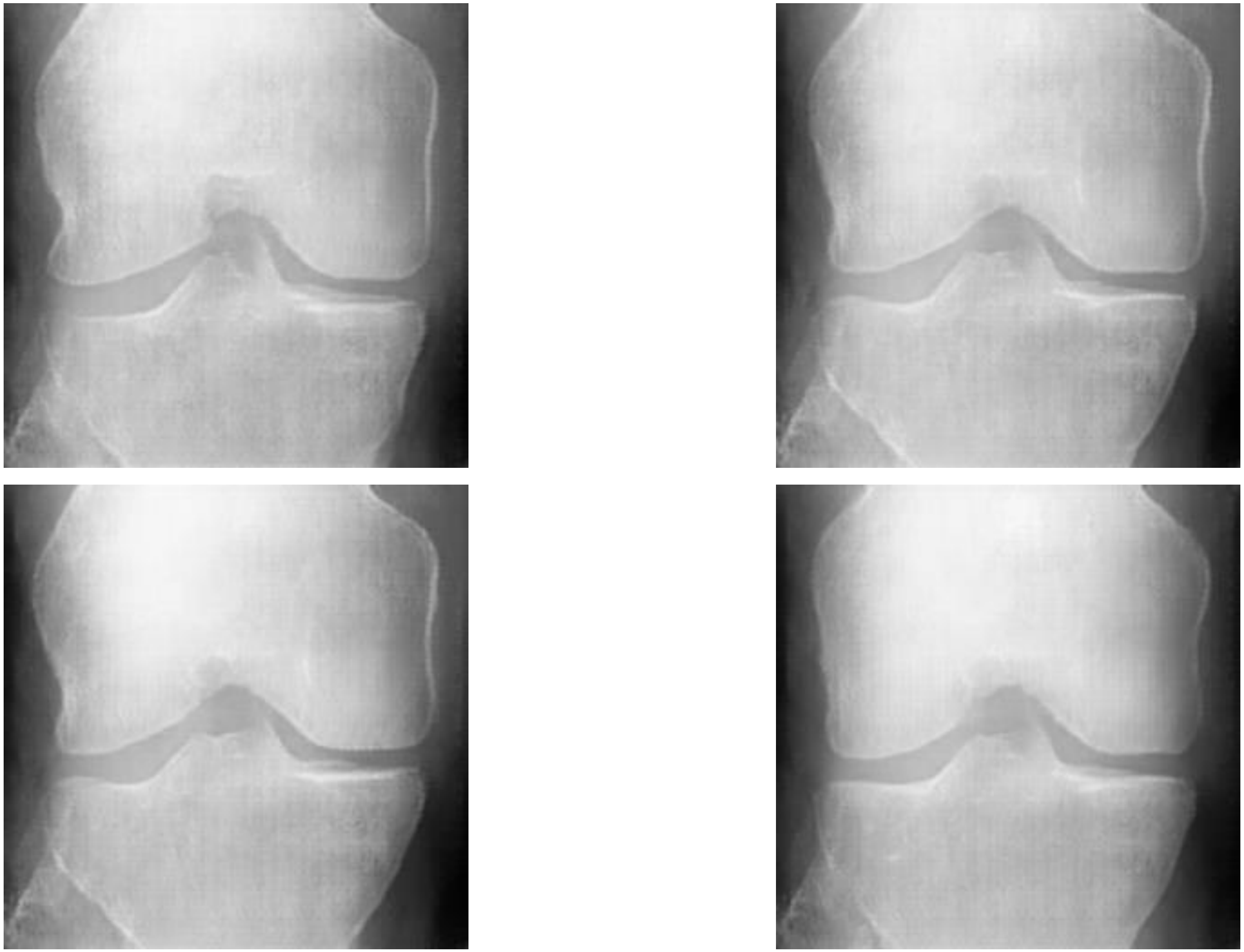
KL-0 KL-1 KL-2 KL-3 KL-4

IntroVAE$_{50}$ - Random Generated Samples



0 1 2 3 4

Live Demo – Latent Space IntroVAE



Summary

- ✓ Applied different VAE Models to high resolution medical images
- ✓ Get very sharp reconstruction results with the VQ-VAE model
- ✓ Able to generate new sharp knee images (IntroVAE)
- ✓ Able to manipulate and distentagle the latent space (IntroVAE)
- ✓ Making the model more complicated not always improves the performance (SVAE, VPGA)
- ✓ Build the basis for many medical applications

Thank you

Contact

Stefan Dünhuber
ORTUM - Center for Orthopaedic Research at TUM

Klinikum rechts der Isar
Technical University of Munich
Department of Orthopaedics and Sportsorthopaedics
Ismaninger Strasse 22
81675 Munich/Germany

Tel.: +49 89 4140 7880
Fax: +49 89 4140 7881
E-Mail: your.name@tum.de
Web: www.ortho.med.tum.de

Sources

[1] Tom Young, Devamanyu Hazarika, Soujanya Poria, and Erik Cambria. Recent trends in deep learning based natural language processing [review article]. IEEE Computational Intelligence Magazine, 13(3):55–75, 2018. ISSN 1556-603X. doi: 10.1109/MCI.2018.2840738.
[2] AhmadEL Sallab, Mohammed Abdou, Etienne Perot, and Senthil Yogamani. Deep reinforcement learning framework for autonomous driving. *Electronic Imaging*, 2017 (19):70–76, 2017. ISSN 2470-1173. doi: 10.2352/ISSN.2470-1173.2017.19.AVM-023.
[3] Asim, K. M.; Martínez-Álvarez, F.; Basit, A.; Iqbal, T. Earthquake magnitude prediction in Hindukush region using machine learning techniques,2017. doi: 10.1007/s11069-016-2579-3
[4] Hongbing Ouyang, Xiaowei Zhang, Hongju Yan. Index tracking based on deep neural network, 2019. https://doi.org/10.1016/j.cogsys.2018.10.022
[5] Garam Lee, Kwangsik Nho, Byungkon Kang, Kyung-Ah Sohn, and Dokyoon Kim. Predicting alzheimer’s disease progression using multi-modal deep learning approach. Scientific reports, 9(1):1952, 2019. doi: 10.1038/s41598-018-37769-z.

Sources

[6] Wenya Linda Bi, Ahmed Hosny. Artificial intelligence in cancer imaging: Clinical challenges and applications. CA: a cancer journal for clinicians, 69(2):127–157, 2019. doi: 10.3322/caac.21552.

[7] Daniil Pakhomov, Vittal Premachandran, Max Allan, Mahdi Azizian, and Nassir Navab. Deep residual learning for instrument segmentation in robotic surgery. URL http://arxiv.org/pdf/1703.08580v1.

[8] Diederik P. Kingma and Max Welling. Auto-encoding variational bayes. URL https://arxiv.org/pdf/1312.6114.pdf.