Semantic Latent Space Regression of Diffusion Autoencoders for Vertebral Fracture Grading

Matthias Keicher[1*], Matan Atad[1*], MD David Schinz[2], MD Alexandra S. Gersing[2,3], MD Sarah C. Foreman[2], MD Sophia S. Goller[3], MD Juergen Weissinger[3], MD Jon Rischewski[3], MD Anna-Sophia Dietrich[2], MD Benedikt Wiestler[2], MD Jan S. Kirschke[2], and Nassir Navab[1]

[1] Chair for Computer-Aided Medical Procedures and Augmented Reality, Technische University of Munich, Germany

[2] Department of Diagnostic and Interventional Radiology, School of Medicine, Technical University of Munich, Germany

Department of Neuroradiology, University Hospital of Munich, Ludwig-Maximilians University, Germany

* *Contributed equally.*

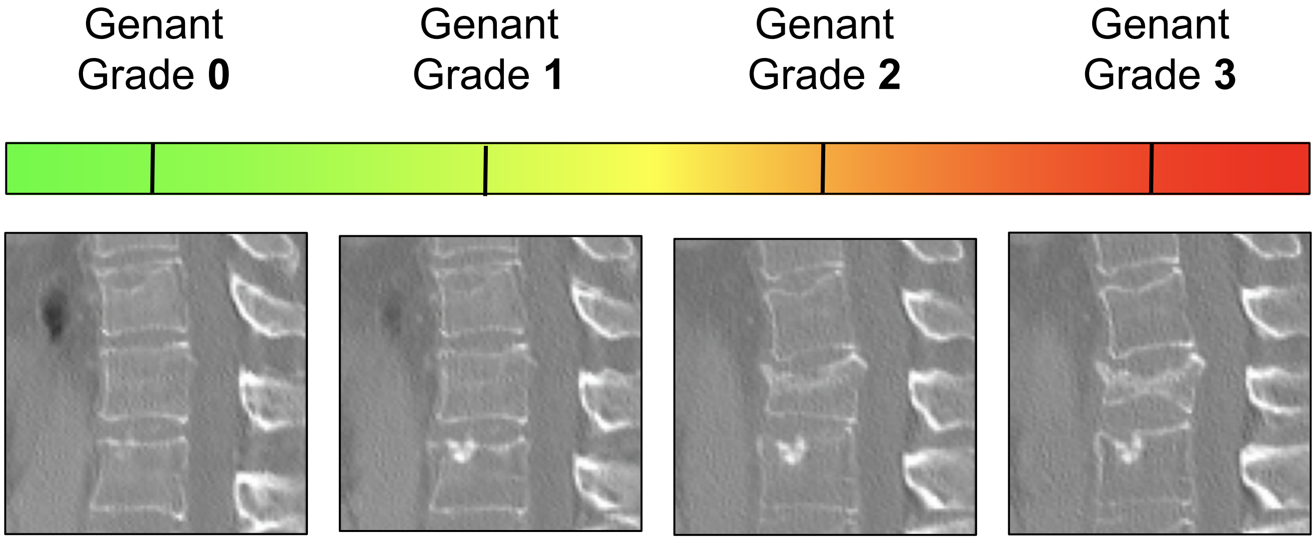
MICCAI 2023 Vancouver CANADA

Technical University of Munich

Computer Aided Medical Procedures

Klinikum rechts der Isar Technische Universität München

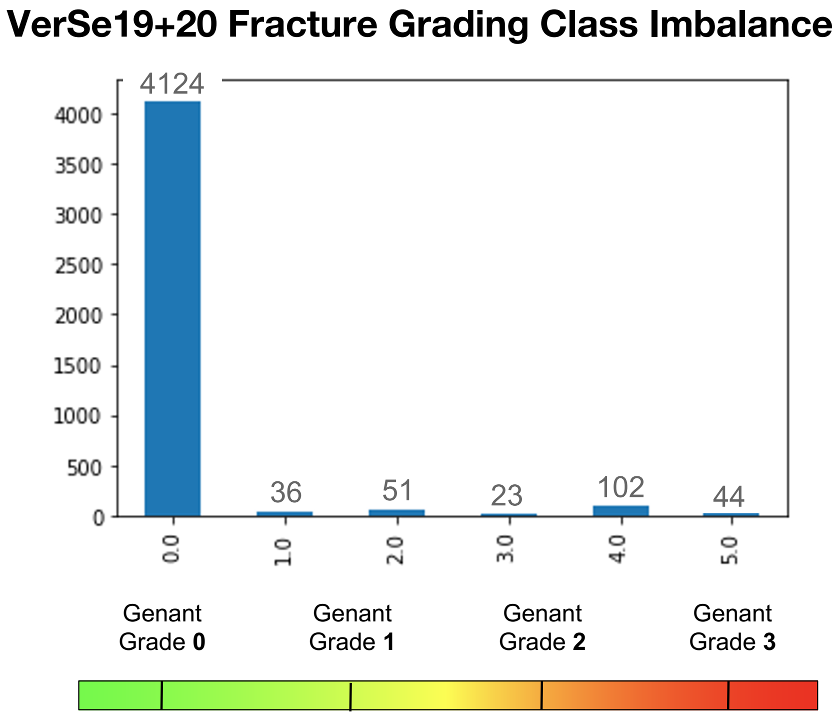
Vertebral fracture grading with the genant scale



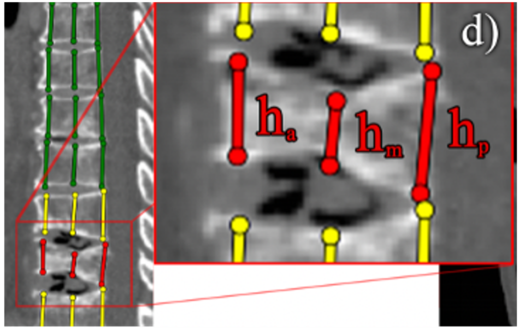
❌ Light fractures are hard to detect

❌ Low inter-rater agreement

Severe class imbalance

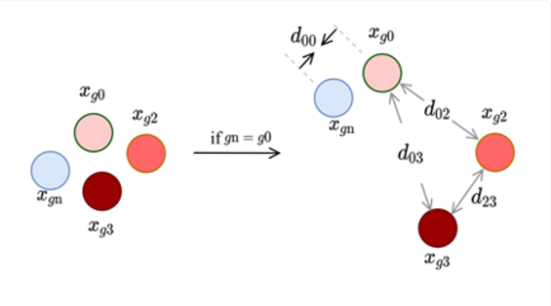


Previous approaches rely on noisy annotations and model as multiclass



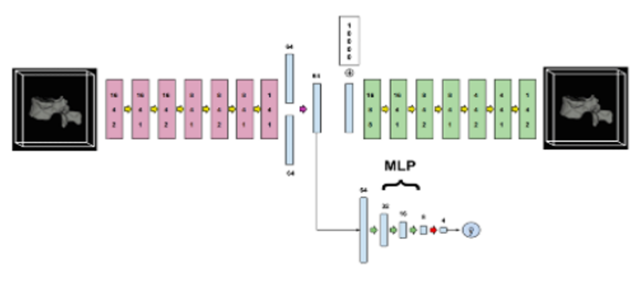
Keypoint Detection [1,2]

❌ Fully supervised



Contrastive Learning [3,4]

❌ Lack of explainability



VAE latents [5]

❌ No grading

[1] Pisov, M., Kondratenko, V., Zakharov, A., Petraikin, A., Gombolevskiy, V., Morozov, S., & Belyaev, M. (2020, October). Keypoints localization for joint vertebra detection and fracture severity quantification. In International Conference on Medical Image Computing and Computer-Assisted Intervention (pp. 723-732). Springer, Cham.

[2] Zakharov, A., Pisov, M., Bukharaev, A., Petraikin, A., Morozov, S., Gombolevskiy, V., & Belyaev, M. (2022). Interpretable Vertebral Fracture Quantification via Anchor-Free Landmarks Localization. arXiv preprint arXiv:2204.06818.

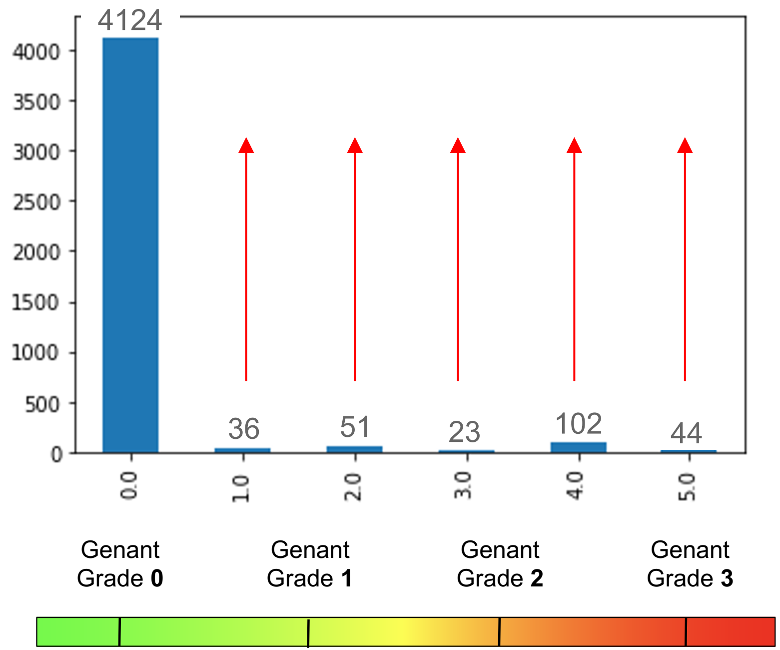
[3] Husseini, M., Sekuboyina, A., Loeffler, M., Navarro, F., Menze, B. H., & Kirschke, J. S. (2020, October). Grading loss: a fracture grade-based metric loss for vertebral fracture detection. In International Conference on Medical Image Computing and Computer-Assisted Intervention (pp. 733-742). Springer, Cham.

[4] Wei, X., Cong, H., Zhang, Z., Peng, J., Chen, G., & Li, J. (2022). Faint Features Tell: Automatic Vertebrae Fracture Screening Assisted by Contrastive Learning. arXiv preprint arXiv:2208.10698.

[5] Husseini, M., Sekuboyina, A., Bayat, A., Menze, B. H., Loeffler, M., & Kirschke, J. S. (2020). Conditioned variational auto-encoder for detecting osteoporotic vertebral fractures. In Computational Methods and Clinical Applications for Spine Imaging: 6th International Workshop and Challenge, CSI 2019, Shenzhen, China, October 17, 2019, Proceedings 6 (pp. 29-38). Springer International Publishing.

Can we address class imbalance with a generative model?

VerSe19+20 Fracture Grading Class Imbalance



Requirements:

- Augment underrepresented classes
- Model ordinal regression of fractures continuously
- Interpretability by counterfactuals

Proposed method: Regression in the latent space of a generative model

1. Train generative model

2. Find latent directions for fractures

3. Continuous regression

Invertible generative feature extractor:

- StyleGAN2 + Encoder
- Autoencoder architectures (AE, VAE)

Linear classifiers:

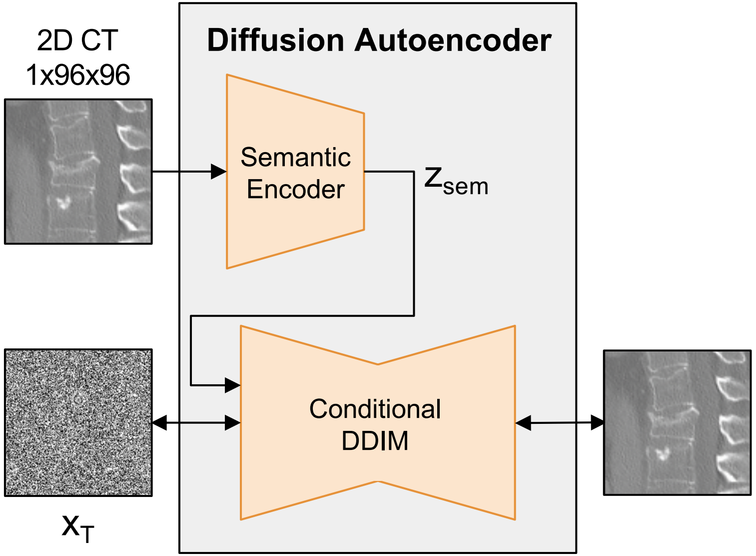
- Linear layer
- SVM

Regression:

- linear
- polynomial

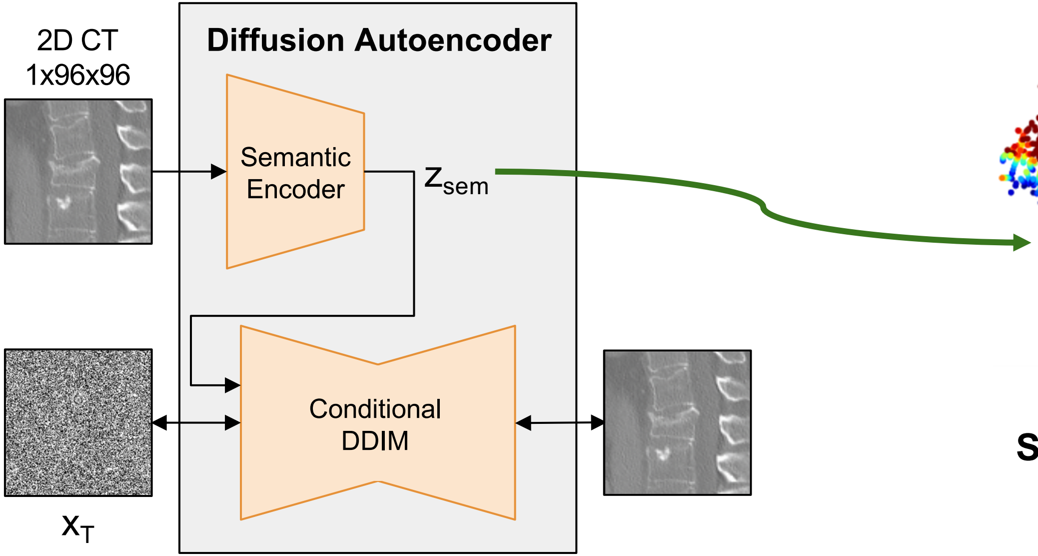
[6] Preechakul, K., Chatthee, N., Wizadwongsa, S., & Suwajanakorn, S. (2022). Diffusion autoencoders: Toward a meaningful and decodable representation. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (pp. 10619-10629).

1. **<u>Unsupervised</u> feature extraction to semantic latent**

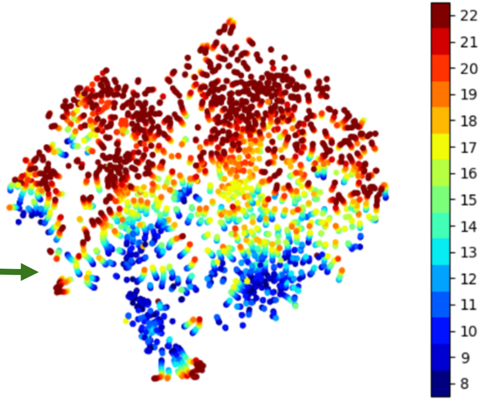


[6] Preechakul, K., Chatthee, N., Wizadwongsa, S., & Suwajanakorn, S. (2022). Diffusion autoencoders: Toward a meaningful and decodable representation. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (pp. 10619-10629).

1. Unsupervised feature extraction to semantic latent



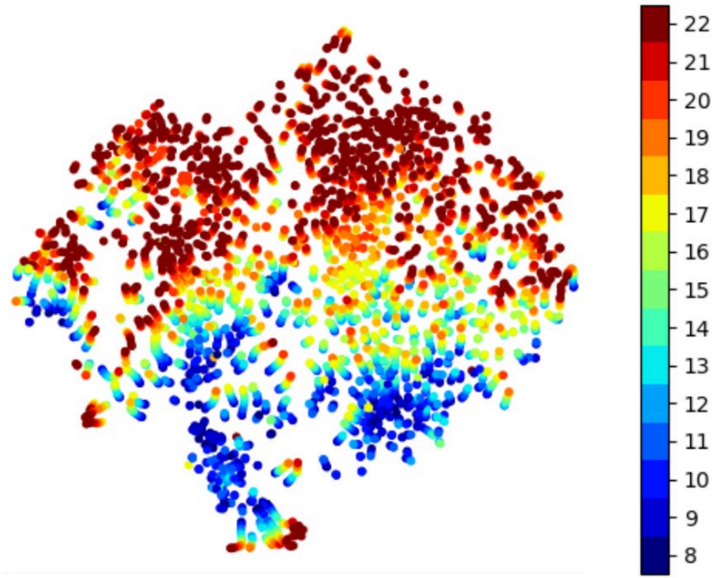
Encoding the VerSe training dataset



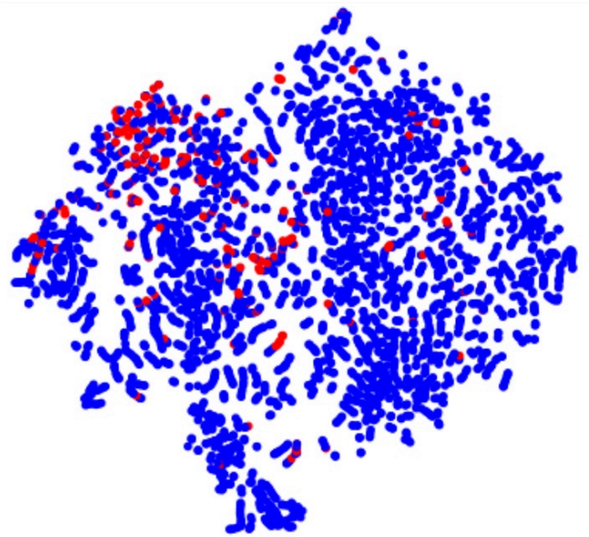
Semantic latent space

[6] Preechakul, K., Chatthee, N., Wizadwongsa, S., & Suwajanakorn, S. (2022). Diffusion autoencoders: Toward a meaningful and decodable representation. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (pp. 10619-10629).

TSNE plots show clustering of vertebra level in the spine



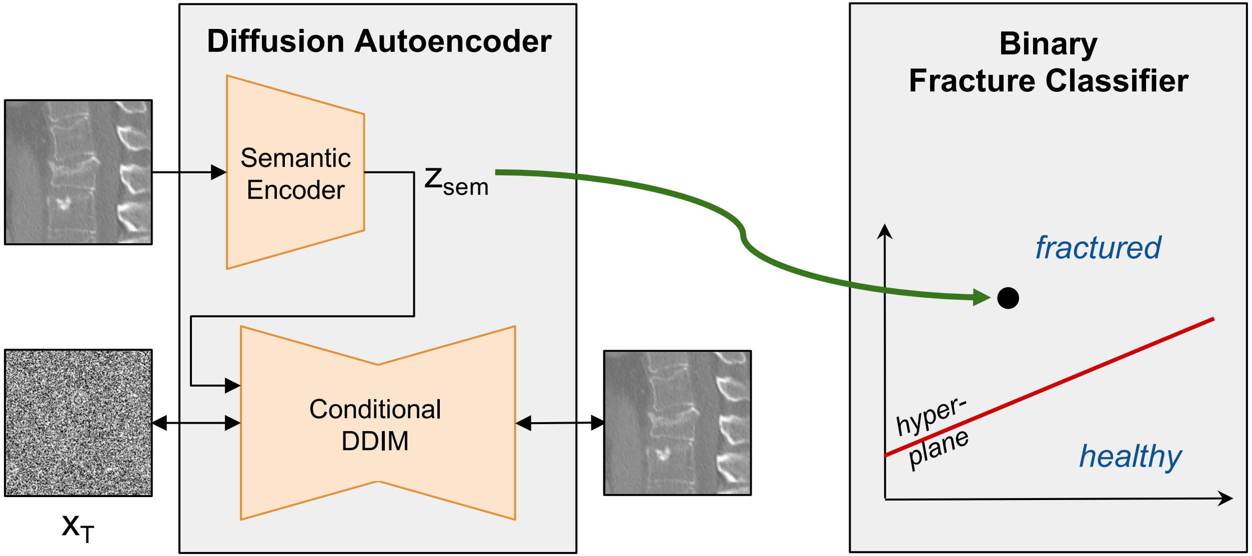
Spine levels T2-L5



Fractures

1. <u>Unsupervised</u> feature extraction to semantic latent

2. <u>Supervised</u> classification with hyperplane

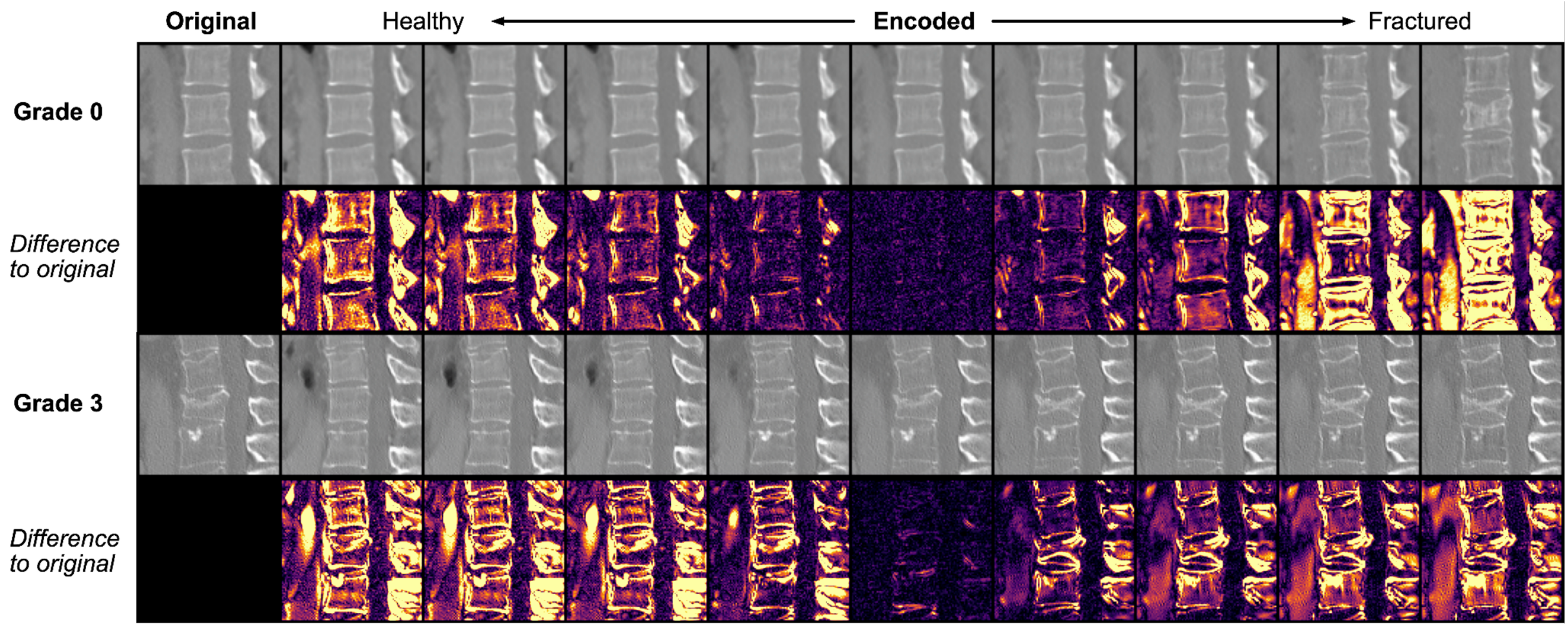


[6] Preechakul, K., Chatthee, N., Wizadwongsa, S., & Suwajanakorn, S. (2022). Diffusion autoencoders: Toward a meaningful and decodable representation. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (pp. 10619-10629).

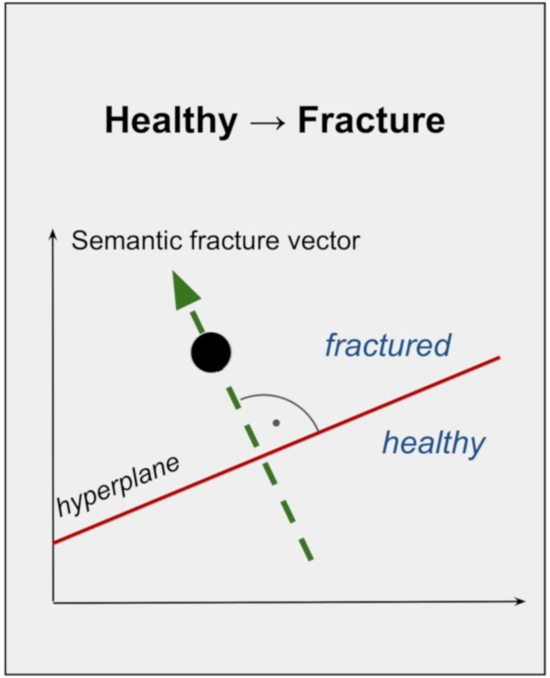
Results: DAE outperforms StyleGAN-based approach and previous AE

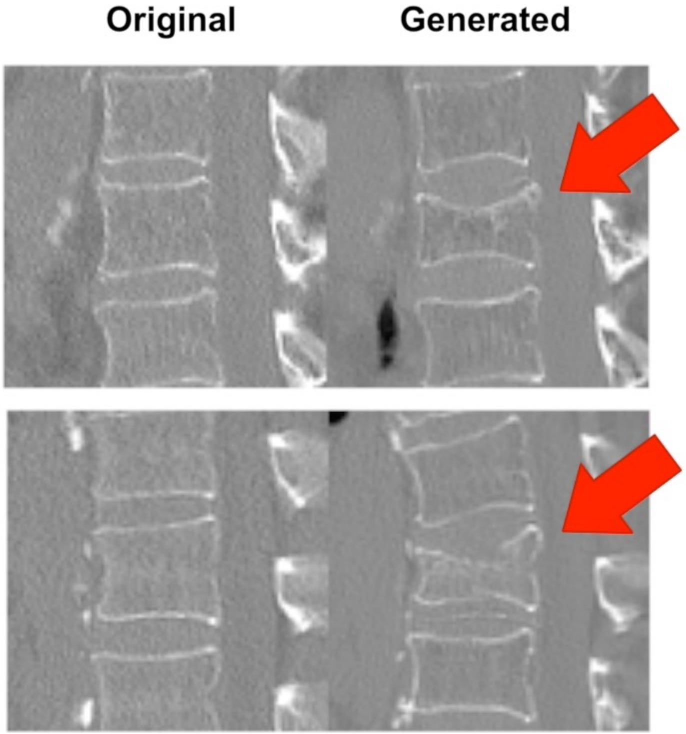
Model	Encoder	Classifier	Detection (AUC $\uparrow$)	Grading (F_1 $\uparrow$)
Linear probing trained on G0, G2 and G3 with frozen encoder of generative model				
AE [5]	AE	Linear layer	0.70*	-
VAE [5]	VAE	Linear layer	0.77*	-
Spine-VAE [5]	VAE	Linear layer	0.81*	-
StyleGAN2	E4E	SVM	0.74	-
DAE	DAE	Linear layer	0.96	(0.23)
DAE	DAE	SVM	0.93	0.59

Visualizing the semantic latent space across the decision boundary

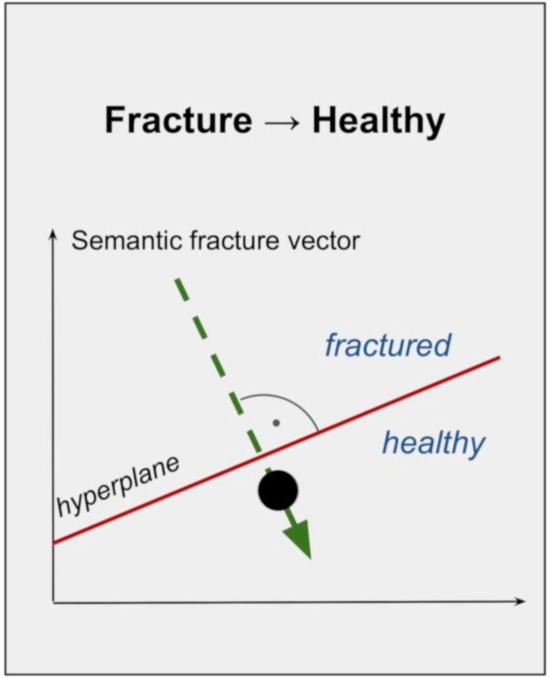


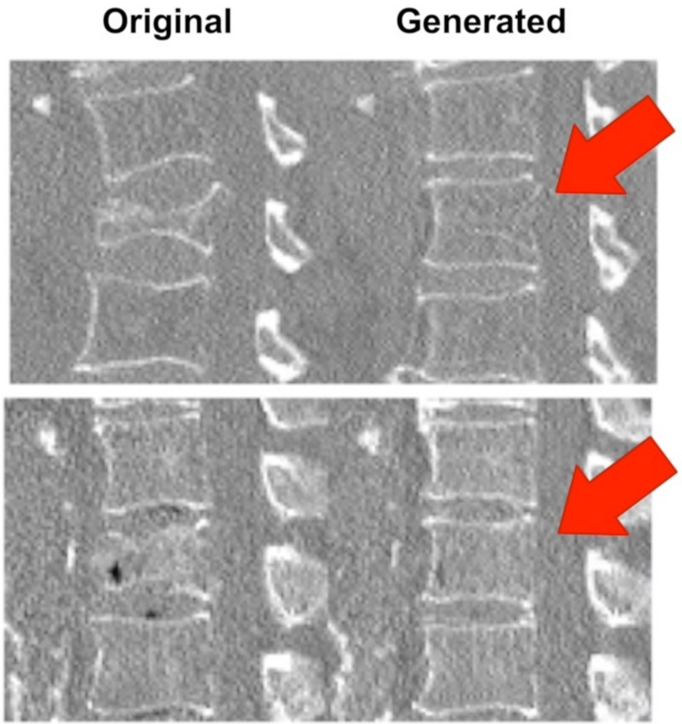
Counterfactual generation: Healthy to fractured





Counterfactual generation: Fractured to healthy

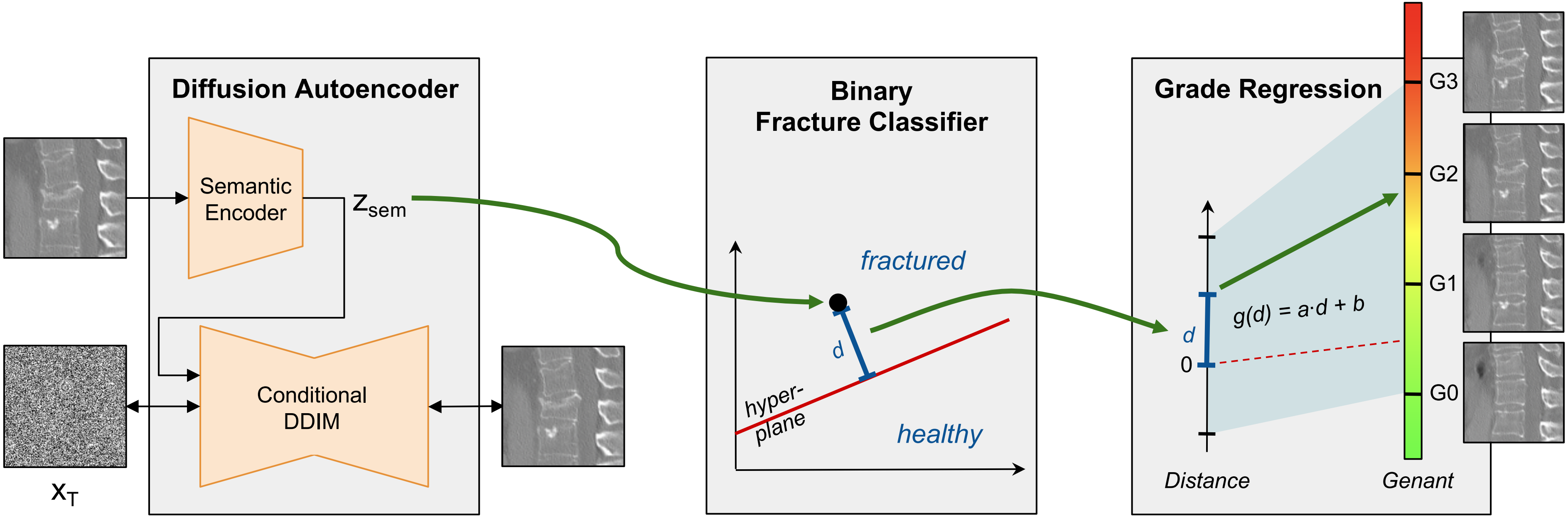




1. Unsupervised feature extraction to semantic latent

2. Supervised classification with hyperplane

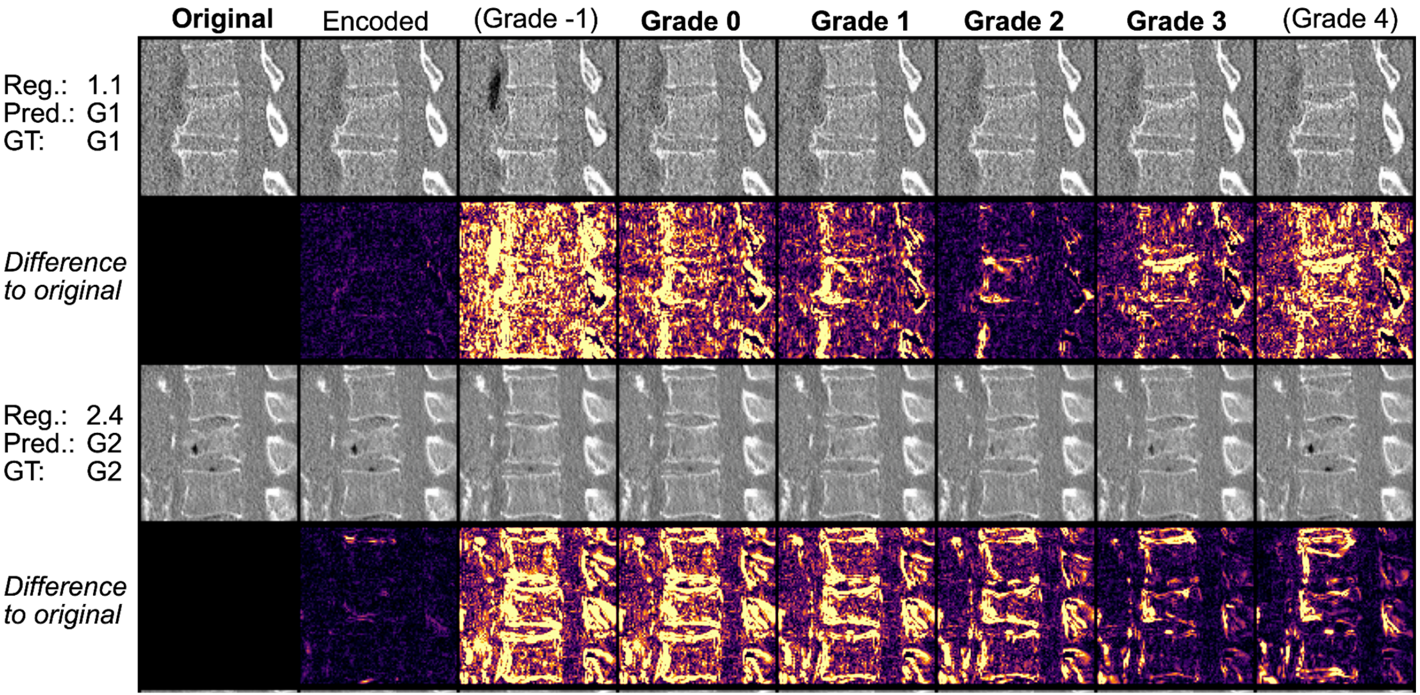
3. Calibration of latent distance to grades



[6] Preechakul, K., Chatthee, N., Wizadwongsa, S., & Suwajanakorn, S. (2022). Diffusion autoencoders: Toward a meaningful and decodable representation. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (pp. 10619-10629).

What would a vertebra with Genant Grade X look like?

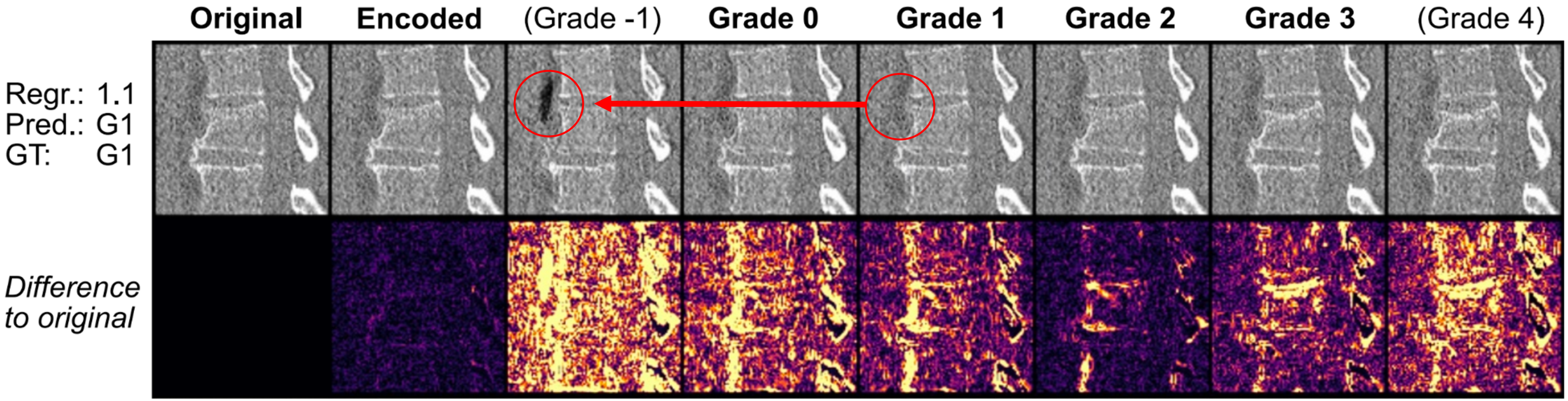
Generation of **Ordinal Counterfactuals**



Results: Polynomial calibration improves grading

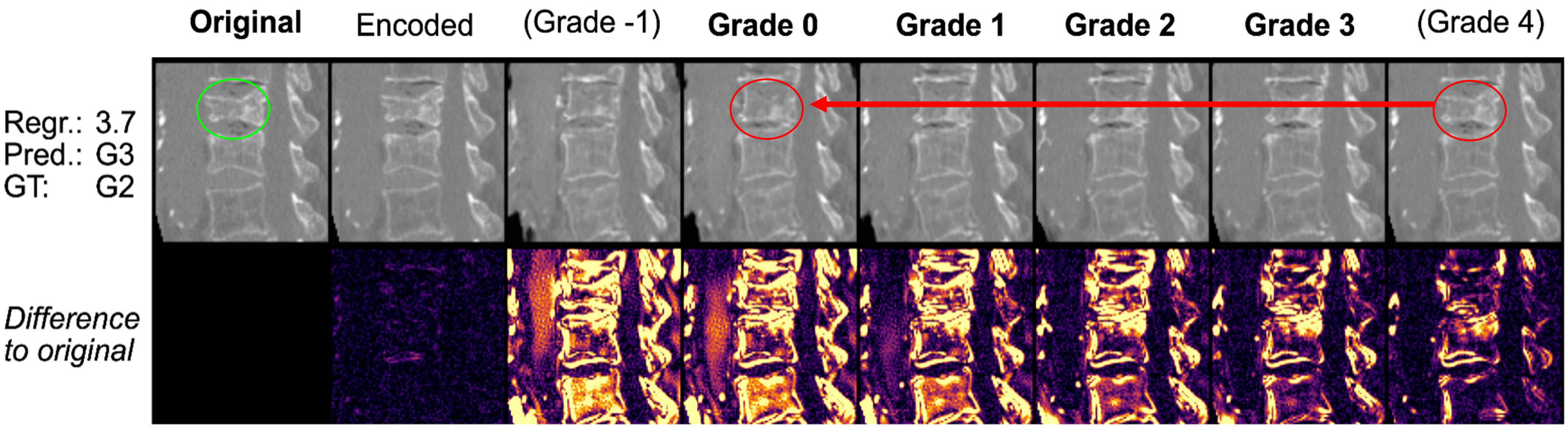
Model	Encoder	Classifier	Detection (AUC ↑)	Grading (F_1 ↑)
Linear probing trained on G0, G2 and G3 with frozen encoder of generative model				
AE [5]	AE	Linear layer	0.70*	-
VAE [5]	VAE	Linear layer	0.77*	-
Spine-VAE [5]	VAE	Linear layer	0.81*	-
StyleGAN2	E4E	SVM	0.74	-
DAE	DAE	Linear layer	0.96	(0.23)
DAE	DAE	SVM	0.93	0.59
Linear regression with distance to hyperplane, calibrated with means of G0 and G3				
DAE	DAE	Linear layer	0.96	0.44
DAE	DAE	SVM	0.93	0.51
Polynomial regression with distance to hyperplane, calibrated with G0, G2 and G3				
DAE	DAE	SVM (deg=1)	0.93	0.42
DAE	DAE	SVM (deg=3)	0.93	0.56
End-to-end training with full supervision (G0, G2 and G3)				
DenseNet121, baseline			0.98	0.65
3D SE-ResNet50 with SupCon loss [4]			0.99*	0.86*
DAE Semantic Encoder, fine-tuned			0.96	0.44

Model bias: Hallucination of lung artifact



→ Associates "healthiness" with thoracic vertebrae in the higher spine region

Model bias: adjacent vertebra changes



→ Attends to the vertebrae above

Conclusion

Physiology-inspired **continuous regression** of fractures

Successful generation of **ordinal counterfactuals** – a great tool for communication!

Unsupervised feature extraction allows the addition of more unlabeled training data.

Future work:

Extension to 3D volumes

Apply non-linear, invertible models on semantic latent space

Predict progression, add longitudinal data

Semantic Latent Space Regression of Diffusion Autoencoders for Vertebral Fracture Grading

Matthias Keicher[1*], Matan Atad[1*], MD David Schinz[2], MD Alexandra S. Gersing[2,3], MD Sarah C. Foreman[2], MD Sophia S. Goller[3], MD Juergen Weissinger[3], MD Jon Rischewski[3], MD Anna-Sophia Dietrich[2], MD Benedikt Wiestler[2], MD Jan S. Kirschke[2], and Nassir Navab[1]

MICCAI2023 Vancouver CANADA

Klinikum rechts der Isar
Technische Universität München