Neural Radiance Fields for Building 3D Reconstructions of the Operating Room

BEEREND GERATS | TECHMED EVENT | 01-11-2023

meander MEDISCH CENTRUM

UNIVERSITY OF TWENTE.

RAM

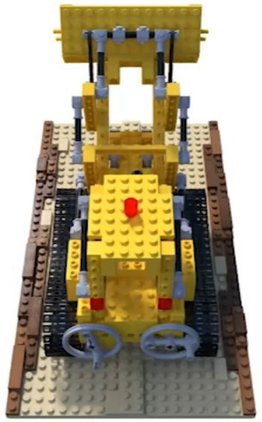
Mathematics of Imaging & AI

NeRF

NeRF

NeRF

NeRF
Neural Radiance Fields

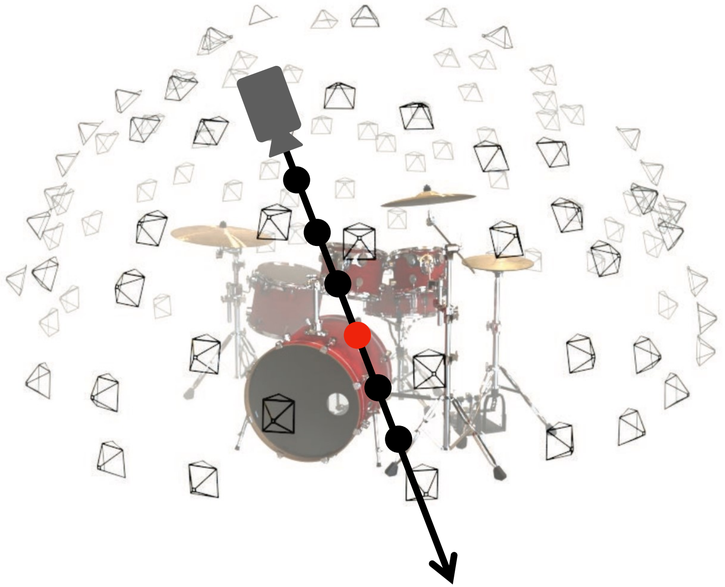


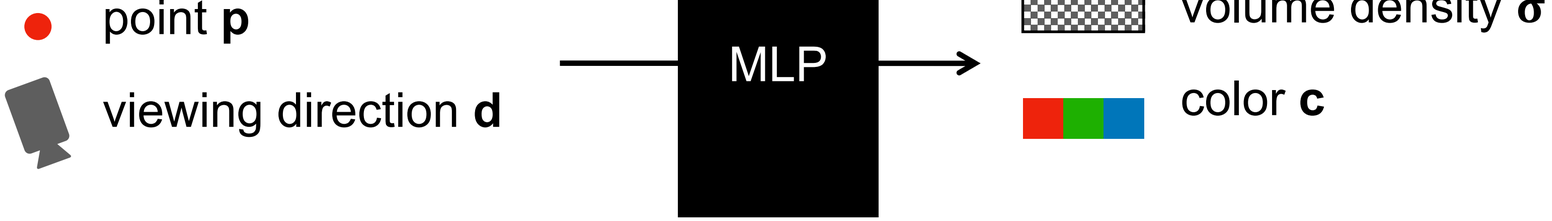


View synthesis with NeRF[1]

1. Mildenhall, B., et al. NeRF: representing scenes as neural radiance fields for view synthesis, ECCV 2021.

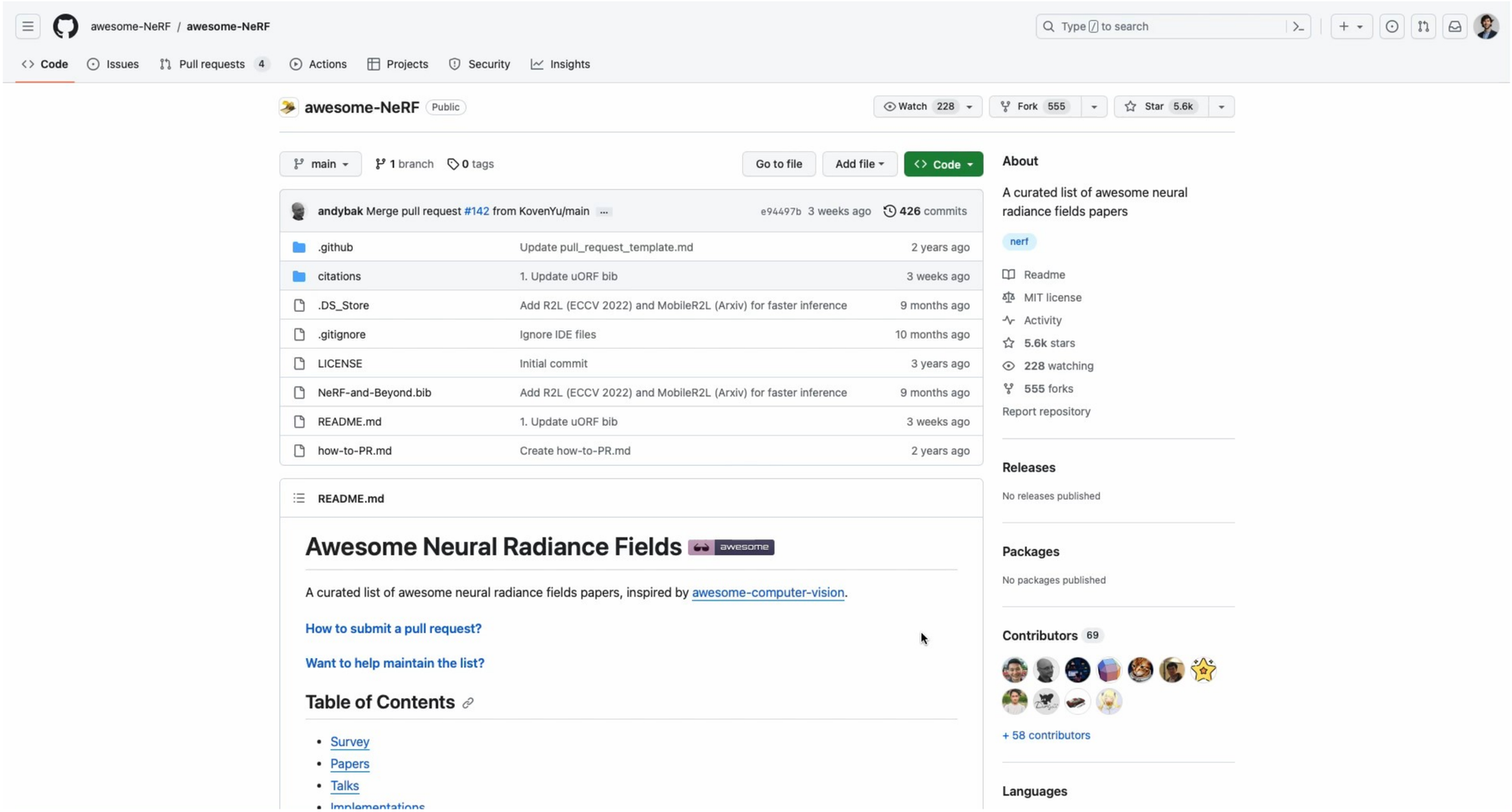
View synthesis with NeRF[1]





1. Mildenhall, B., et al. NeRF: representing scenes as neural radiance fields for view synthesis, ECCV 2021.

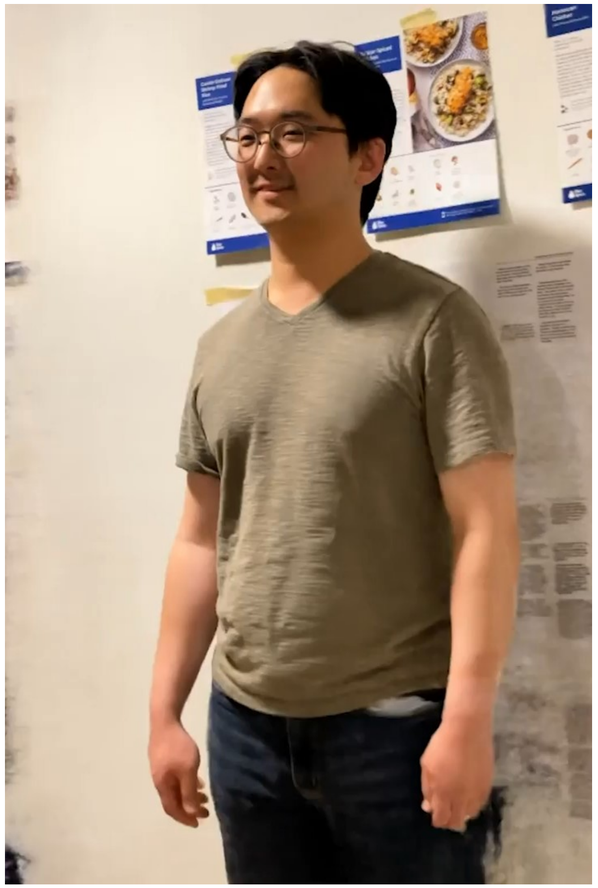
The “NeRF explosion”



https://github.com/awesome-NeRF/awesome-NeRF

The “NeRF explosion”

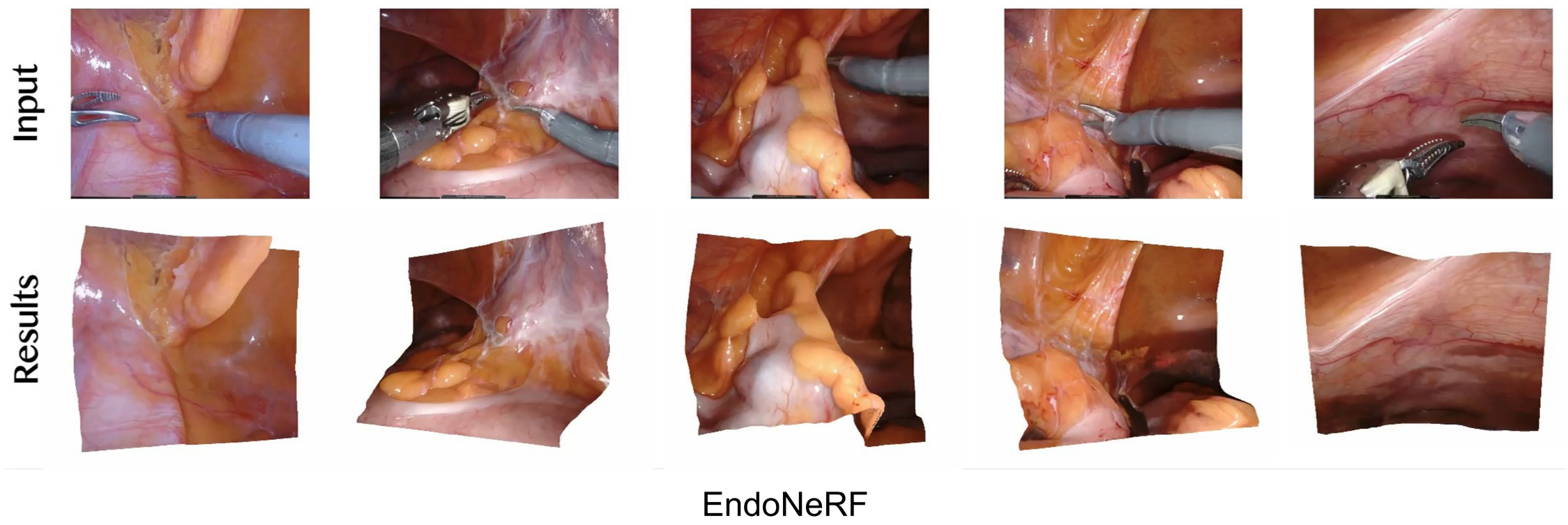
Zip-NeRF[1]



“Nerfie”[2]

1. Barron, J.T., et al., Zip-NeRF: anti-aliased grid-based neural radiance fields, ICCV 2023.
2. Park, K., et al., Nerfies: deformable neural radiance fields, ICCV 2021.

The “NeRF explosion”

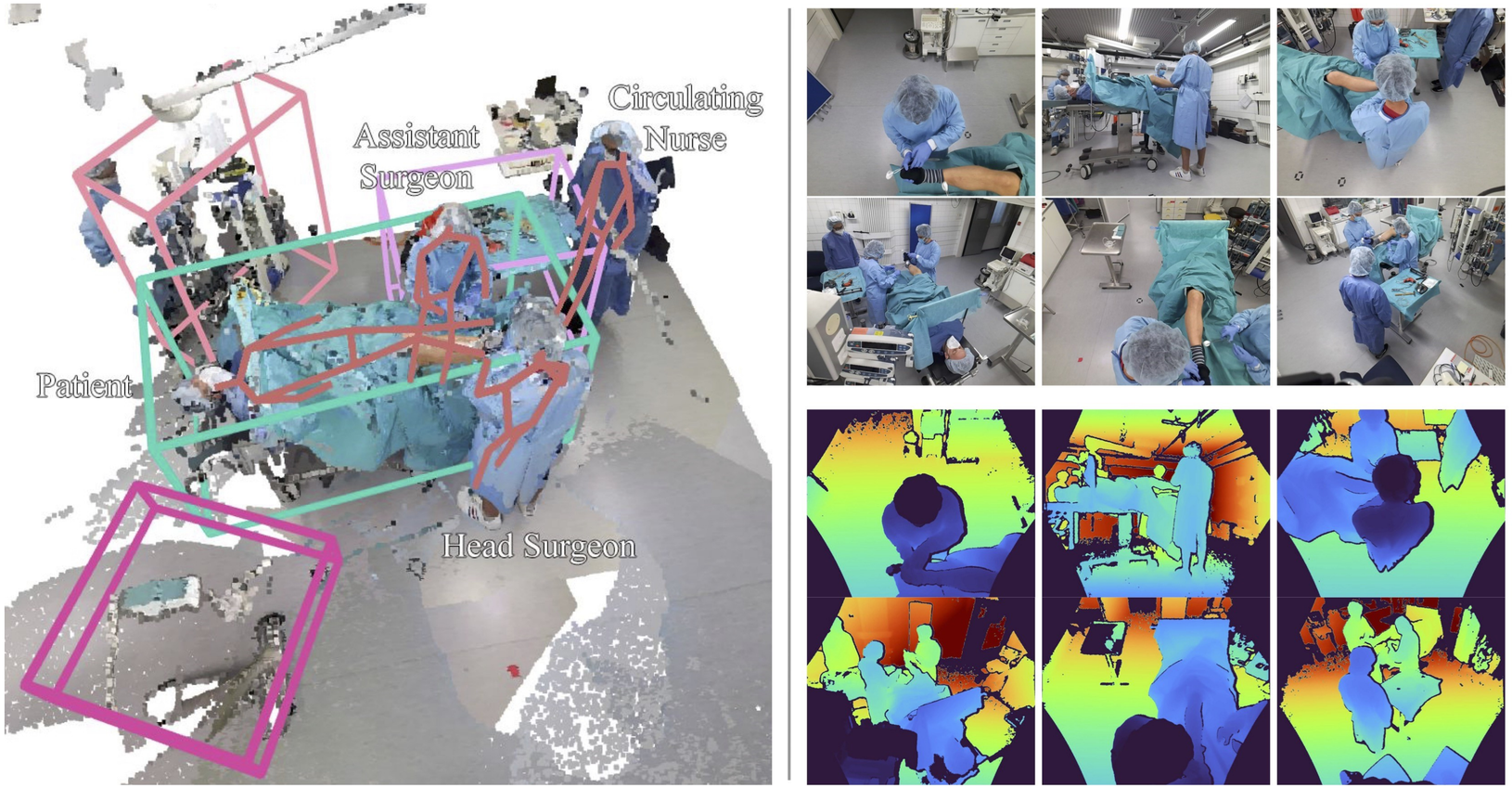


1. Wang, Y., et al., Neural rendering for stereo 3D reconstruction of deformable tissues in robotic surgery, MICCAI 2022.

Our research:

Scene understanding in operating room videos

Scene understanding in OR videos



Example tasks:

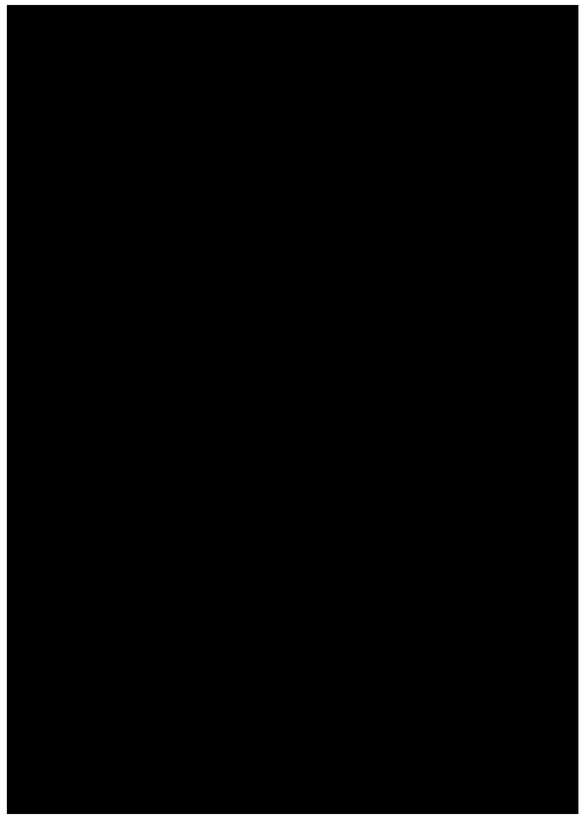
- Object detection
- Human pose estimation
- Domain modeling
- Role identification

Example purposes:

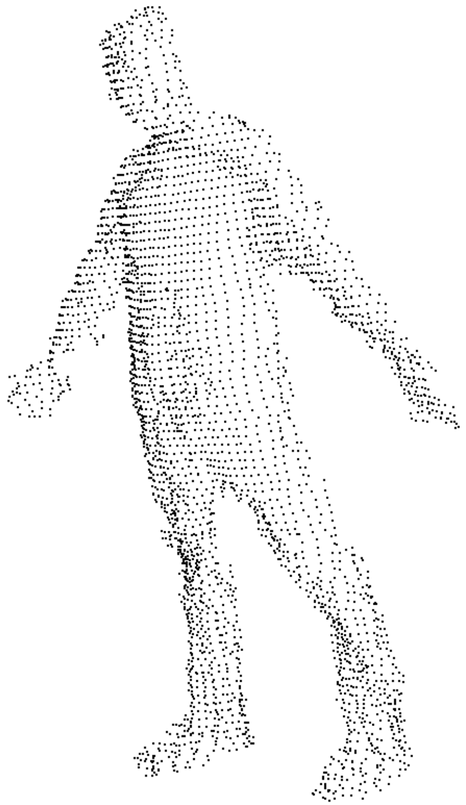
- Register adverse events and distractions
- Monitor X-ray radiation exposure

Image: Özsoy, E., et al. 4D-OR: semantic scene graphs for OR domain modeling, MICCAI 2022.

Algorithms operate on different domains



Images (2D)

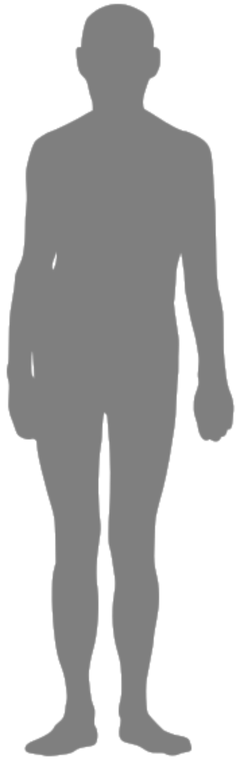


Point clouds (3D)

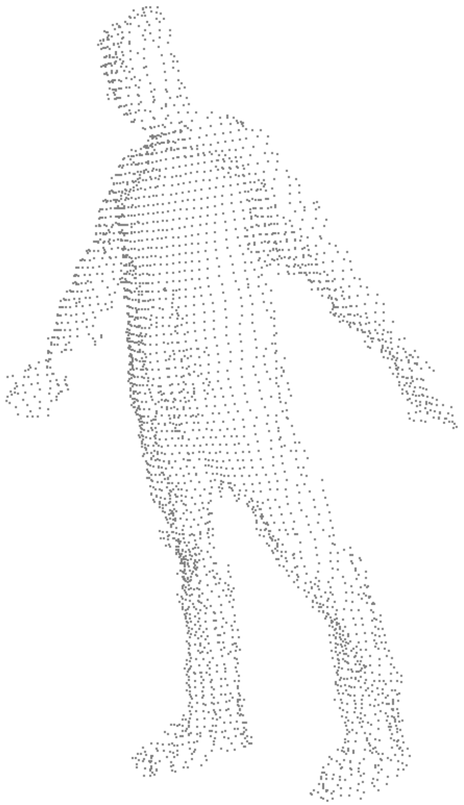
NeRF (3D)[1]

1. Işık, M., et al., HumanRF: high-fidelity neural radiance fields for humans in motion, ArXiv 2023.

Algorithms operate on different domains



Images (2D)

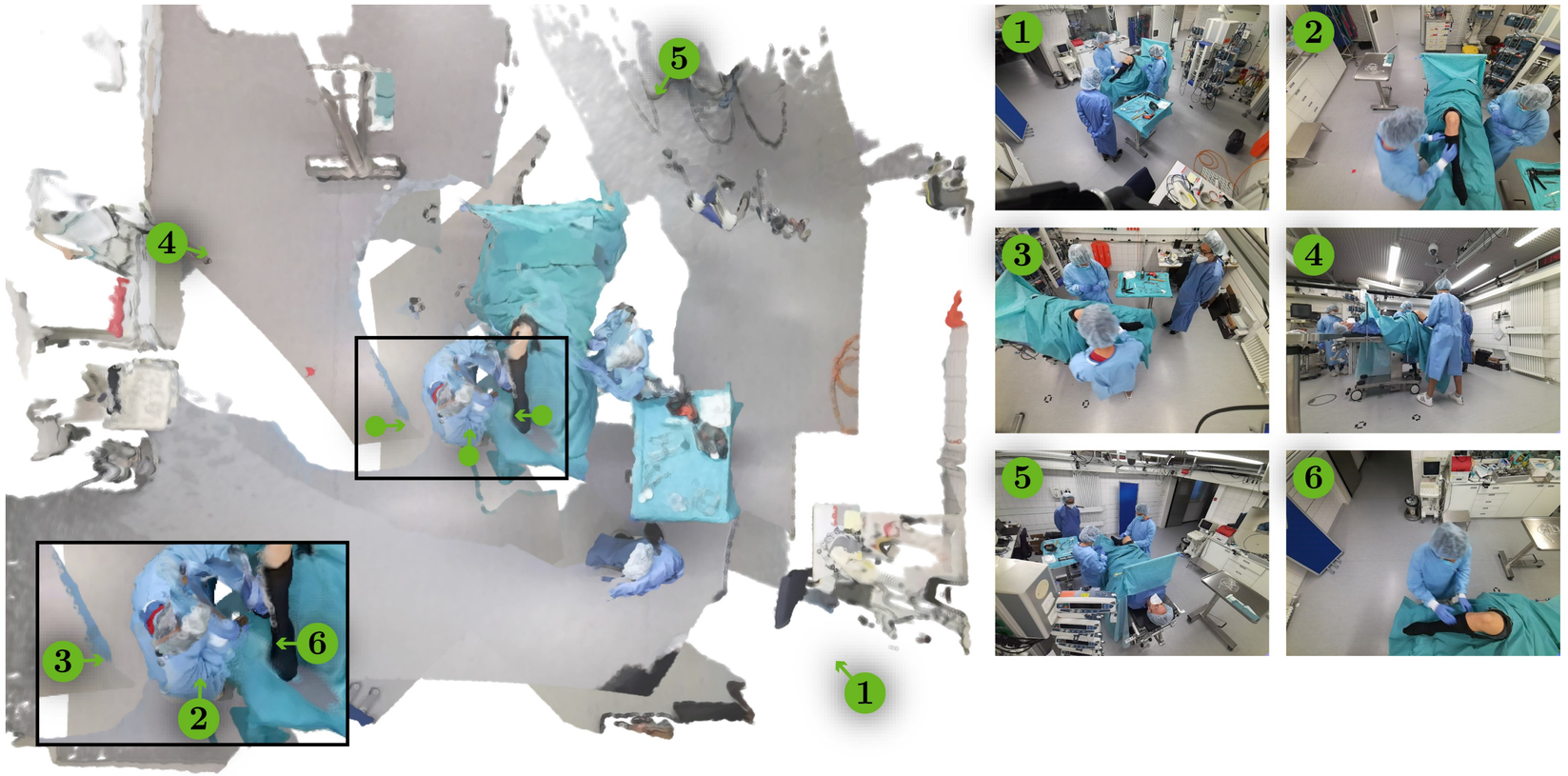


Point clouds (3D)

NeRF (3D)[1]

1. Işık, M., et al., HumanRF: high-fidelity neural radiance fields for humans in motion, ArXiv 2023.

NeRF for operating room videos



4D-OR Dataset[1]:

Three cameras around surgical field

Three distant cameras

RGB-D imaging

1. Özsoy, E., et al. 4D-OR: semantic scene graphs for OR domain modeling, MICCAI 2022.

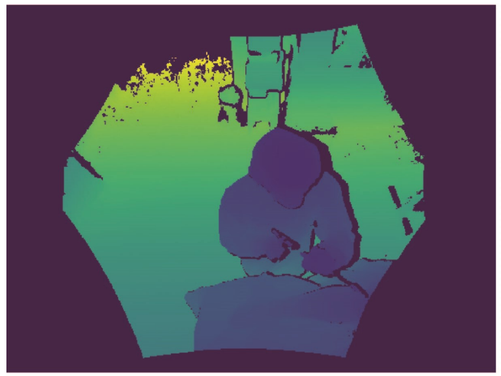
Dynamic depth-supervised NeRF

Depth-Supervision

Use regularisation from DS-NeRF[1]

Additional depth loss term

True depth values from RGB-D sensors



Ground truth

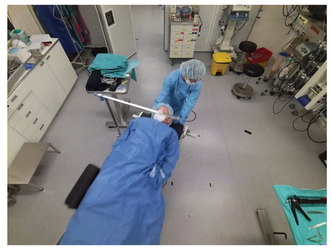
Predicted depth

Dynamics

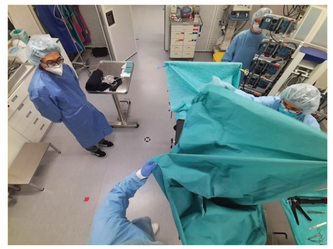
Inspired by DyNeRF[2]

Additional input variable **t**

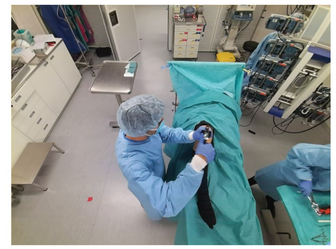
View synthesis dependent on time variable



t = 100s



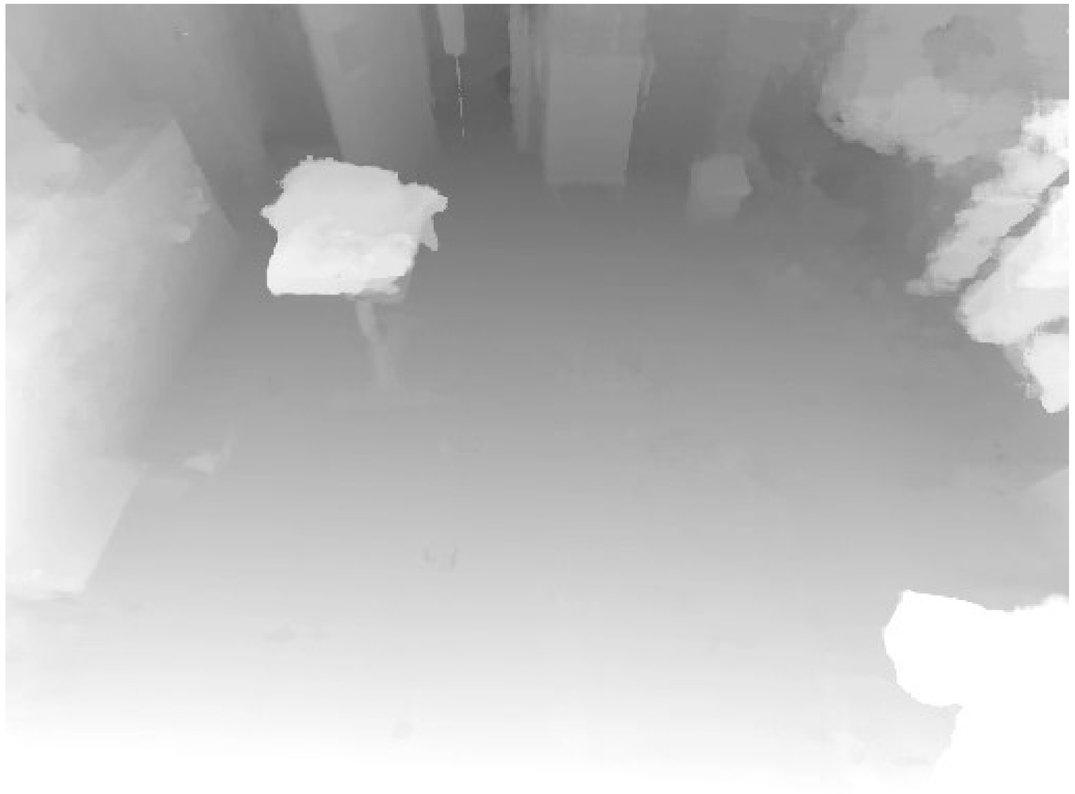
t = 200s



t = 300s

1. Deng, K., et al. Depth-supervised NeRF: fewer views and faster training for free, CVPR 2022.
2. Li, T., et al. Neural 3D video synthesis from multi-view video, CVPR 2022.

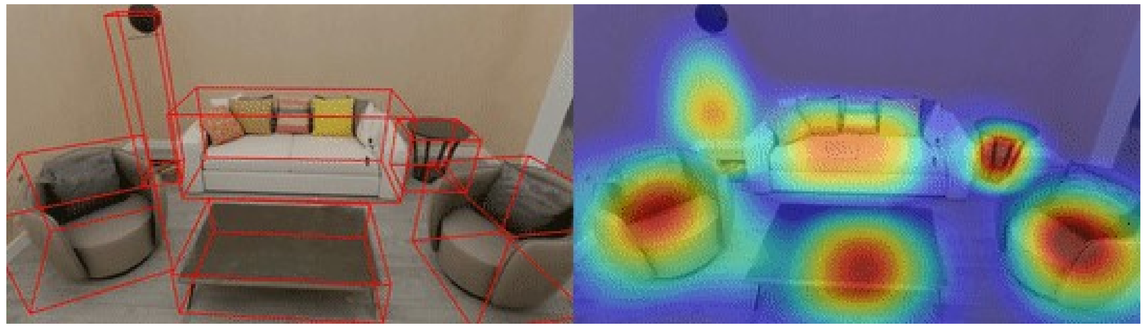
Qualitative results



But now…

What to do with the NeRF representations?

3D continuous data domain for further processing



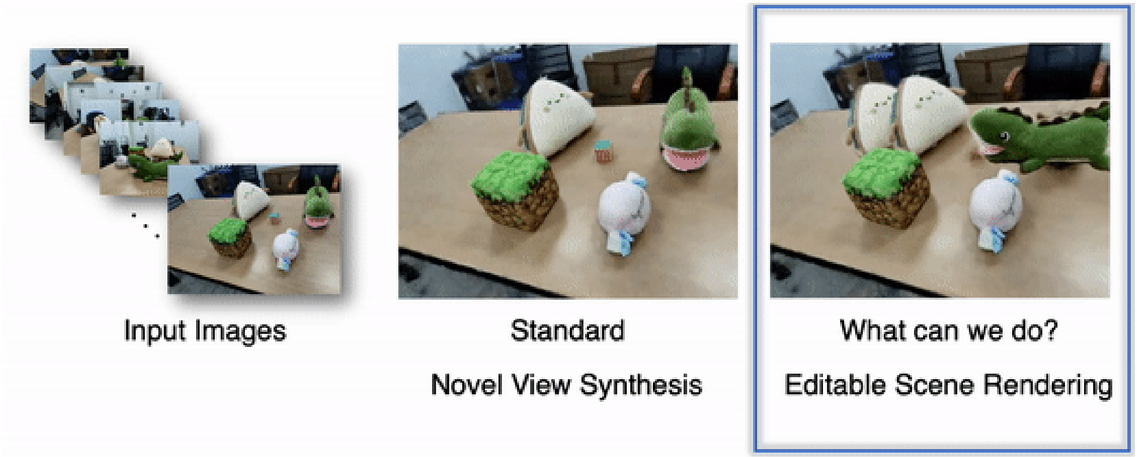
Object detection[1]

Instance segmentation[2]

1. Hu, B., et al., NeRF-RPN: a general framework for object detection in NeRFs, CVPR 2023.
2. Liu, Y., et al., Instance neural radiance field, ICCV 2023.

Data augmentation to a whole new level

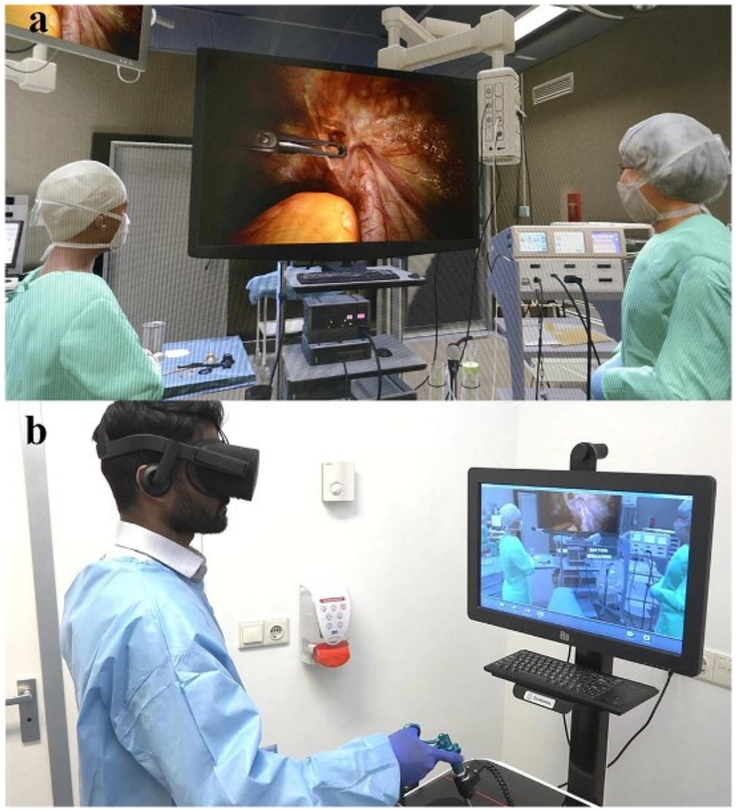
Novel-view synthesis[1]



Compositional NeRFs[2]

1. Zhou, Al, et al., NeRF in the palm of your hand: corrective augmentation for robotics via novel-view synthesis, CVPR 2023.
2. Yang, B., et al., Learning object-compositional neural radiance field for editable scene rendering, ICCV 2021.

Virtual presence in the OR



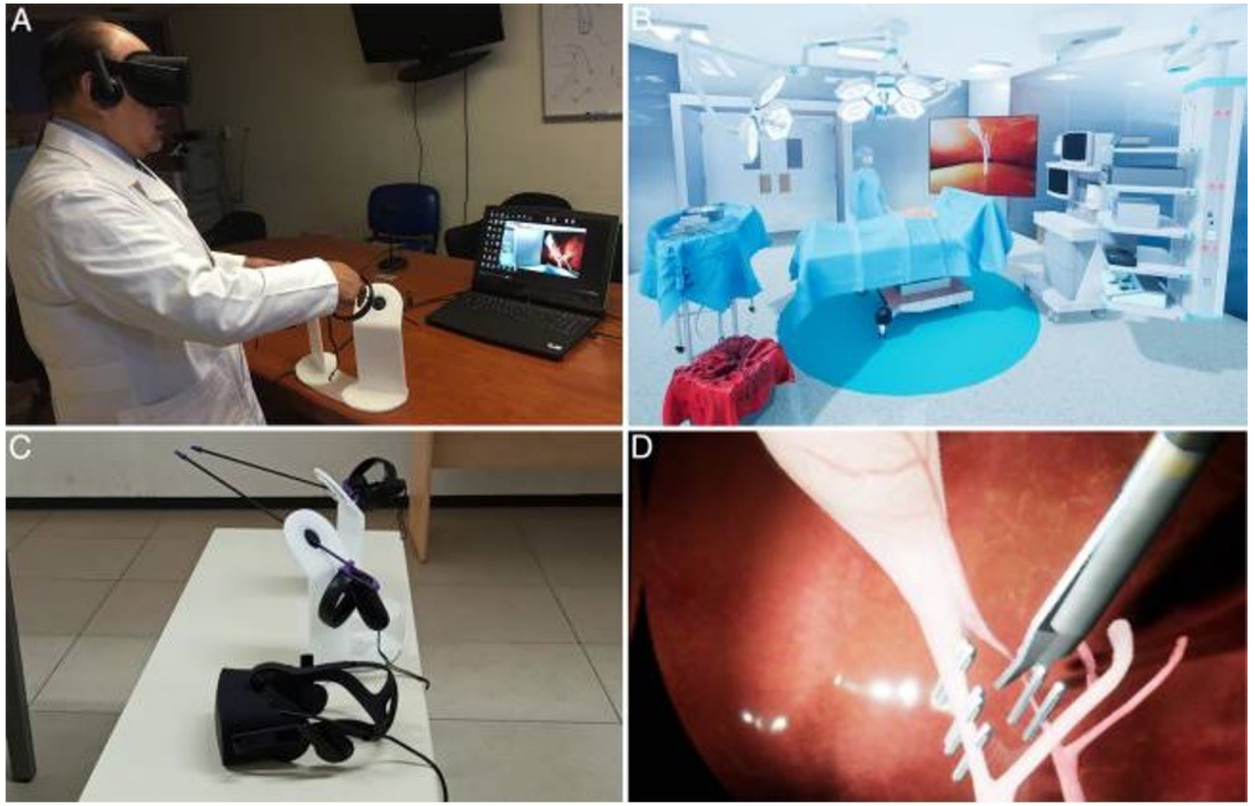
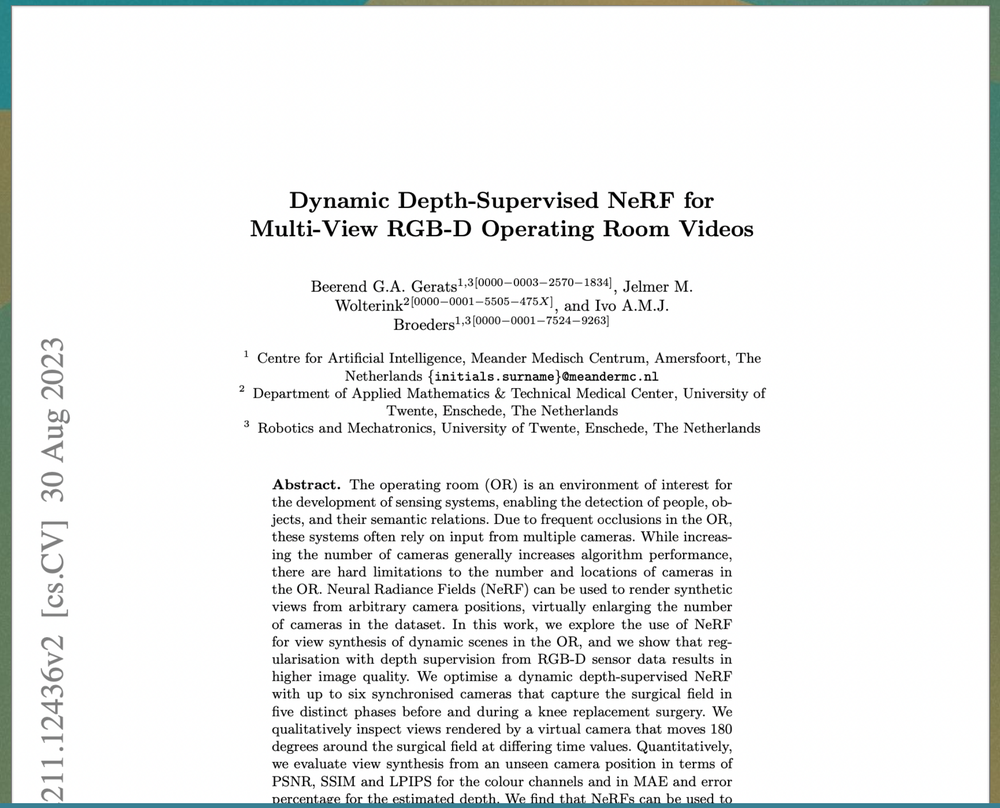


Image (left): Ganni, S., et al., Virtual operating room simulation setup (VORSS) for procedural training in minimally invasive surgery - a pilot study, Indian J. of Surg. 2020. Image (right): Pérez-Escamirosa, F., et al., Immersive virtual operating room for surgical resident education during COVID-19, Surg. Innovation 2020.

Thank you listening!
Any Questions?

BGA.GERATS@MEANDERMC.NL

Find our paper on arXiv:
https://arxiv.org/abs/2211.12436



Dynamic Depth-Supervised NeRF for Multi-View RGB-D Operating Room Videos

Beerend G.A. Gerats[1,3][0000−0003−2570−1834], Jelmer M. Wolterink[2][0000−0001−5505−475X], and Ivo A.M.J. Broeders[1,3][0000−0001−7524−9263]

[1] Centre for Artificial Intelligence, Meander Medisch Centrum, Amersfoort, The Netherlands {initials.surname}@meandermc.nl
[2] Department of Applied Mathematics & Technical Medical Center, University of Twente, Enschede, The Netherlands
[3] Robotics and Mechatronics, University of Twente, Enschede, The Netherlands

Abstract. The operating room (OR) is an environment of interest for the development of sensing systems, enabling the detection of people, objects, and their semantic relations. Due to frequent occlusions in the OR, these systems often rely on input from multiple cameras. While increasing the number of cameras generally increases algorithm performance, there are hard limitations to the number and locations of cameras in the OR. Neural Radiance Fields (NeRF) can be used to render synthetic views from arbitrary camera positions, virtually enlarging the number of cameras in the dataset. In this work, we explore the use of NeRF for view synthesis of dynamic scenes in the OR, and we show that regularisation with depth supervision from RGB-D sensor data results in higher image quality. We optimise a dynamic depth-supervised NeRF with up to six synchronised cameras that capture the surgical field in five distinct phases before and during a knee replacement surgery. We qualitatively inspect views rendered by a virtual camera that moves 180 degrees around the surgical field at differing time values. Quantitatively, we evaluate view synthesis from an unseen camera position in terms of PSNR, SSIM and LPIPS for the colour channels and in MAE and error percentage for the estimated depth. We find that NeRFs can be used to

1 Introduction