University of Zurich[UZH] | USZ Universitäts Spital Zürich

Department of Neurosurgery

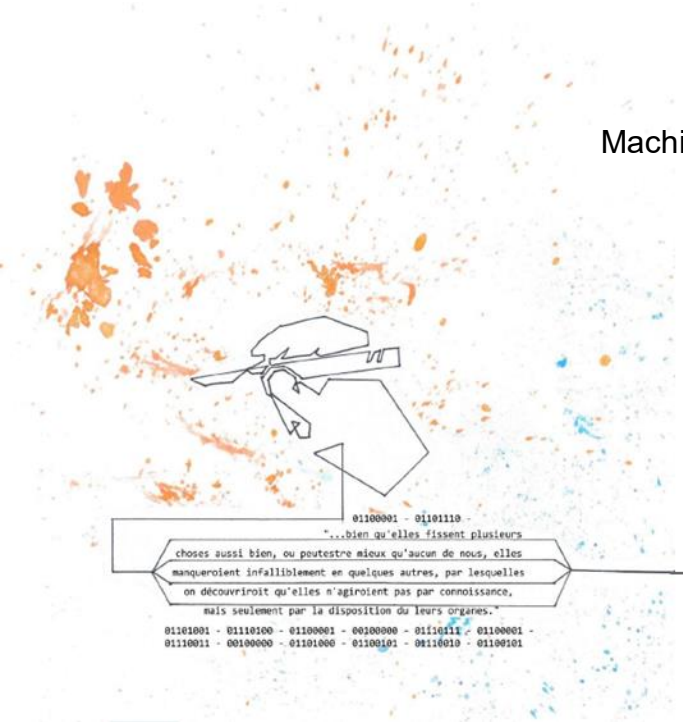
Klinisches Neurozentrum

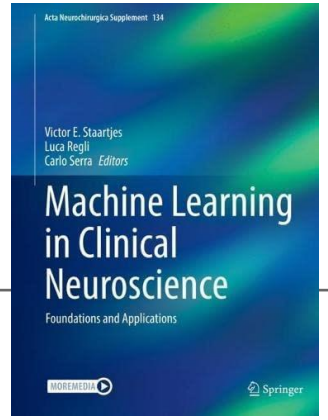
MICN Lab

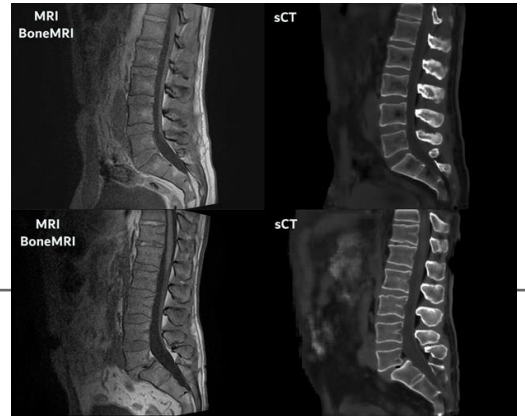
Augmenting Intraoperative Imaging in Surgery with Machine Learning

Friday, May 23rd, 2025 – 1:30 to 2:00 p.m.
Symposium: Intraoperative Workflow, Solutions, and Innovations
Karolinska Universitetssjukhuset
Stockholm, Sweden

Victor Staartjes, MD, PhD
Group Leader
Machine Intelligence in Clinical Neuroscience & Microsurgical Neuroanatomy (MICN) Laboratory
Neurosurgery Resident
Capio Spine Center Stockholm, Upplands Väsby, Sweden
Department of Neurosurgery, University Hospital Zurich, University of Zurich







No Conflicts of Interest relevant to presented content

Funding Disclosures:

- Research RELATED to presentation funded by:
  - Medtronic: BoneMRI
  - Swiss National Science Foundation (SNSF): AENEAS
  - Bangerter-Rhyner Stiftung; Max Cloetta Stiftung: TomoRay
- Research NOT RELATED to presentation funded by:
  - n/a

2D vs. 3D - Operative Visualization


Microneurosurgery

in 4 Volumes

M.G.Yasargil

Stereoscopic Perspective

A few neurosurgical procedures such as microvascular anastomoses and nerve repairs are performed on the surface of the operating field, and in these the magnification and depth of field are primary considerations. However, most neurosurgical operations take place in a small space at the base of the brain through a narrow gap and in these cases it is more important to the neurosurgeon that he maintains well-lit binocular vision in the recesses of the field. This stereoscopic perspective is thus the more useful function of the surgical microscope in these situations (Fig 190A).The operating microscope allows stereoscopic vision in small spaces by reducing the necessary interpupillary distance required for binocular vision. The distance between the anterior lenses of the binocular tube of the microscope is only 16 mm, whereas the average interpupillary distance is around 60 mm. This means that light reflected from deep basal structures towards the operating microscope during surgical procedures employing fissure, sulci or transcortical approaches, will result in a stereoscopic image when only a 16 mm image enters the microscope aided eye. Even when assisted with magnification loupes, the eyes are unable to maintain stereoscopic vision in such a narrow space. Thus the real importance of the surgical microscope as it relates to most neurological procedures is not the magnification it supplies, but in the clear visual perspective it provides. With this the surgeon can avoid excessive brain retraction and yet still reach every point in the central nervous system, adequately visualizing deep structures either along the basal cisterns or through a transcerebral tunnel.

2D vs. 3D - Microscopy

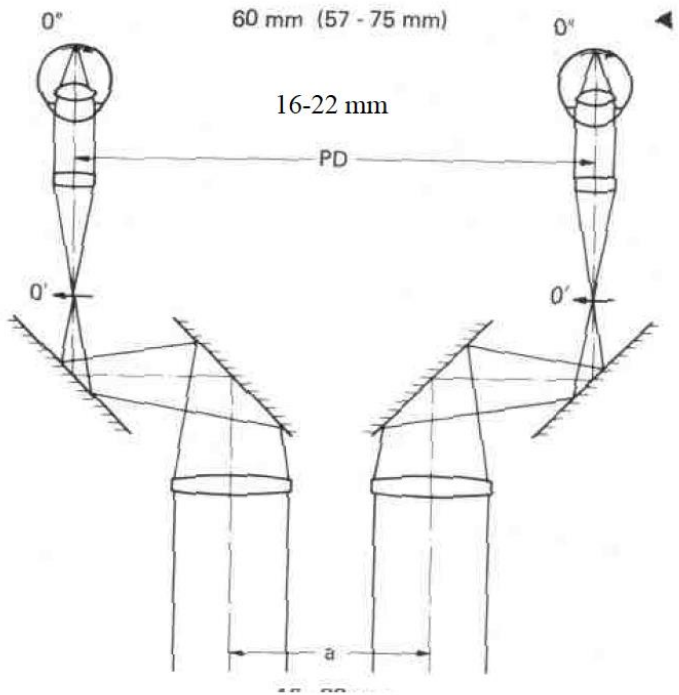


Fig 190 A Diagram of the difference I pillary distance (PD) and interocular erating microscope (a).

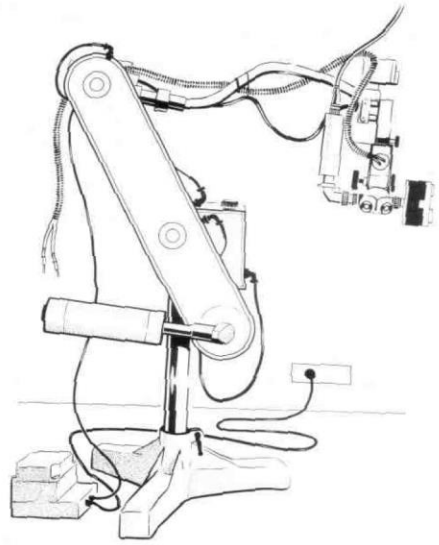


Fig **190B** The prototype counterbalanced Contraves stand for the operating microscope in use at the University Hospital of Zurich since 1972.

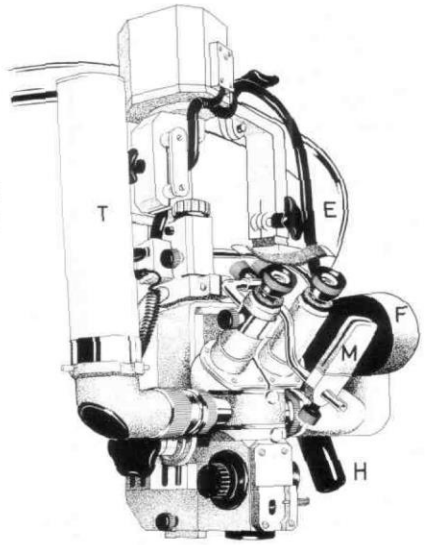
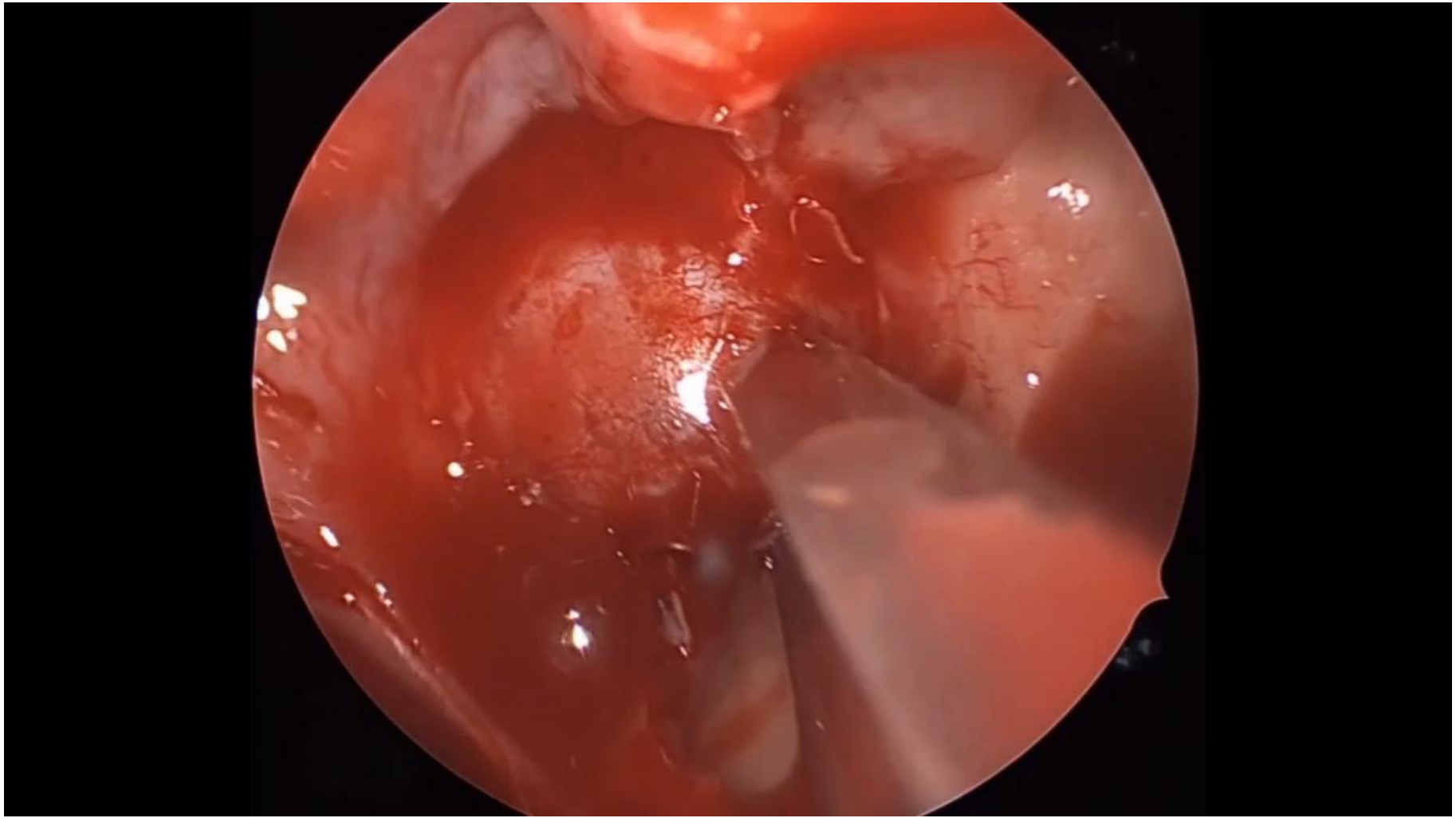


Fig 190C H = hand-switch, M = mouth-switch, F = photo- or movie camera, T = Hitachi colour TV camera now available with three tubes.

2D vs. 3D – Endoscopy in 2D

→ experienced surgeons compensate for lack of depth



What if we could grant stereoscopic vision in endoscopy?

→ without a 3D endoscope
→ without specialized hardware, costs, etc.
→ power of software / machine vision

Approaches to the Problem:

- **Multi-View Geometry** such as **Structure-From-Motion** (SFM or SLAM algorithms) ≅ photogrammetry
  - Computationally expensive and difficult to execute in real time
- **Photometry / "Depth-From-Shading"** based on lighting changes
  - Assumes stable lighting conditions
- **Deep Learning**-based methods
  - Computationally expensive to develop but applicable in real-time
  - Can deal with changing lighting conditions

Zanier, Serra, Staartjes; *Real-time intraoperative depth estimation in transsphenoidal surgery using deep learning: a pilot study*, submitted (2025)

DinoV2 in Action: Still Images

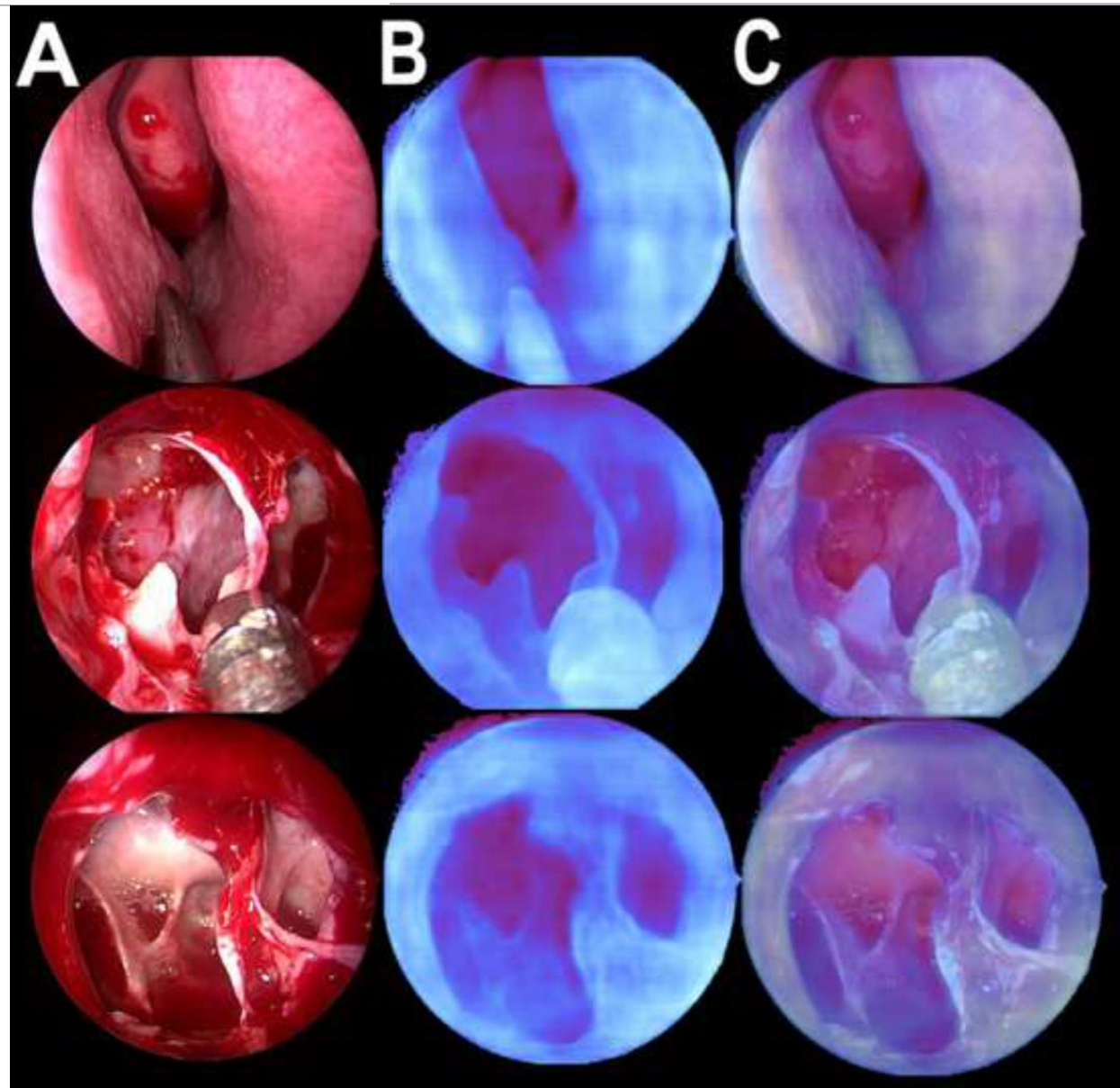
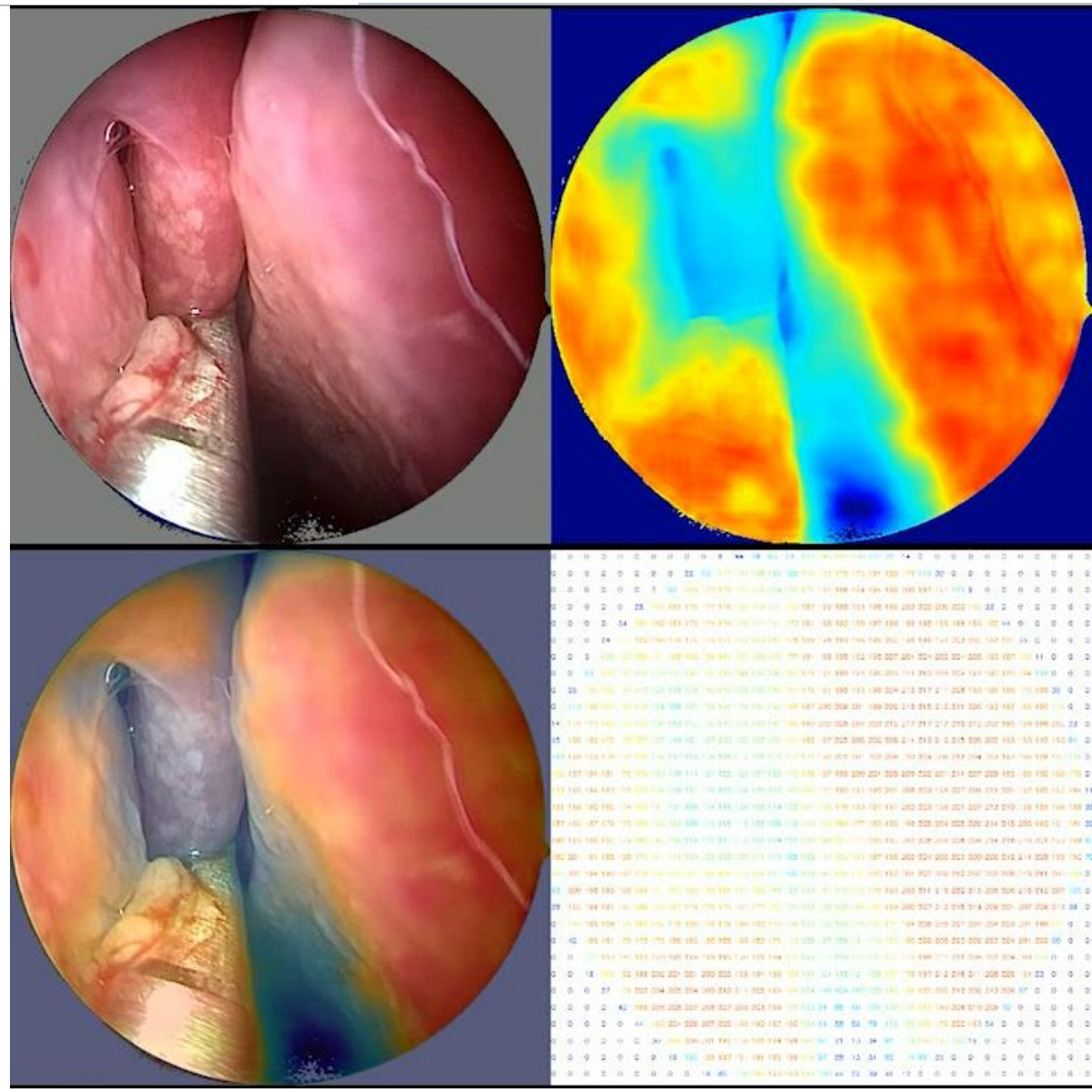


Figure 1: Example frames of DinoV2 running on a surgical video.

A: input frame, B: depth map as calculated by the model, C: overlay of A and B.

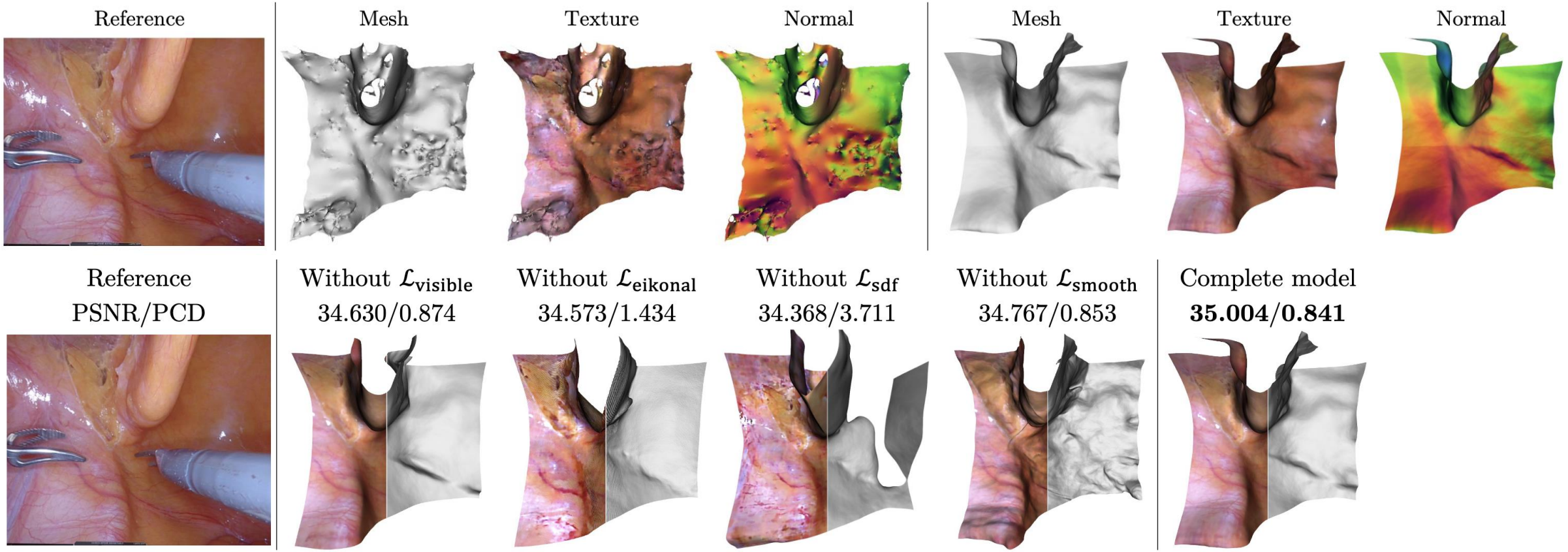
Zanier, Serra, Staartjes; *Real-time intraoperative depth estimation in transsphenoidal surgery using deep learning: a pilot study*, submitted (2025)

DinoV2 in Action: Real-Time Render



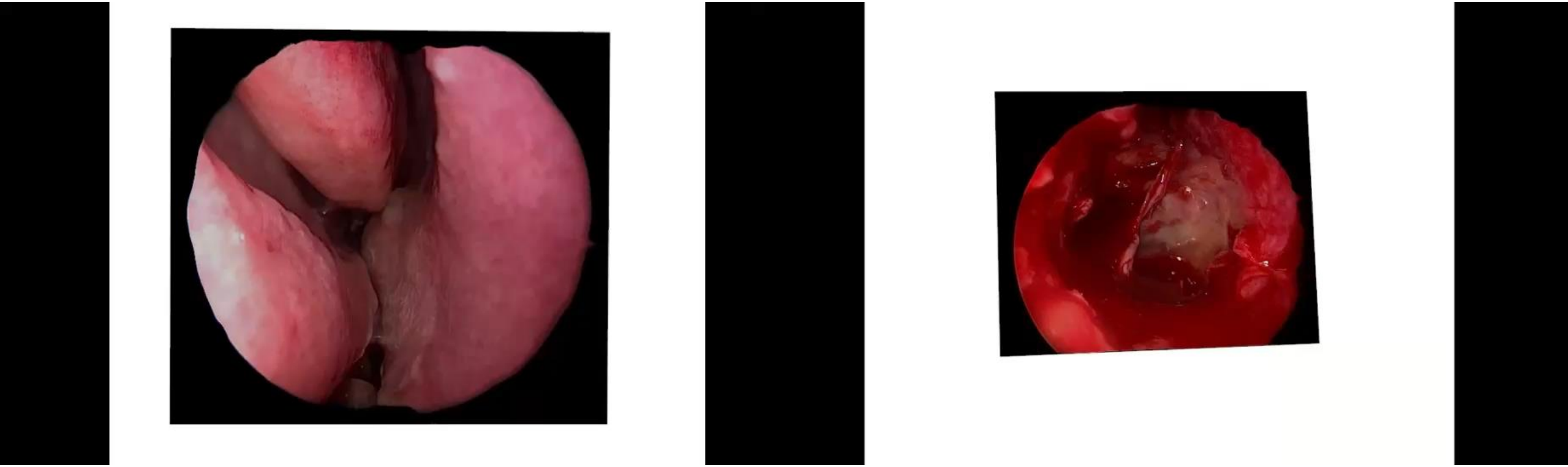
Zanier, Serra, Staartjes; *Real-time intraoperative depth estimation in transsphenoidal surgery using deep learning: a pilot study*, submitted (2025)

How to get from a synthetic 3D image to binocular vision?
→ e.g. EndoSurf 3D → 3D glasses or AR/VR



Zha et al. *EndoSurf: Neural Surface Reconstruction of Deformable Tissues with Stereo Endoscope Videos*, arXiv 2023

How to get from a synthetic 3D image to binocular vision?
→ 3D reconstructions from 2D endoscopy



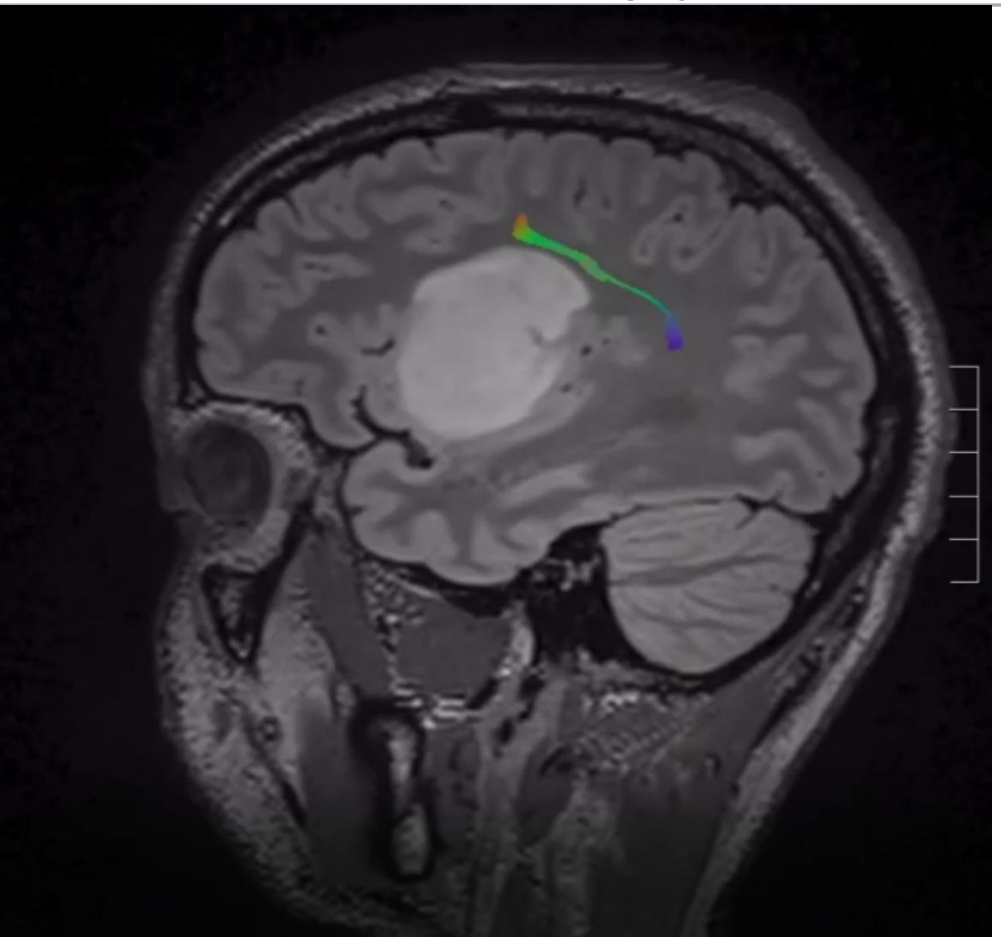
University of Zurich[UZH] | USZ Universitäts Spital Zürich

Department of Neurosurgery

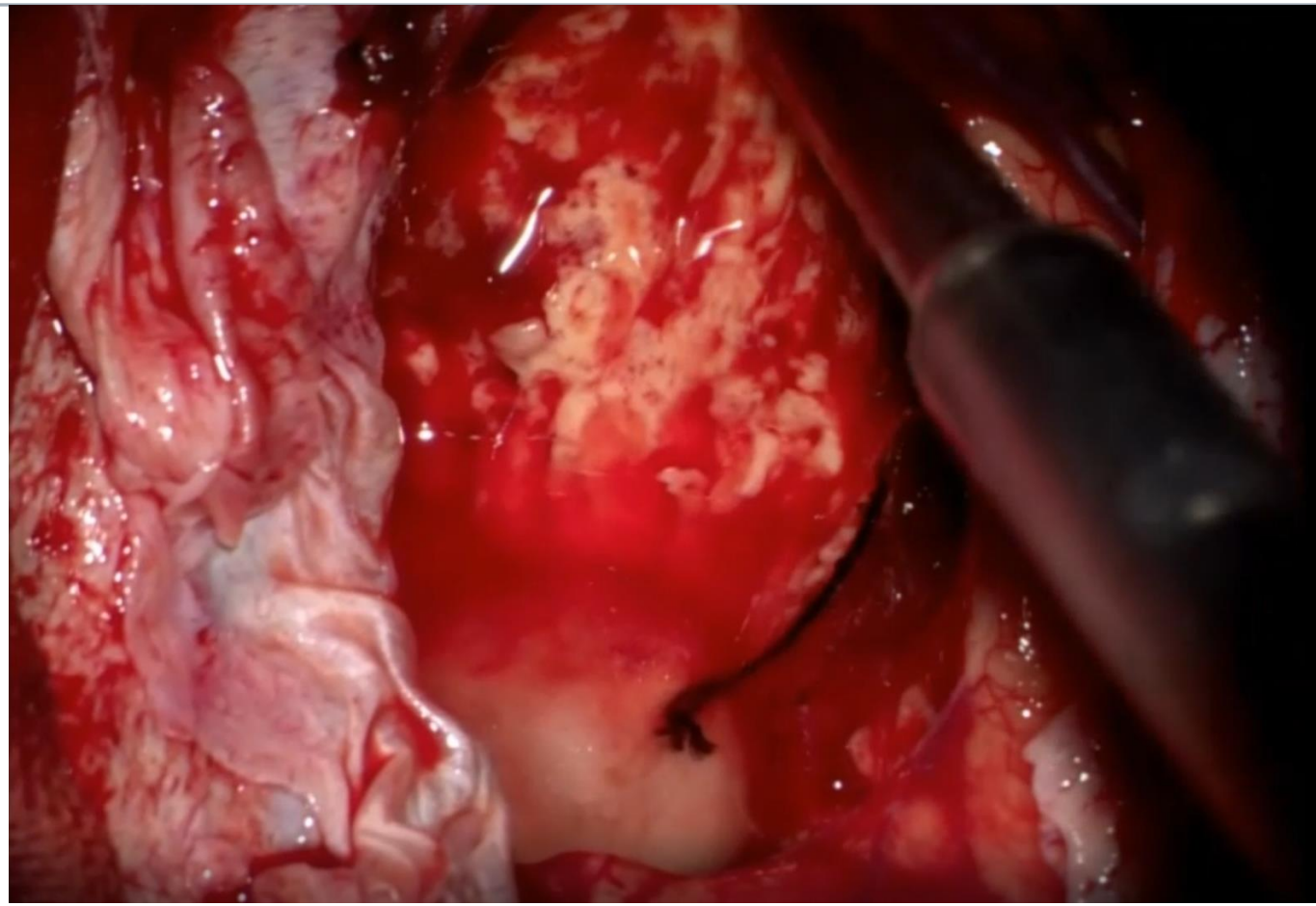
Project AENEAS
Anatomical Orientation

ETH Eidgenössische Technische Hochschule Zürich Swiss Federal Institute of Technology Zurich

MICN Lab



Clarity



Reality

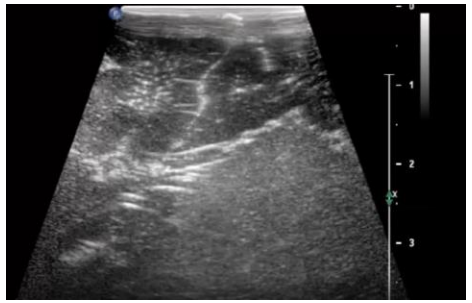
Anatomical Orientation

University of Zurich[UZH] | USZ Universitäts Spital Zürich

Department of Neurosurgery

ETH Eidgenössische Technische Hochschule Zürich Swiss Federal Institute of Technology Zurich | MICN Lab

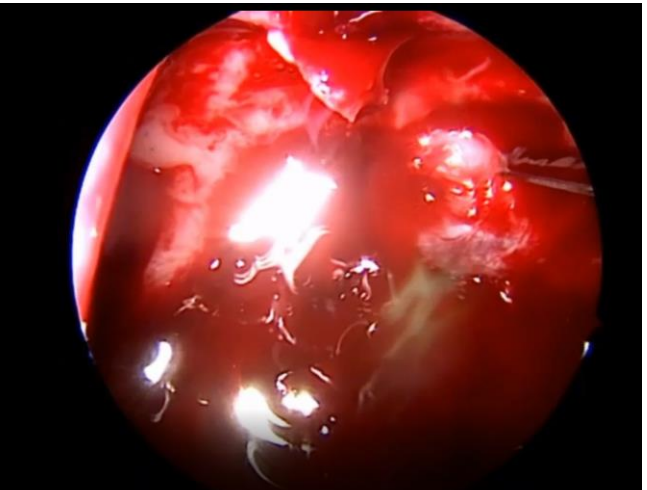
1. The Problem



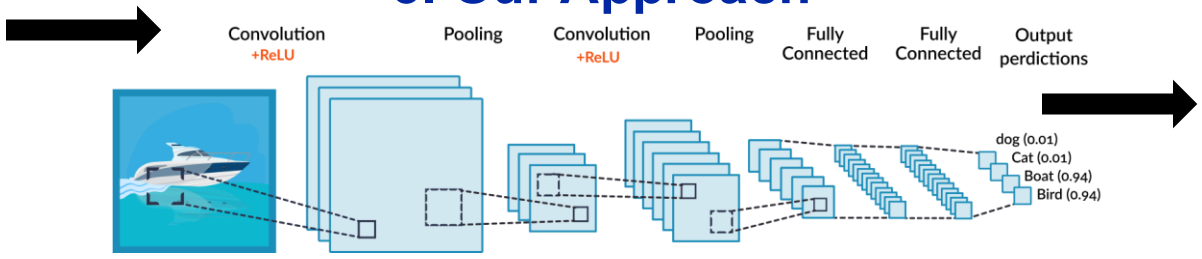
2. The Solution

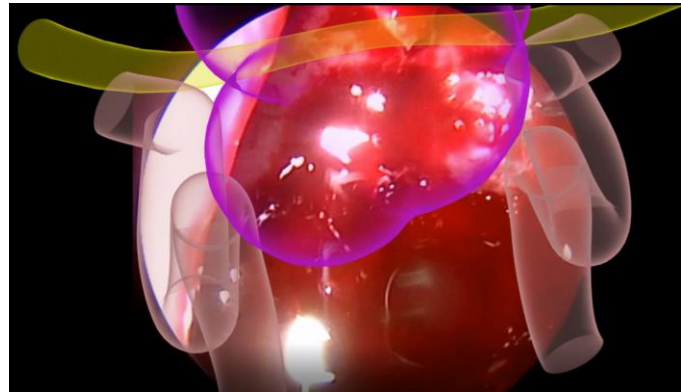
Automatically detect anatomy in real-time:

- Eliminate preop dependence/brain shift
- Eliminate costs
- Eliminate logistics
- Enhanced user-variability/specificity problems



3. Our Approach





AENEAS Concept

Without reliance on prior anatomical knowledge/maps!

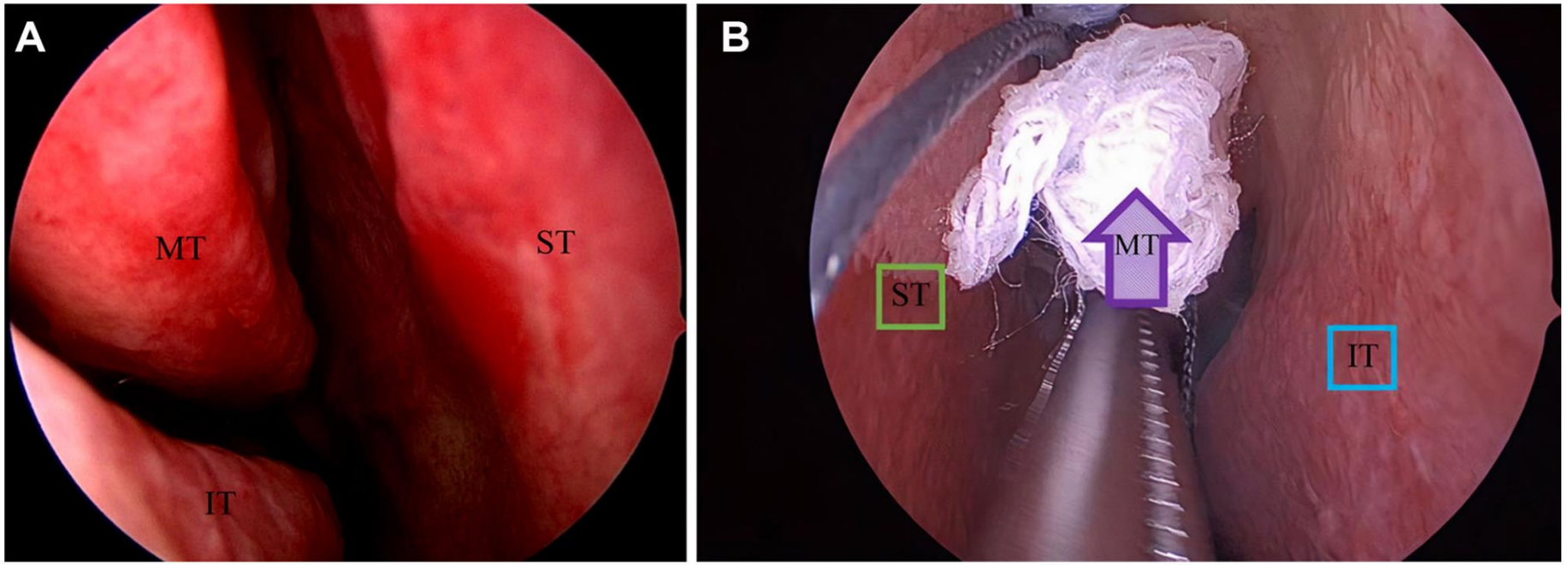


FIGURE 1. **A**, *Demonstrates an endoscopic view of the endonasal anatomy, specifically the middle and inferior turbinates as well as the nasal septum. An experienced endoscopic pituitary surgeon labeled these 3 structures once each second with a single pixel.* **B**, *Demonstrates the intended output of the machine vision model: anatomic structures in sight such as the septum and inferior turbinate are identified and marked by the user interface. In the future, the model is intended to learn to anticipate structures, eg, the middle turbinate hidden behind the cotton patty. MT, middle turbinate; IT inferior turbinate; ST, nasal septum.*

Staartjes VE, Volokitin A, Regli L, Konukoglu E, Serra C.
Machine Vision for Real-Time Intraoperative Anatomic Guidance: A Proof-of-Concept Study in Endoscopic Pituitary Surgery.
Operative Neurosurgery. 2021;(opab187). doi:10.1093/ons/opab187

Machine Vision for Real-Time Intraoperative Anatomic Guidance: A Proof-of-Concept Study in Endoscopic Pituitary Surgery

Victor E. Staartjes, BMed *
Anna Volokitin, MEng‡
Luca Regli, MD *
Ender Konukoglu, PhD‡
Carlo Serra, MD*

*Machine Intelligence in Clinical Neuroscience (MICN) Laboratory, Department of Neurosurgery, University Hospital Zurich, Clinical Neuroscience Centre, University of Zurich, Zurich, Switzerland; ‡Computer Vision Lab (CVL), ETH Zurich, Zurich, Switzerland

Correspondence:
Carlo Serra, MD,
Machine Intelligence in Clinical Neuroscience (MICN) Lab,
Department of Neurosurgery,
Clinical Neuroscience Center,
University Hospital Zurich,
University of Zurich,
Frauenklinikstrasse 10,
8091 Zurich, Switzerland.
Email: carlo.serra@usz.ch

Received, November 10, 2020.
Accepted, April 4, 2021.

BACKGROUND: Current intraoperative orientation methods either rely on preoperative imaging, are resource-intensive to implement, or difficult to interpret. Real-time, reliable anatomic recognition would constitute another strong pillar on which neurosurgeons could rest for intraoperative orientation.

OBJECTIVE: To assess the feasibility of machine vision algorithms to identify anatomic structures using only the endoscopic camera without prior explicit anatomo-topographic knowledge in a proof-of-concept study.

METHODS: We developed and validated a deep learning algorithm to detect the nasal septum, the middle turbinate, and the inferior turbinate during endoscopic endonasal approaches based on endoscopy videos from 23 different patients. The model was trained in a weakly supervised manner on 18 and validated on 5 patients. Performance was compared against a baseline consisting of the average positions of the training ground truth labels using a semiquantitative 3-tiered system.

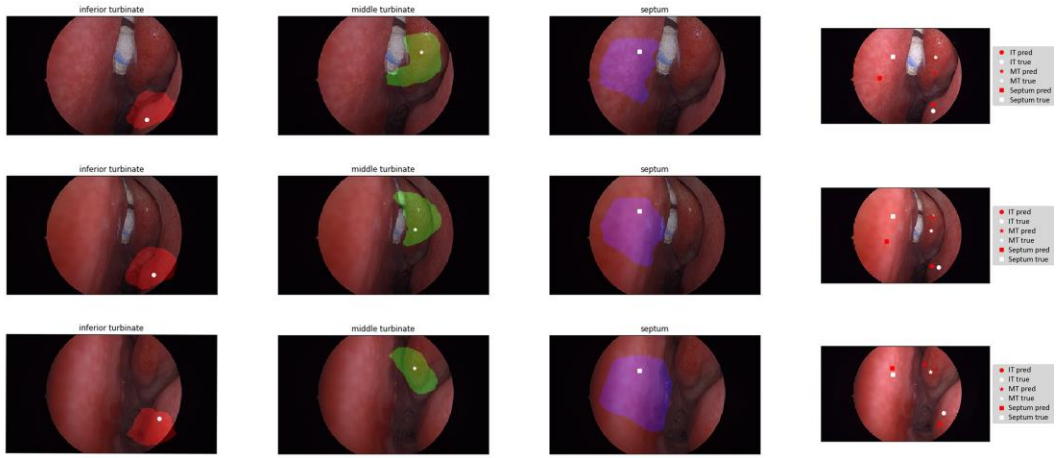
RESULTS: We used 367 images extracted from the videos of 18 patients for training, as well as 182 test images extracted from the videos of another 5 patients for testing the fully developed model. The prototype machine vision algorithm was able to identify the 3 endonasal structures qualitatively well. Compared to the baseline model based on location priors, the algorithm demonstrated slightly but statistically significantly ($P < .001$) improved annotation performance.

CONCLUSION: Automated recognition of anatomic structures in endoscopic videos by means of a machine vision model using only the endoscopic camera without prior explicit anatomo-topographic knowledge is feasible. This proof of concept encourages further development of fully automated software for real-time intraoperative anatomic guidance during surgery.

KEY WORDS: Machine vision, Anatomic recognition, Machine learning, Artificial intelligence, Anatomic guidance

Operative Neurosurgery 0:1–6, 2021 DOI: 10.1093/ons/opab187

Weakly Supervised Deep Learning to Generate Real-Time Anatomical Heatmaps



AENEAS: Roadmap Concept

IMAGING IN NEUROSURGERY

OPEN

AENEAS Project: Live Image-Based Navigation and Roadmap Generation in Endoscopic Neurosurgery Using Machine Vision

Victor E. Staartjes, MD, PhD*, Gary Sarwin, MSc‡, Alessandro Carretta, MD§, Matteo Zoli, MD, PhD §, Diego Mazzatenta, MD, PhD §, Luca Regli, MD*, Ender Konukoglu, PhD‡, Carlo Serra, MD *

*Machine Intelligence in Clinical Neuroscience (MICN) Laboratory, Department of Neurosurgery, University Hospital Zurich, Clinical Neuroscience Centre, University of Zurich, Zurich, Switzerland; ‡Computer Vision Lab (CVL), ETH Zurich, Zurich, Switzerland; §Department of Biomedical and Neuromotor Sciences (DIBINEM), University of Bologna, Bologna, Italy

Part of this publication have been published in the preprint server ArXiv under the name "Live image-based neurosurgical guidance and roadmap generation using unsupervised embedding" 31 March 2023.

Correspondence: Carlo Serra, MD, Machine Intelligence in Clinical Neuroscience (MICN) Laboratory, Department of Neurosurgery, University Hospital Zurich, Clinical Neuroscience Centre, University of Zurich, Frauenklinikstrasse 10, Zurich 8091, Switzerland. Email: carlo.serra@usz.ch

Received, October 08, 2024; **Accepted,** December 26, 2024; **Published Online,** April 28, 2025.

Operative Neurosurgery 00:1–7, 2025 https://doi.org/10.1227/ons.0000000000001583

BACKGROUND AND OBJECTIVES: Artificial intelligence algorithms have proven capable of replicating cognitive processes. Our aim was to replicate human roadmap generation for endoscopic neurosurgery with a live image-based machine vision method.

METHODS: Surgical videos of a highly standardized surgical approach are labeled and used for algorithm training. After object detection (YOLOv7) to generate bounding boxes for landmark anatomical structures, an autoencoder first encodes the currently detected structures into an estimated position within this anatomical roadmap and then enables extrapolation of structures that are expected to be encountered in forward or backward directions. Average precision of the model applied to the test videos at an intersection-over-union threshold of 0.5 is reported.

RESULTS: In total, 166 anonymized endoscopic recording (3×10^6 labeled video frames) were included. We performed model development using 146 videos and held out 20 videos for evaluation (test set). The performance regarding bounding box detection among the 20 test set videos on average was 53.4. Evaluation of the performance of the autoencoder model in detecting the current position within the roadmap of the surgical approach is evaluated semiquantitatively, showing that the first detection of anatomical structures by the model corresponds well to their label distribution along the latent variable encoding the anatomical roadmap. We also provide videos demonstrating the mixed reality head's up display for anatomical navigation.

CONCLUSION: Our method enables reliable identification of key anatomical structures during endoscopic endonasal trans-sphenoidal surgery in mixed reality. Through encoding detected landmark anatomical structures, a surgical roadmap is encoded. This approach allows for detection of visible anatomical structures and enables extrapolation toward the location of those yet to be dissected in deeper anatomical layers. Further development of such algorithms may pave the way toward adding a mixed reality, real-time anatomical navigation software to the neurosurgeon's armamentarium.

KEY WORDS: Machine vision, Anatomical recognition, Machine learning, Artificial intelligence, Anatomical guidance

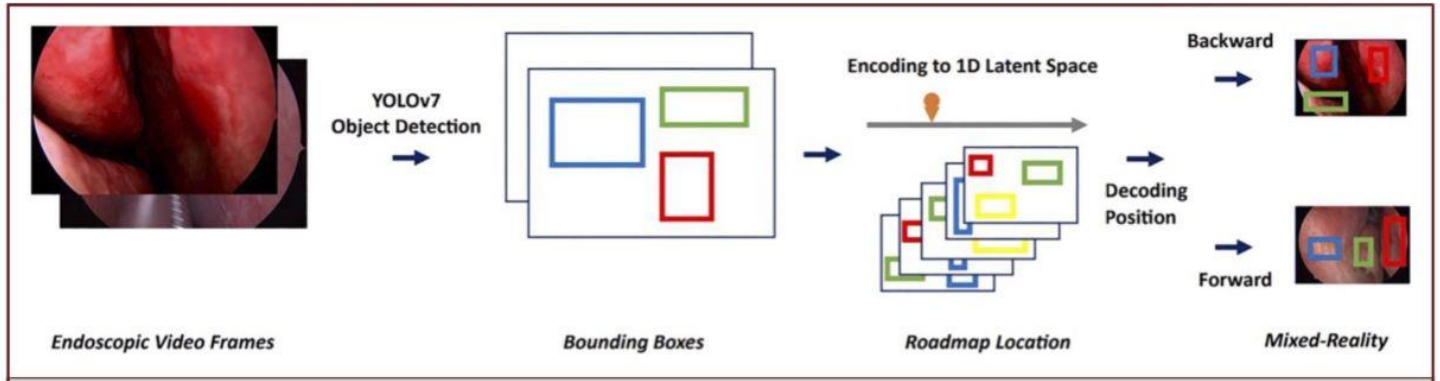
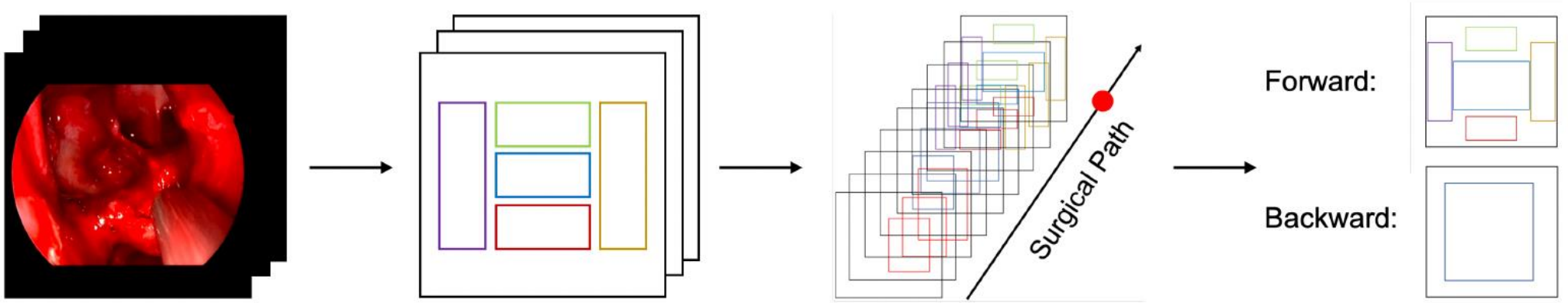


FIGURE 1. *Representation of methodological approach. A sequence of video frames processed to detect bounding boxes relative to various anatomical structures. A neural network then encodes the sequence of detections into a 1D latent variable correlating with the position along the surgical path. Given the current position along the surgical path, an estimation of anatomical structures in the forward or backward directions can be obtained, by extrapolating the current value of the latent variable.*

The "Roadmap Concept"



Sarwin G, Carretta A, Staartjes VE, Zoli M, Mazzatenta D, Regli L, Serra C, Konukoglu E
Live Image-Based Neurosurgical Guidance and Roadmap Generation Using Unsupervised Embedding
IPMI 2023.

University of Zurich[UZH] | USZ Universitäts Spital Zürich

Department of Neurosurgery

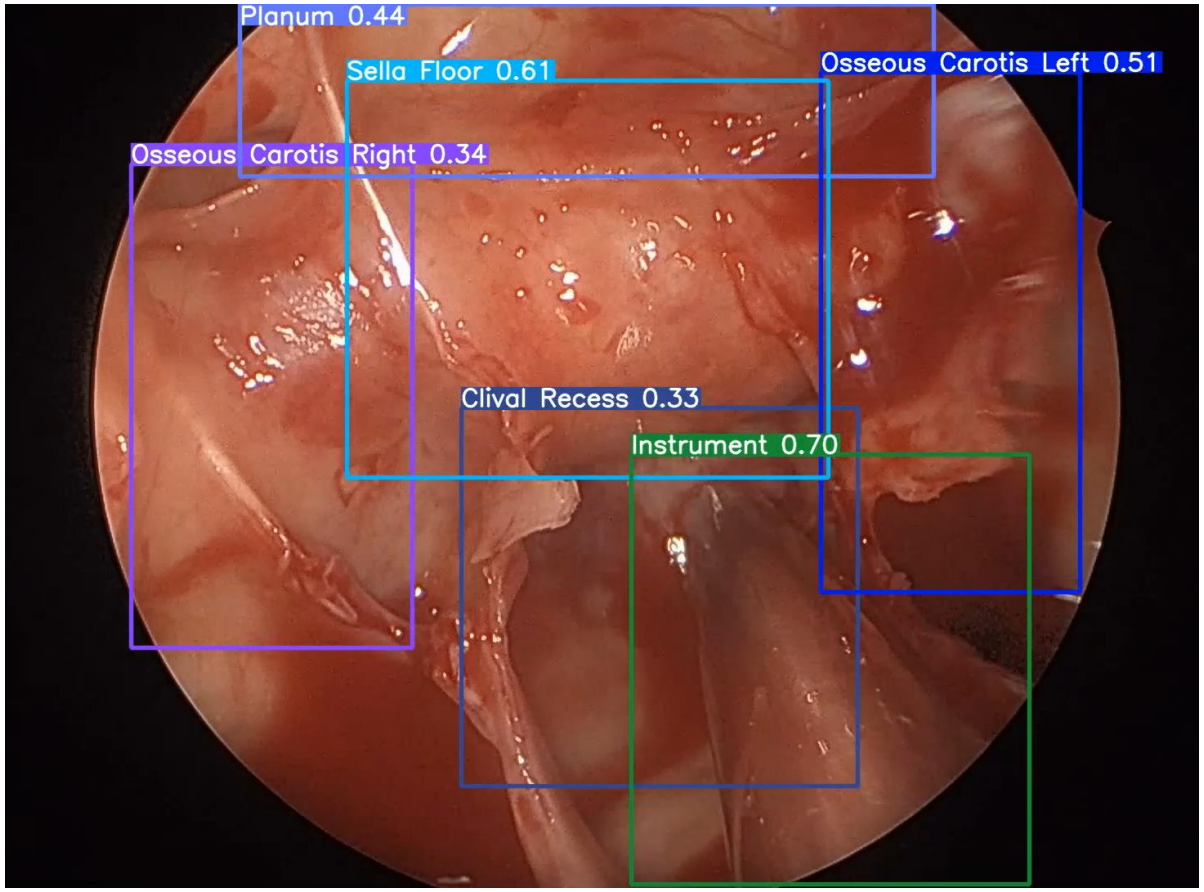
Klinisches Neurozentrum

ETH Eidgenössische Technische Hochschule Zürich Swiss Federal Institute of Technology Zurich

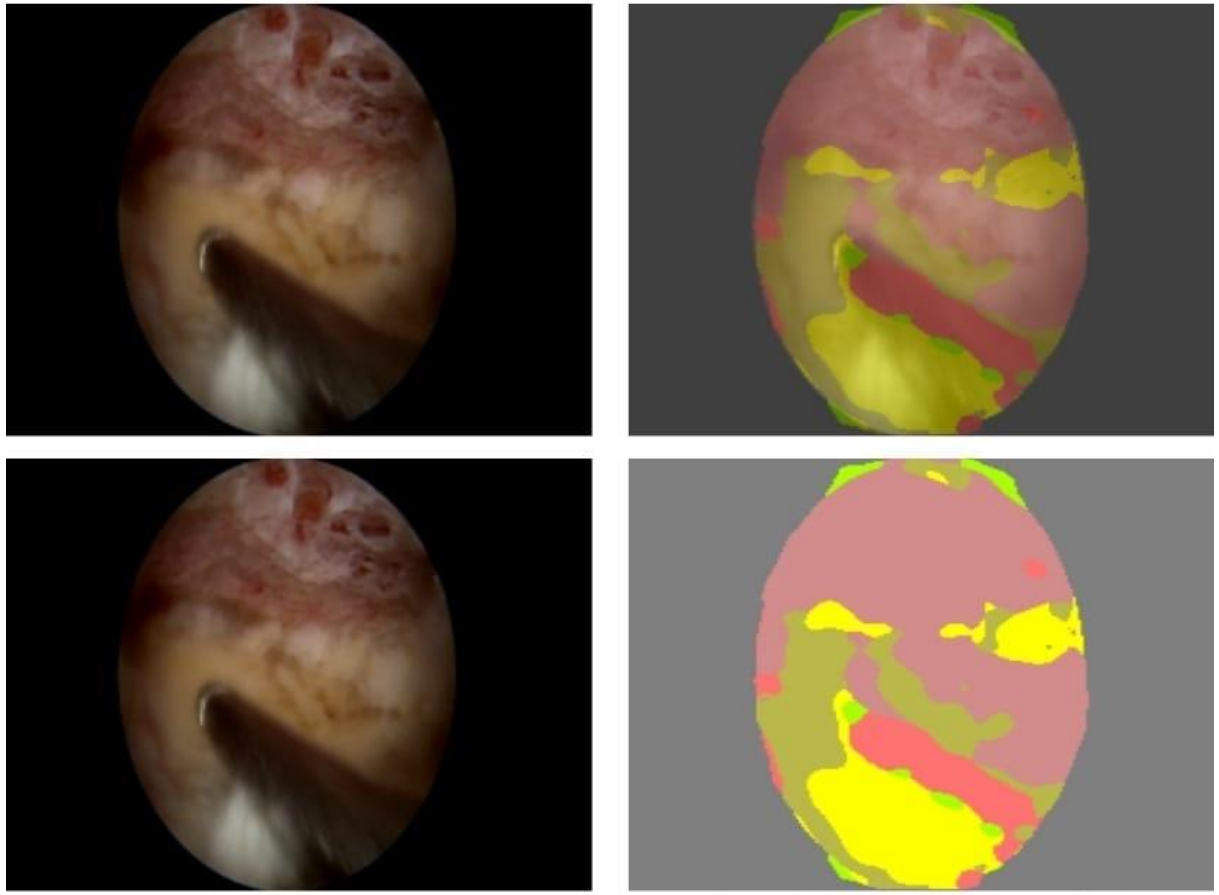
MICN Lab

Project AENEAS

Real-Time Anatomical Navigation in Endo-/Microscopic Brain & Spine Surgery

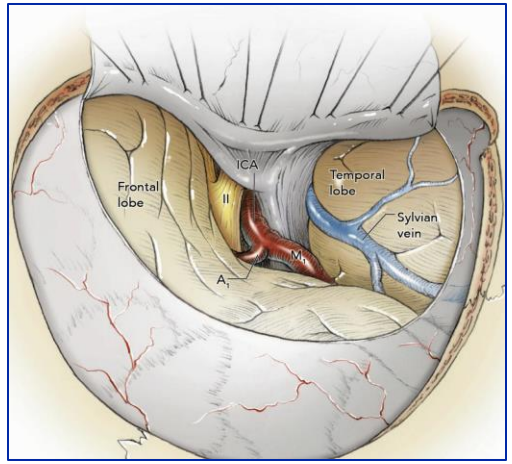


Sarwin G, Konukoglu E, Serra C et al. – Pituitary

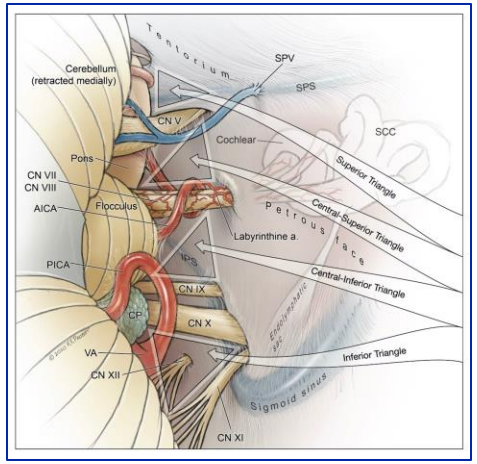


Ryu, Staartjes, Serra et al. – Endoscopic Spine

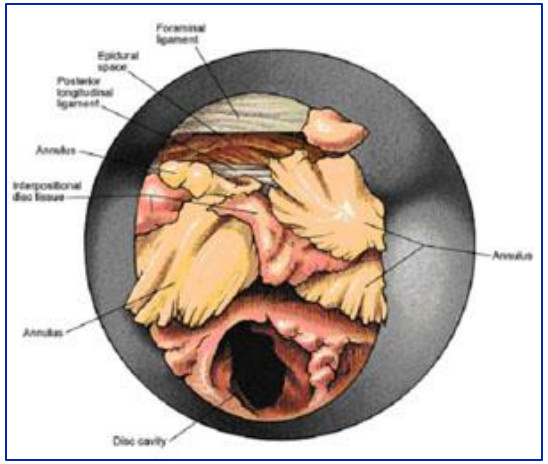
Further Development: Application in Common Craniotomies & Spine



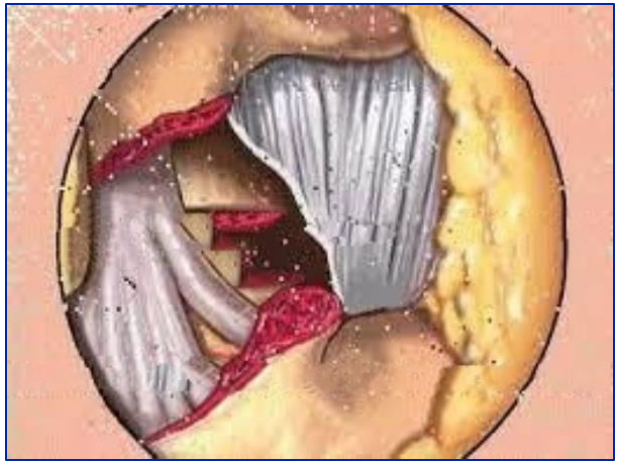
Pterional



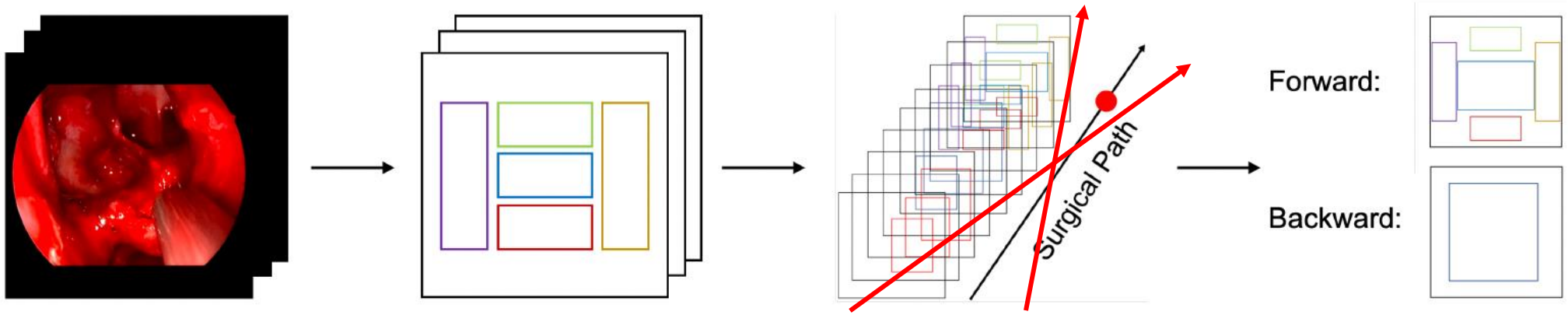
Retrosigmoid



Percutaneous Transforaminal Endoscopic Discectomy

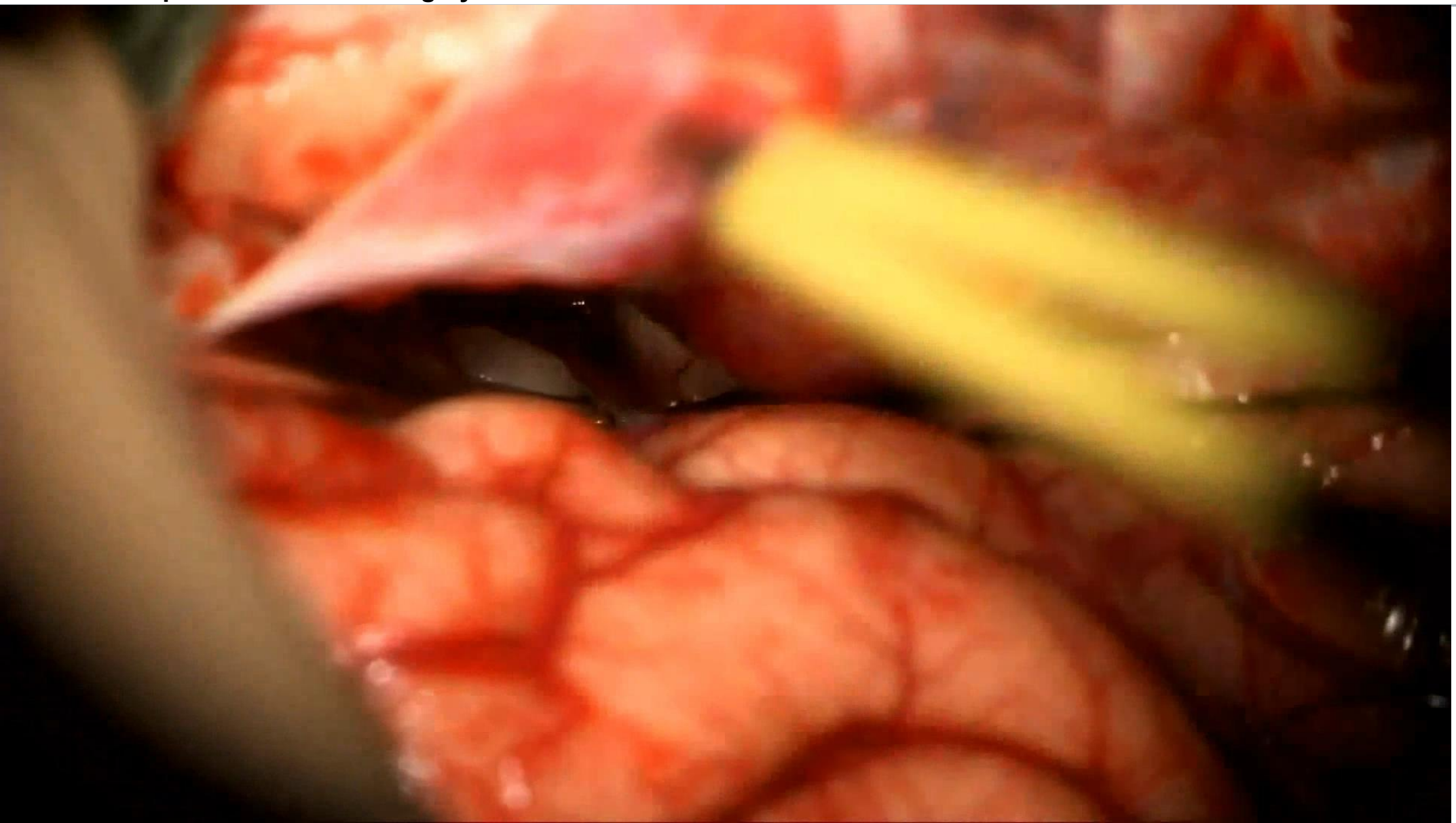


Microdiscectomy



3D polyaxial movement – multiple surgical paths!

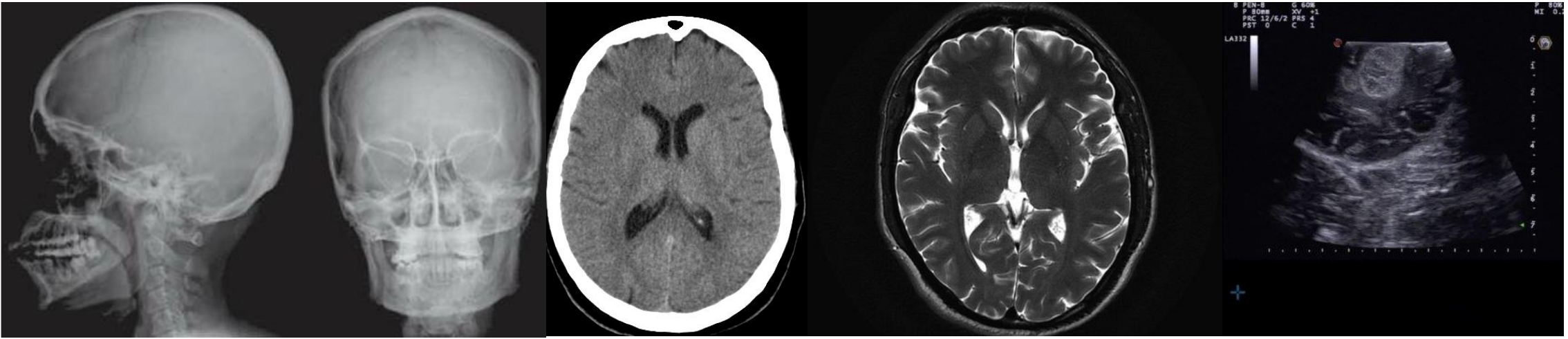
AENEAS Pterional



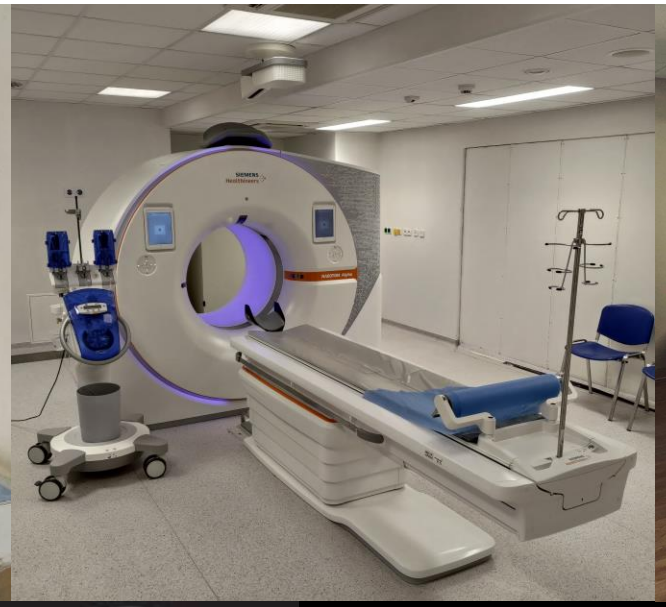
Olei, Sarwin, Staartjes, Regli, Konukoglu, Serra, *unpublished*, 2025

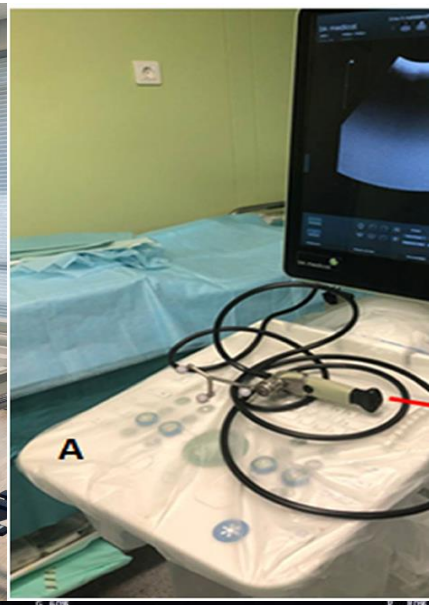
Imaging vs. Synthetic Imaging

- Based on **physical measurement** (+/- processing)
  - X-Ray / Computed Tomography (CT): ionising radiation based radiodensity
  - Magnetic Resonance Imaging (MRI): Powerful magnetic field + RF pulses + complex computation
  - Ultrasound: Sound waves reflected by differing tissue densities

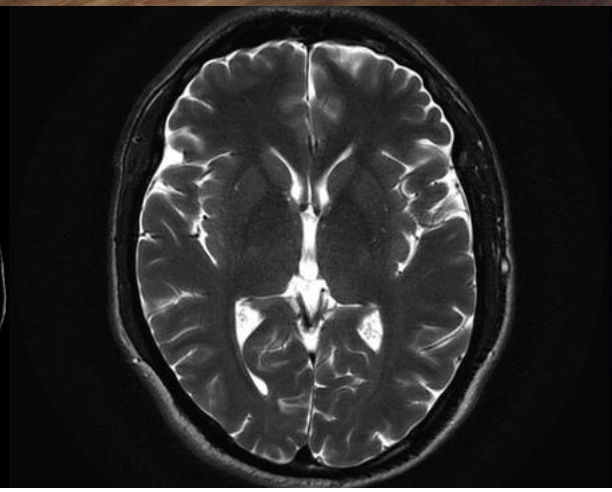


1. Access/Availability - 2. Transport - 3. Costs - 4. Limitations








BoneMRI: Fast synthetic CT from MR of the lumbar spine

University of Zurich[UZH] | USZ Universitäts Spital Zürich

Department of Neurosurgery

Klinisches Neurozentrum | MICN Lab

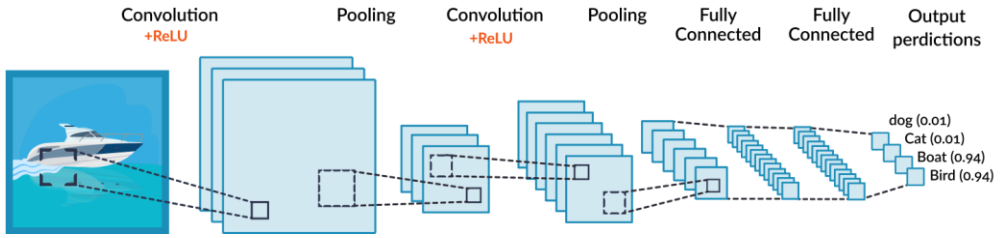
1. The Problem

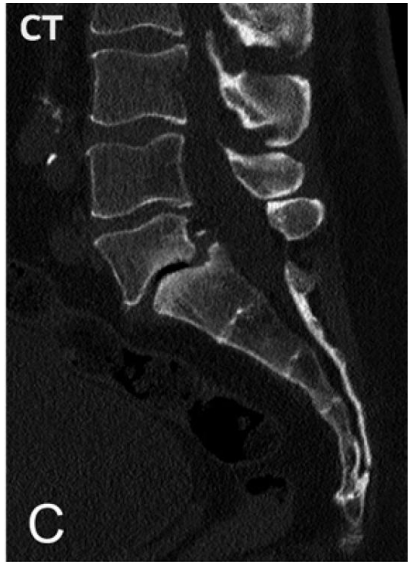
2. The Solution

Generate a **synthetic CT** from a MRI to:

- Eliminate radiation
- Eliminate costs
- Eliminate logistics
- Enhanced diagnostics & planning

3. Our Approach





NEUROSURGICAL FOCUS
Neurosurg Focus 50 (1):E13, 2021

Magnetic resonance imaging–based synthetic computed tomography of the lumbar spine for surgical planning: a clinical proof-of-concept

Victor E. Staartjes, BMed,[1–3] Peter R. Seevinck, PhD,[4,5] W. Peter Vandertop, MD, PhD,[2] Marijn van Stralen, PhD,[4,5] and Marc L. Schröder, MD, PhD[1]

[1]Department of Neurosurgery, Bergman Clinics, Amsterdam; [2]Amsterdam UMC, Vrije Universiteit Amsterdam, Neurosurgery, Amsterdam Movement Sciences, Amsterdam; [4]Image Sciences Institute, University Medical Center Utrecht; and [5]MRIguidance B.V., Utrecht, The Netherlands; and [3]Machine Intelligence in Clinical Neuroscience (MICN) Laboratory, Department of Neurosurgery, University Hospital Zurich, Clinical Neuroscience Centre, University of Zurich, Switzerland

OBJECTIVE Computed tomography scanning of the lumbar spine incurs a radiation dose ranging from 3.5 mSv to 19.5 mSv as well as relevant costs and is commonly necessary for spinal neuronavigation. Mitigation of the need for treatment-planning CT scans in the presence of MRI facilitated by MRI-based synthetic CT (sCT) would revolutionize navigated lumbar spine surgery. The authors aim to demonstrate, as a proof of concept, the capability of deep learning–based generation of sCT scans from MRI of the lumbar spine in 3 cases and to evaluate the potential of sCT for surgical planning.

METHODS Synthetic CT reconstructions were made using a prototype version of the "BoneMRI" software. This deep learning–based image synthesis method relies on a convolutional neural network trained on paired MRI-CT data. A specific but generally available 4-minute 3D radiofrequency-spoiled T1-weighted multiple gradient echo MRI sequence was supplemented to a 1.5T lumbar spine MRI acquisition protocol.

RESULTS In the 3 presented cases, the prototype sCT method allowed voxel-wise radiodensity estimation from MRI, resulting in qualitatively adequate CT images of the lumbar spine based on visual inspection. Normal as well as pathological structures were reliably visualized. In the first case, in which a spiral CT scan was available as a control, a volume CT dose index ($CTDI_{vol}$) of 12.9 mGy could thus have been avoided. Pedicle screw trajectories and screw thickness were estimable based on sCT findings.

CONCLUSIONS The evaluated prototype BoneMRI method enables generation of sCT scans from MRI images with only minor changes in the acquisition protocol, with a potential to reduce workflow complexity, radiation exposure, and costs. The quality of the generated CT scans was adequate based on visual inspection and could potentially be used for surgical planning, intraoperative neuronavigation, or for diagnostic purposes in an adjunctive manner.

https://thejns.org/doi/abs/10.3171/2020.10.FOCUS20801

KEYWORDS lumbar spine; image conversion; imaging; deep learning; machine learning; artificial intelligence

TABLE 1. Exemplary measurements performed comparatively on synthetic CT and spiral CT in case 1

Measurement	sCT	Spiral CT	Difference
L3			
Anterior VBH	26.5	26.5	0.0
Posterior VBH	32.0	31.8	0.2
Spinal canal diameter	14.9	15.0	–0.1
L4			
Anterior VBH	27.2	26.8	0.4
Posterior VBH	29.3	28.9	0.4
Spinal canal diameter	19.1	18.9	0.2
L5			
Anterior VBH	29.1	29.1	0.0
Posterior VBH	11.8	12.0	–0.2
Spinal canal diameter	24.8	25.6	–0.8
Total (MAD ± SD)			0.26 ± 0.24

MAD = mean absolute difference; VBH = vertebral body height.
Measurements are provided in millimeters.

JOURNAL OF NEUROSURGERY:
Case Lessons

J Neurosurg Case Lessons 6(2): CASE23120, 2023
DOI: 10.3171/CASE23120

Robot-assisted screw fixation in a cadaver utilizing magnetic resonance imaging–based synthetic computed tomography: toward radiation-free spine surgery. Illustrative case

*A. Daniel Davidar, MBBS,[1] Brendan F. Judy, MD,[1] Andrew M. Hersh, AB,[1] Carly Weber-Levine, MS,[1] Safwan Alomari, MD,[1] Arjun K. Menta, BS, BBA,[1] Kelly Jiang, MS,[1] Meghana Bhimreddy, BA,[1] Mir Hussain, BS,[2] Neil R. Crawford, PhD,[2] Majid Khan, MBBS, MD,[3] Gary Gong, MD, PhD,[3] and Nicholas Theodore, MD, MS[1]

Departments of [1]Neurosurgery and [3]Radiology, Johns Hopkins University School of Medicine, Baltimore, Maryland; and [2]Globus Medical, Inc, Audubon, Pennsylvania

BACKGROUND Synthetic computed tomography (sCT) can be created from magnetic resonance imaging (MRI) utilizing newer software. sCT is yet to be explored as a possible alternative to routine CT (rCT). In this study, rCT scans and MRI-derived sCT scans were obtained on a cadaver. Morphometric analysis was performed comparing the 2 scans. The ExcelsiusGPS robot was used to place lumbosacral screws with both rCT and sCT images.

OBSERVATIONS In total, 14 screws were placed. All screws were grade A on the Gertzbein-Robbins scale. The mean surface distance difference between rCT and sCT on a reconstructed software model was –0.02 ± 0.05 mm, the mean absolute surface distance was 0.24 ± 0.05 mm, and the mean absolute error of radiodensity was 92.88 ± 10.53 HU. The overall mean tip distance for the sCT versus rCT was 1.74 ± 1.1 versus 2.36 ± 1.6 mm ($p = 0.24$); mean tail distance for the sCT versus rCT was 1.93 ± 0.88 versus 2.81 ± 1.03 mm ($p = 0.07$); and mean angular deviation for the sCT versus rCT was 3.2° ± 2.05° versus 4.04°± 2.71° ($p = 0.53$).

LESSONS MRI-based sCT yielded results comparable to those of rCT in both morphometric analysis and robot-assisted lumbosacral screw placement in a cadaver study.

https://thejns.org/doi/abs/10.3171/CASE23120

KEYWORDS robotics; synthetic CT; bone MRI; screw fixation; lumbar; convolutional neural network

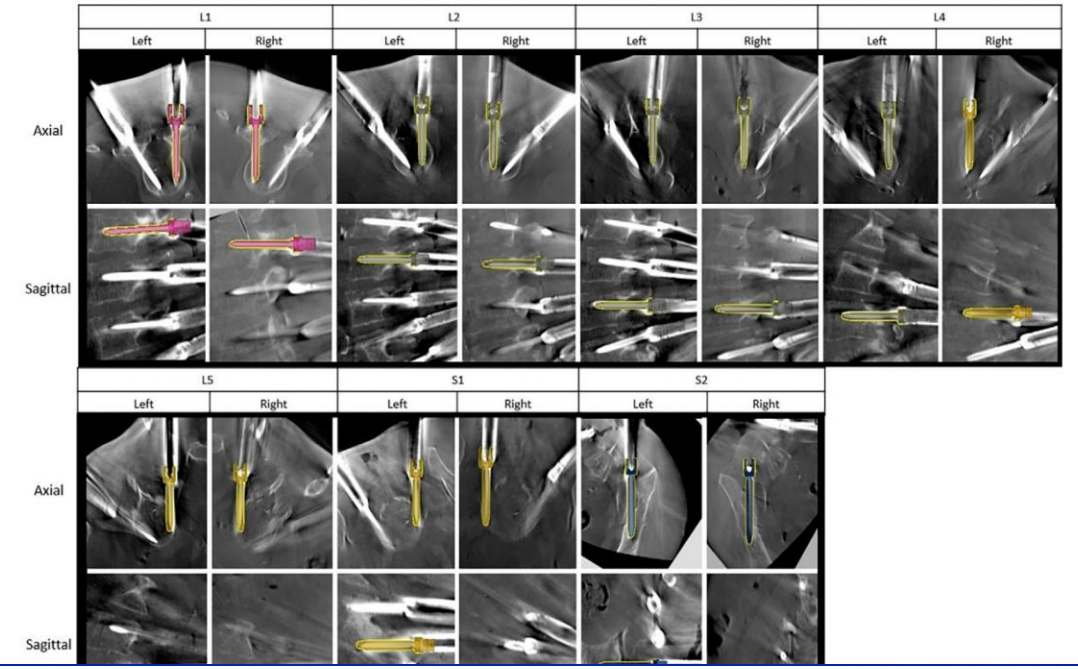
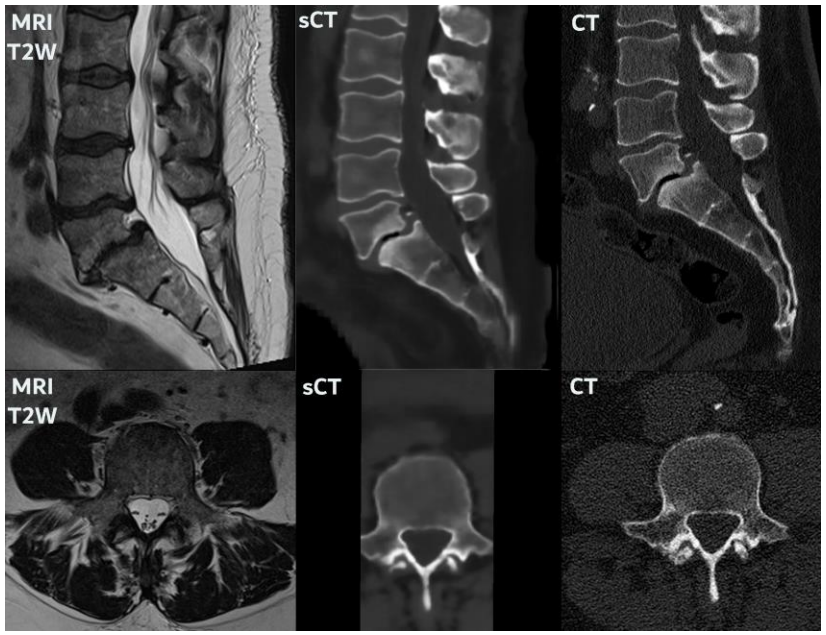


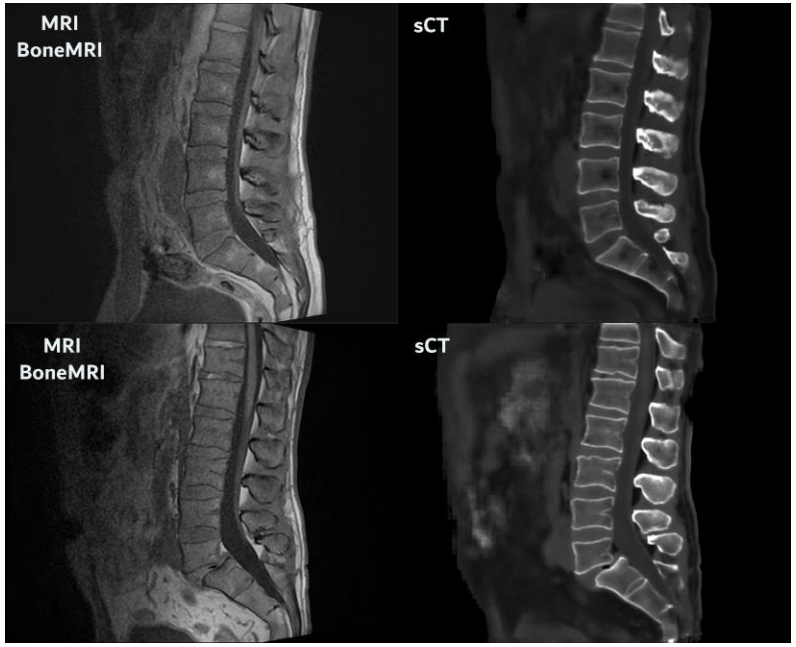
TABLE 2. Screw precision measurements of sCT vs rCT

Level	Tip Distance sCT (mm)	Tip Distance rCT (mm)	Tail Distance sCT (mm)	Tail Distance rCT (mm)	Angular Deviation sCT (degrees)	Angular Deviation rCT (degrees)
L1	0.5	1.1	1.1	2	2.1	2.1
L2	1.1	1.7	2.1	2.3	3	3.7
L3	1	0	0.8	2.2	2.2	2.5
L4	2.4	2.7	1.6	4.5	2.7	8
L5	3.4	2.8	2.9	3.9	3.4	7.8
S1	2.8	4.8	3.2	3	7.6	2.7
S2AI	1	3.4	1.8	1.8	1.4	1.5
Mean ± SD	1.74 ± 1.1	2.36 ± 1.6	1.93 ± 0.88	2.81 ± 1.03	3.2 ± 2.05	4.04 ± 2.71
p value (sCT vs rCT)	0.24		0.07		0.53	

Tip distance, tail distance, and angular deviation were measured using proprietary software where planned sCT screws were overlayed with postoperative CT. Images were aligned in 3D axes and screw tip distance, tail distance, and angular deviation were measured.

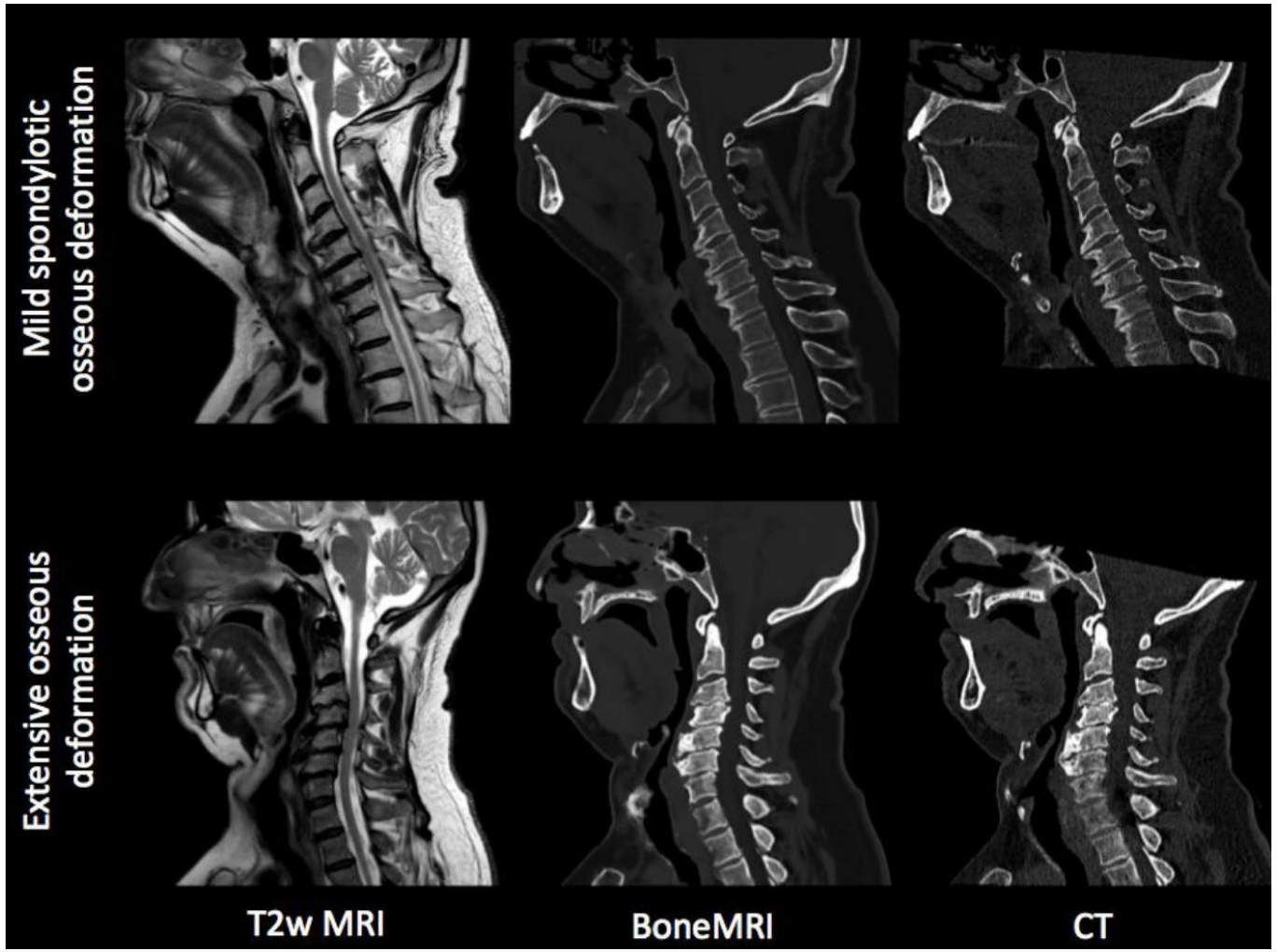
Fast MR- to CT-conversion for lumbar spinal diagnostics & navigation using convolutional neural networks ("synthetic CT/BoneMRI")





Staartjes VE et al.
Magnetic resonance imaging-based synthetic computed tomography of the lumbar spine for surgical planning: a clinical proof-of-concept.
Neurosurg Focus. 2021;50(1):E13. doi:10.3171/2020.10.FOCUS20801

BoneMRI in the cervical spine



Staartjes VE et al.
Magnetic resonance imaging-based synthetic computed tomography of the lumbar spine for surgical planning: a clinical proof-of-concept.
Neurosurg Focus. 2021;50(1):E13. doi:10.3171/2020.10.FOCUS20801

University of Zurich[UZH] | USZ Universitäts Spital Zürich

Department of Neurosurgery

Klinisches Neurozentrum

MICN Lab

MR-only intraoperative robotic pedicle screw placement workflow

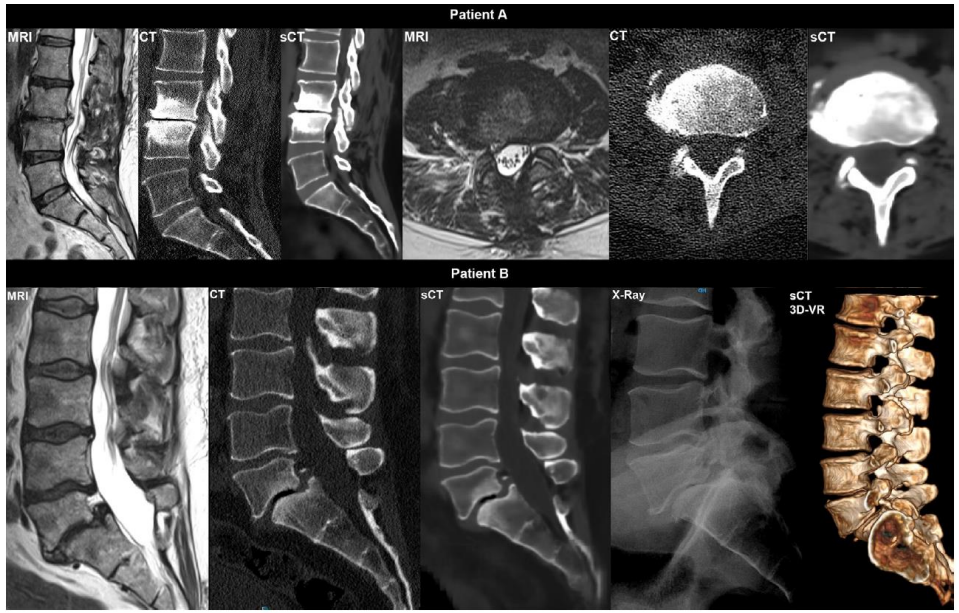


FIG. 15.1 *Patient A* is depicted in the top panel. It demonstrates preoperative imaging in sagittal and axial multiplanar reconstructions. T2-weighted MRI, conventional spiral CT, and synthetic CT (sCT) based on the BoneMRI technique are shown. *Patient B* is depicted in the bottom panel. It demonstrates preoperative imaging in a sagittal reconstruction. T2-weighted MRI, conventional spiral CT, and sCT based on the BoneMRI technique. In addition, the preoperative lumbosacral x-ray and a 3D-volume render based on the synthetic CT demonstrates spondylolisthesis due to the bilateral L5 spondylolysis.

FIG. 15.2 The top panel (*Patient A*) demonstrates the planning process for an L3–L4 pedicle screw instrumentation on the Mazor Robotics SpineAssist workstation. The bottom panel (*Patient B*) demonstrates the planning process for an L5–S1 pedicle screw instrumentation on the Mazor Robotics SpineAssist workstation.

Jin, Michael, Marc Schröder, and Victor E. Staartjes. "15 - Artificial Intelligence and Machine Learning in Spine Surgery." In *Robotic and Navigated Spine Surgery*, edited by Anand Veeravagu and Michael Y. Wang, 213–29. New Delhi: Elsevier, 2023. https://doi.org/10.1016/B978-0-323-71160-9.00015-0.

Synthetic Imaging - Primer

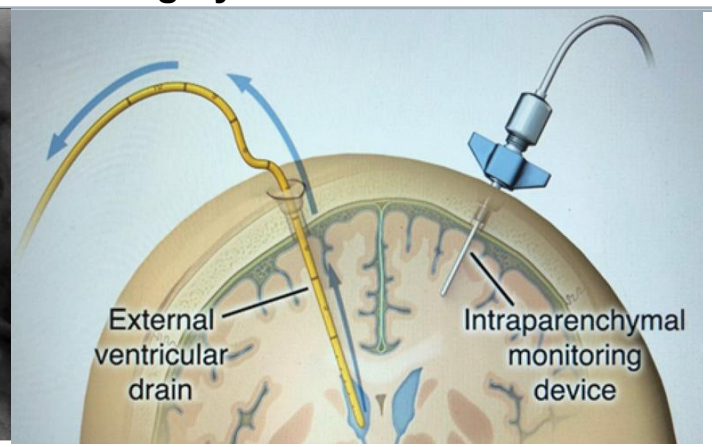
- Image Modality Conversion (MRI to CT, CT to MRI, Sonography to MRI)
- **Dimensionality enhancement (Tomographic imaging from radiographic imaging, 3D from 2D)**
- Extracting adjuvant information from existing imaging (Waveform inversion imaging, BOLD contrast-free perfusion imaging)

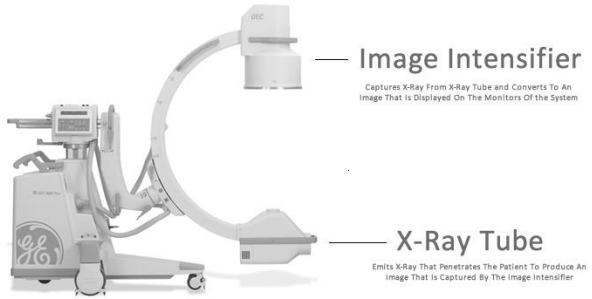
University of Zurich[UZH] | USZ Universitäts Spital Zürich

Department of Neurosurgery

Klinisches Neurozentrum | MICN Lab

1. The Problem



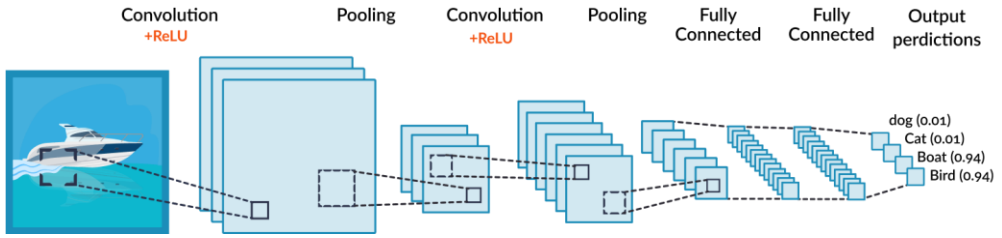


2. The Solution

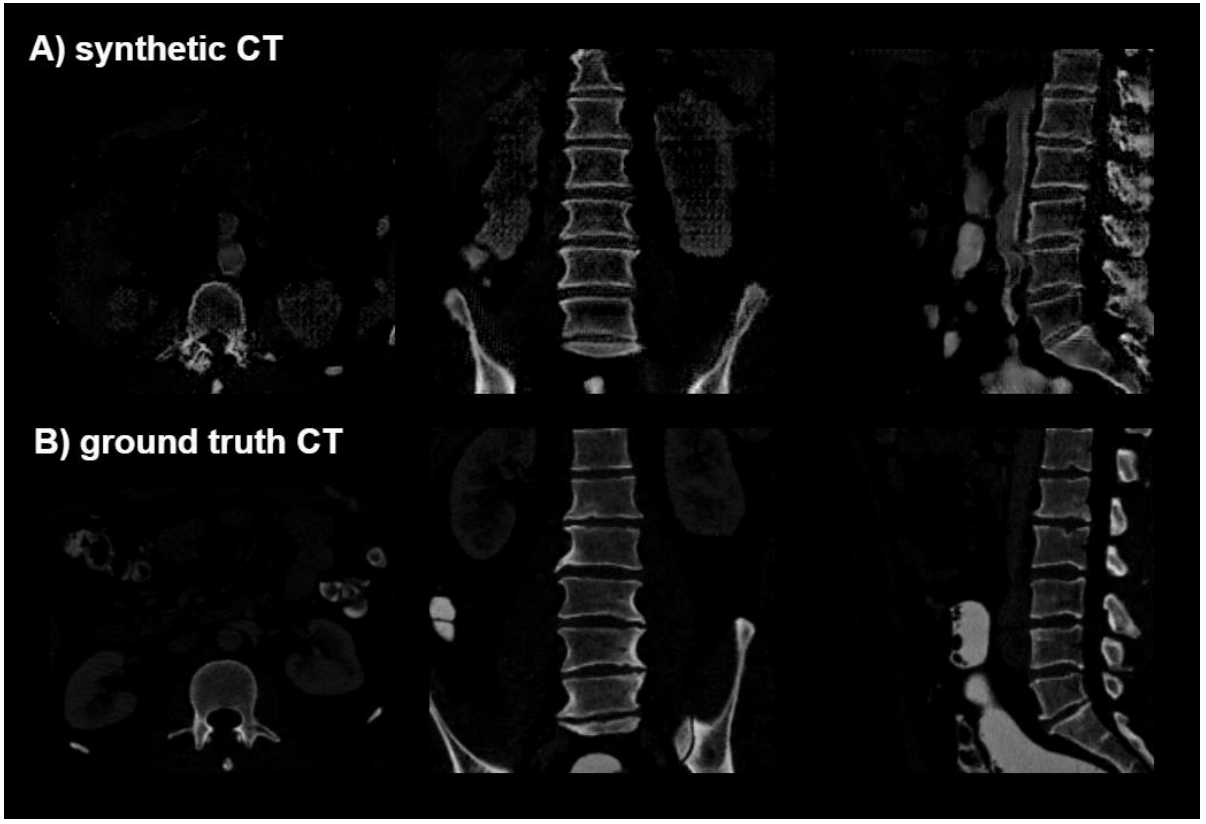
Generate a **synthetic 3D CT** from **biplanar x-ray**:

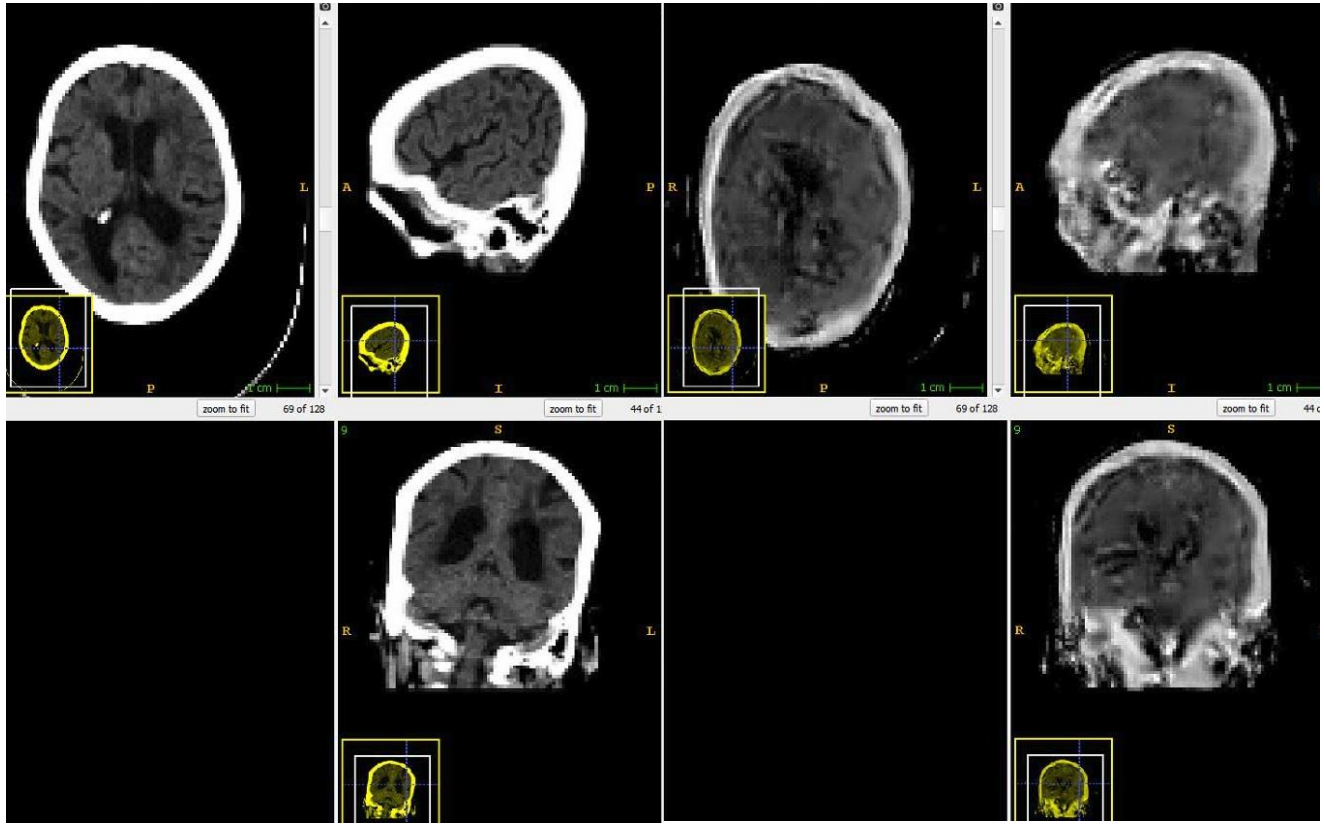
- Eliminate radiation
- Eliminate costs
- Eliminate logistics
- Enhanced and repeatable diagnostics & planning

3. Our Approach



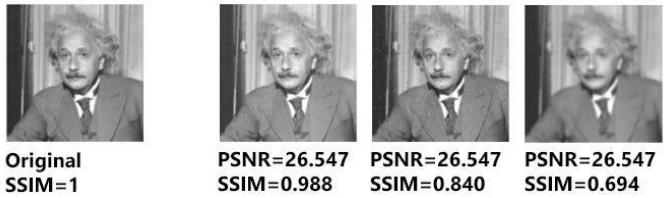
Fast synthetic computed tomography (CT) generation from biplanar x-ray of the skull and spine



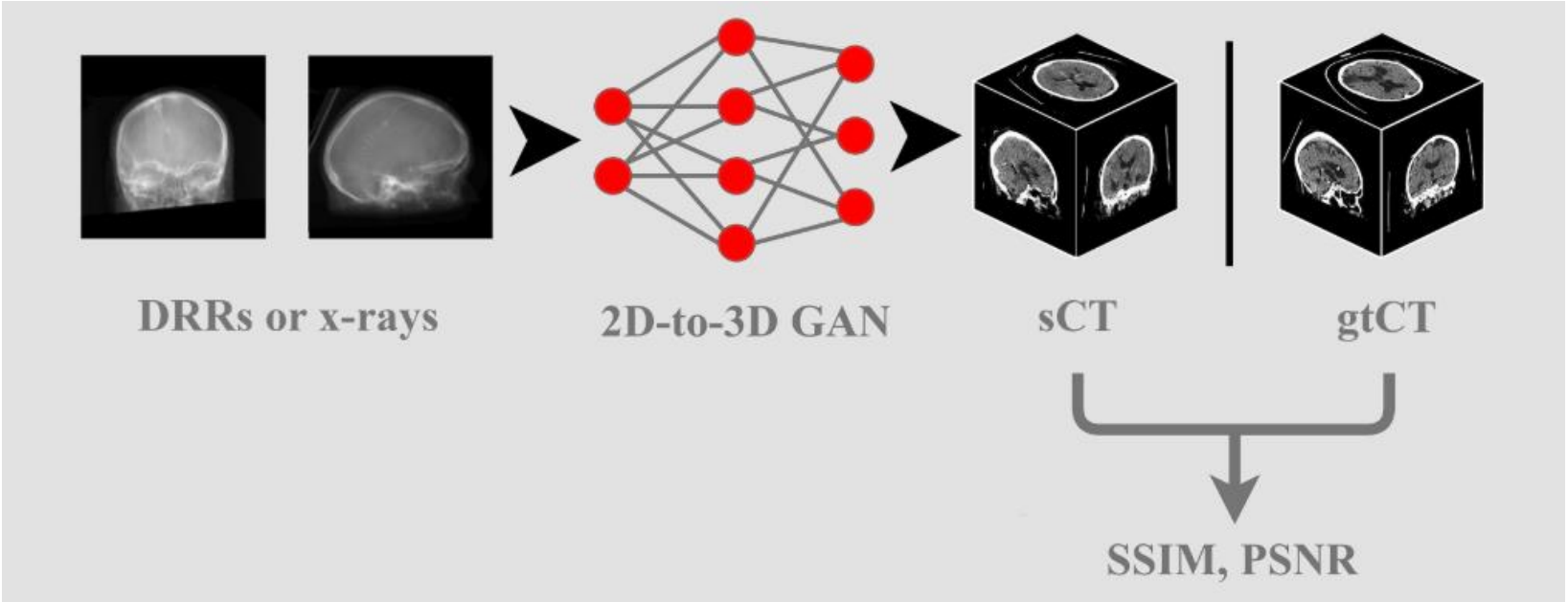


Zanier O, Da Mutten R, Staartjes VE (October 2023) work in progress

TomoRay: Cranial



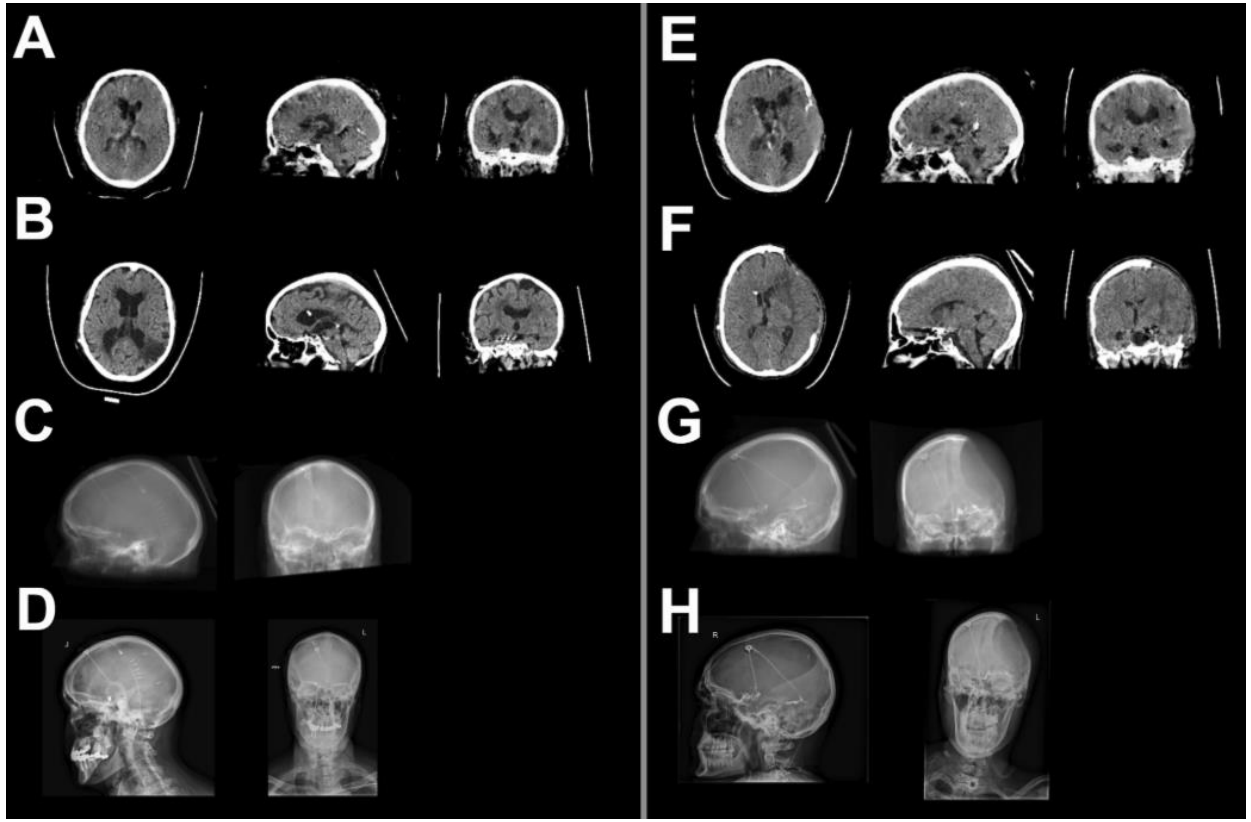
– Basic Principle: CNN-based GAN

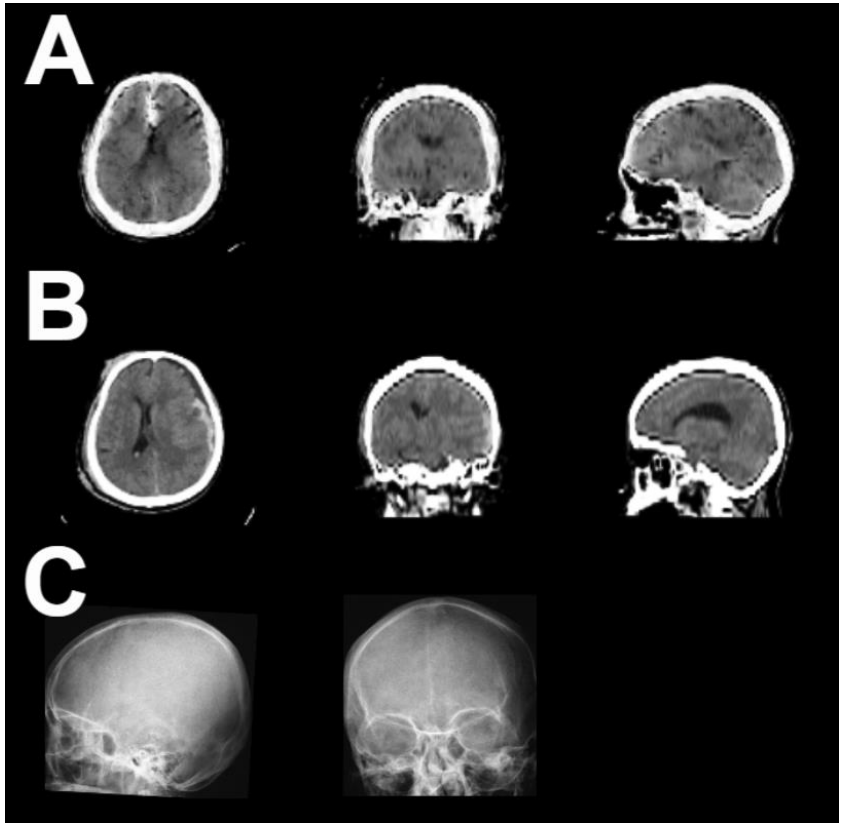


TomoRay Cranial: Synthesis of Cranial CT Imaging from Biplanar Radiographs using a Generative Adversarial Model
Olivier Zanier, Seung-Jun Ryu, Raffaele Da Mutten, Alessandro Carretta, Giorgio Palandri, Diego Mazzatenta, Luca Regli, Carlo Serra, Victor E. Staartjes
Submitted

TomoRay: Cranial

– Two models: Based on DRRs (Model 1) or based on x-rays (Model 2)

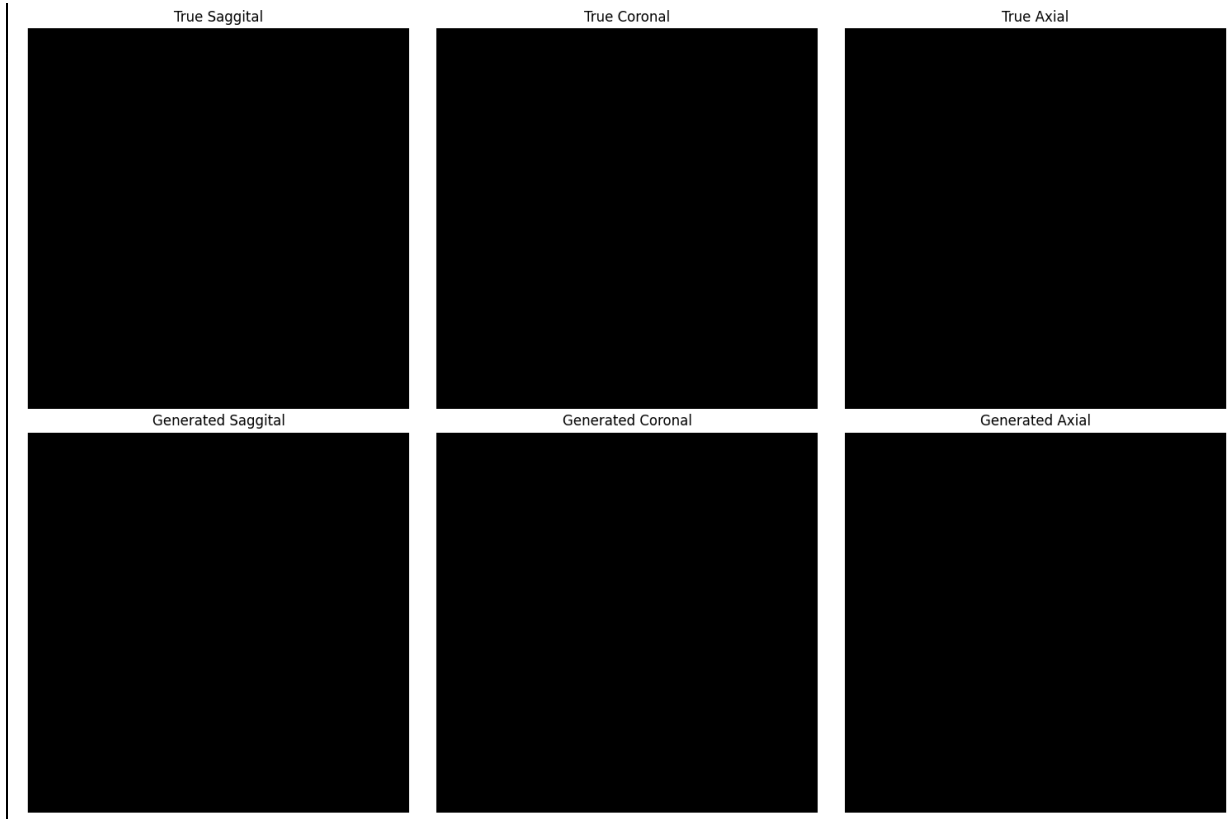


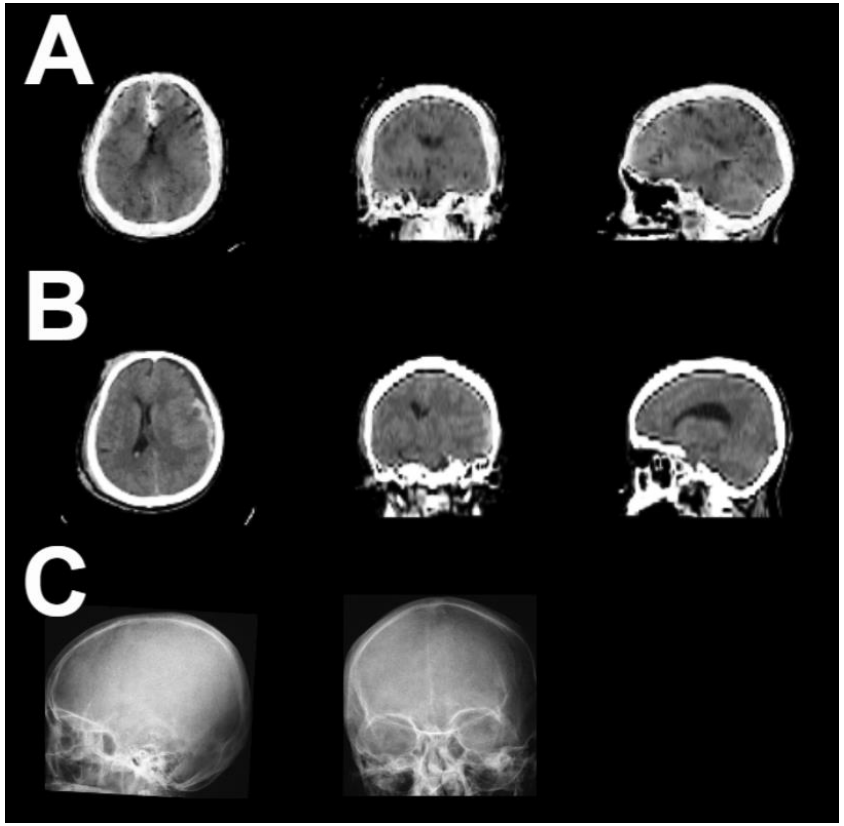


TomoRay Cranial: Synthesis of Cranial CT Imaging from Biplanar Radiographs using a Generative Adversarial Model
Olivier Zanier, Seung-Jun Ryu, Raffaele Da Mutten, Alessandro Carretta, Giorgio Palandri, Diego Mazzatenta, Luca Regli, Carlo Serra, Victor E. Staartjes
Submitted

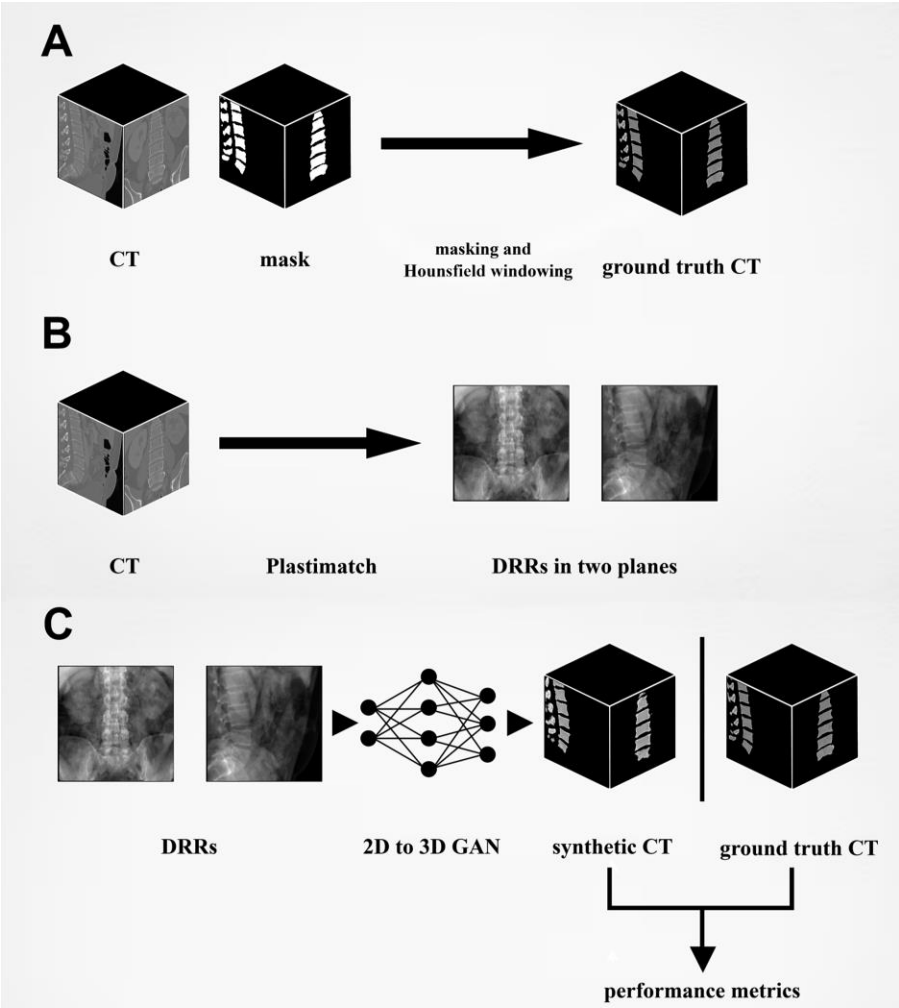
TomoRay: Cranial

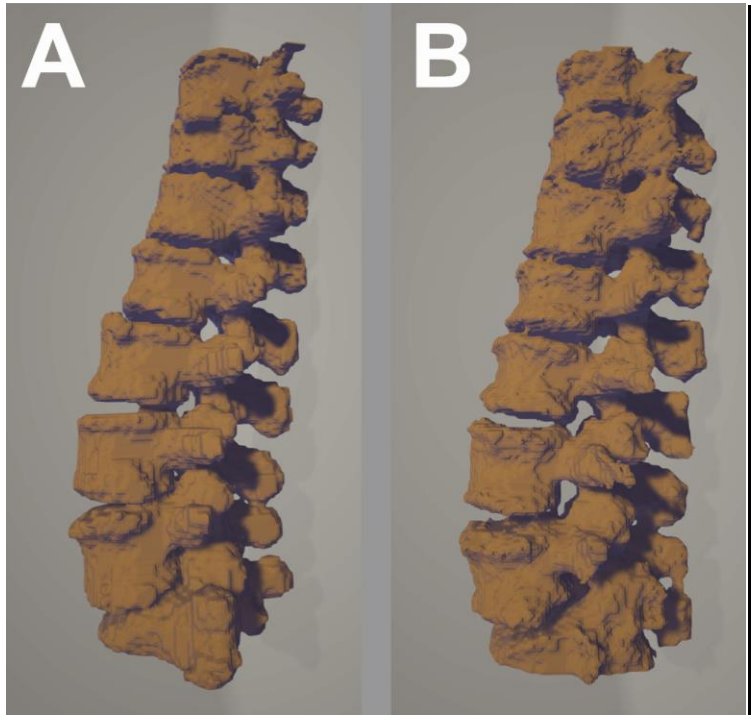
– Two models: Based on DRRs (Model 1) or based on x-rays (Model 2)





TomoRay Cranial: Synthesis of Cranial CT Imaging from Biplanar Radiographs using a Generative Adversarial Model
Olivier Zanier, Seung-Jun Ryu, Raffaele Da Mutten, Alessandro Carretta, Giorgio Palandri, Diego Mazzatenta, Luca Regli, Carlo Serra, Victor E. Staartjes
Submitted





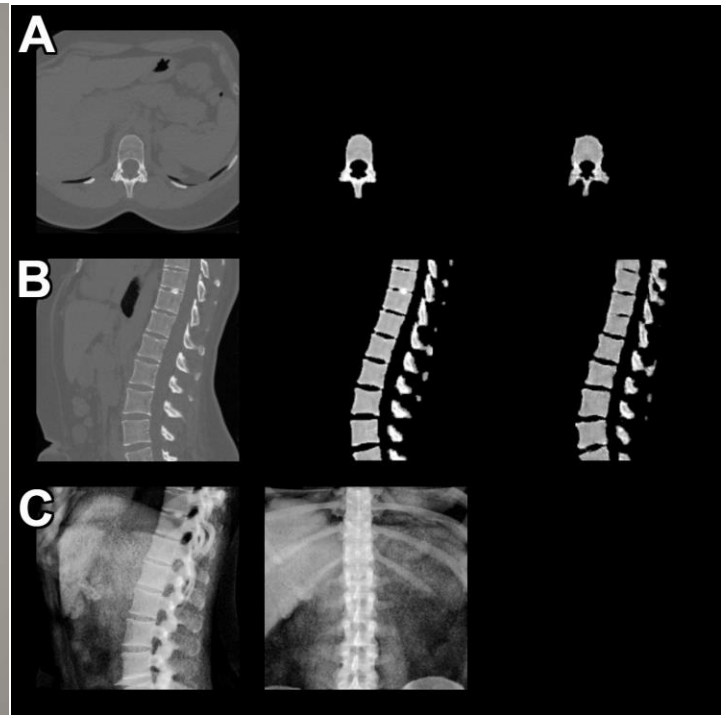


Table 2. Model performance on training and holdout sets

Cohort	PSNR (3D) [dB]	SSIM (2D)	Cosine similarity
Training (VerSe20) (n = 209)			
Mean ± SD	27.011 ± 2.955	0.972 ± 0.008	0.908 ± 0.081
Median (IQR)	26.743 (25.989–27.554)	0.973 (0.968–0.977)	0.927 (0.907–0.935)
Internal validation (VerSe20) (n = 55)			
Mean ± SD	22.206 ± 1.027	0.953 ± 0.009	0.709 ± 0.195
Median (IQR)	22.419 (21.458–22.801)	0.953 (0.948–0.961)	0.779 (0.728–0.817)
External validation (CTSpine1k subset) (n = 56)			
Mean ± SD	21.139 ± 1.018	0.947 ± 0.010	0.671 ± 0.078
Median (IQR)	21.273 (20.514–21.821)	0.948 (0.942–0.953)	0.691 (0.652–0.716)

SSIM is calculated for each dimension individually and then averaged for each image. The final values reported are the mean and median over the whole dataset.
PSNR, peak signal to noise ratio; SSIM, structural similarity index; 3D, 3-dimensional; 2D, 2-dimensional; SD, standard deviation; IQR, interquartile range.

TomoRay: Generating Synthetic Computed Tomography of the Spine From Biplanar Radiographs

Olivier Zanier[1], Sven Theiler[1], Raffaele Da Mutten[1], Seung-Jun Ryu[2], Luca Regli[1], Carlo Serra[1], Victor E Staartjes[1]

Neurospine, 2024 Mar 21

University of Zurich[UZH] | USZ Universitäts Spital Zürich

Department of Neurosurgery

Klinisches Neurozentrum

MICN Lab

Thank you for your attention!

MICN Lab | USZ Universitäts Spital Zürich | University of Zurich[UZH]

Fourth

Zurich Machine Intelligence in Clinical Neuroscience Symposium

"Towards the Horizon – and Beyond"

Thursday, June 19th 2025
Auditorium Monakow, USZ
9:00 a.m. – 4:30 p.m. UTC +2 (Zurich)
3:00 a.m. – 10:30 a.m. UTC -4 (New York)
5:00 p.m. – 0:30 a.m. UTC +10 (Sydney)

Registration:
www.micnlab.com/symposium2025

Registration and participation (online or on-site) are completely free of charge! The event will be held in a hybrid format: **Online over Zoom** and at the University Hospital Zurich, Haldenbach Building, Auditorium Monakow.
You will receive the Zoom link and further information in time and after registering using the abovementioned form.
A recording of the symposium will be made available on the MICN Lab website (**www.micnlab.com**)

Director
Carlo Serra Luca Regli Victor Staartjes

Machine Intelligence in Clinical Neuroscience & Microsurgical Neuroanatomy (MICN) Laboratory
Department of Neurosurgery & Clinical Neuroscience Center
University Hospital Zurich, University of Zurich

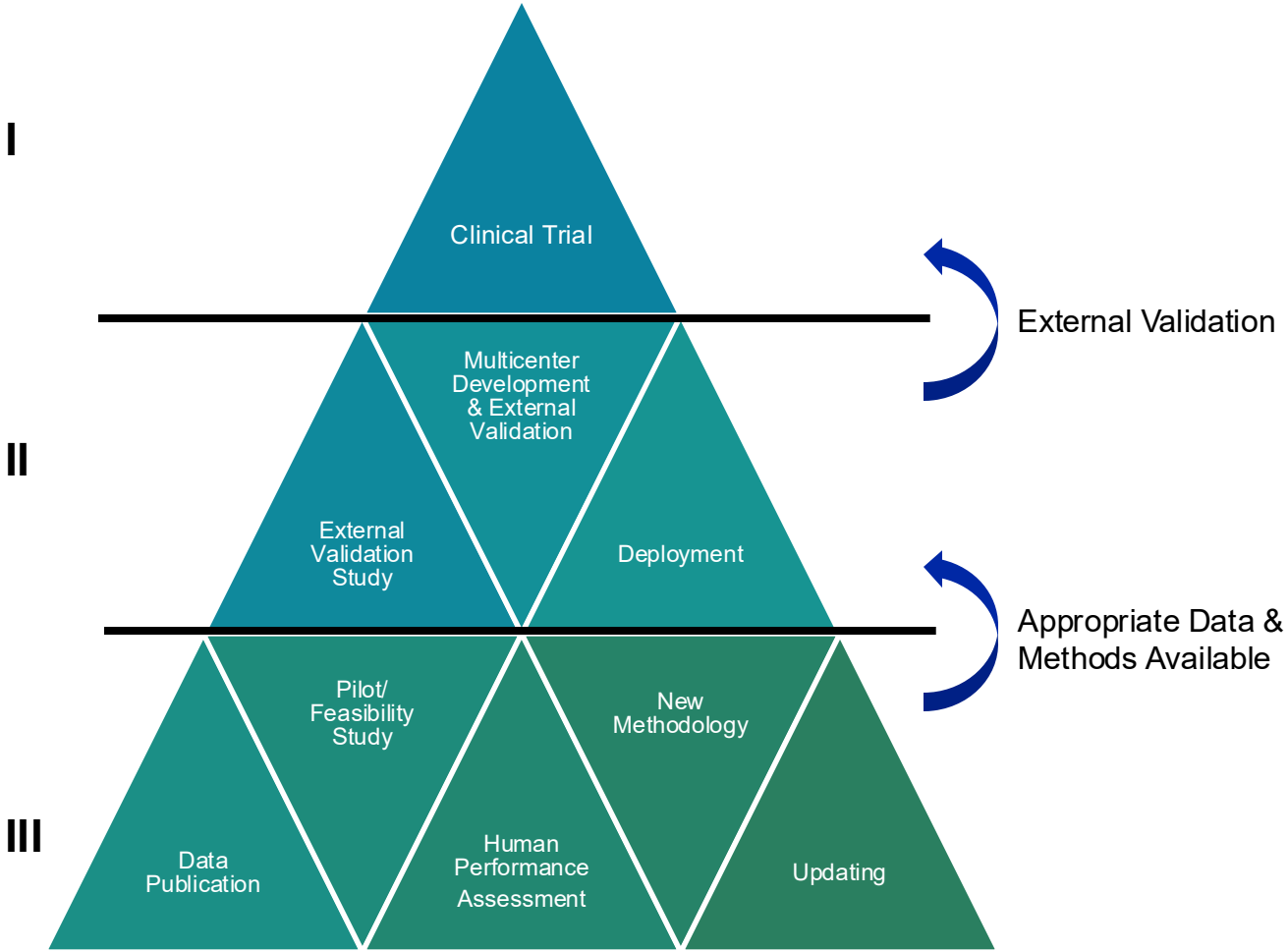
Invited Faculty

Nick Ramsey (Utrecht, NL)	Live	Christoph Laux (Zurich, CH)	Live
Antonio Di Ieva (Sydney, AU)	Online	Alessandro Blasimme (Zurich, CH)	Live
Miikka Korja (Helsinki, FI)	Live	Alessandro Boaro (Verona, IT)	Live
Peter Seevinck (Utrecht, NL)	Live	Vivek Buch (Stanford, US)	Online
Daniel Donoho (Washington, US)	Online	Alexandru Paunoiu (Zurich, CH)	Live
Gary Sarwin (Zurich, CH)	Live	Julius Kernbach (Heidelberg, DE)	Live
Eric Suero Molina (Münster, DE)	Live	Benjamin Hopkins (LA, US)	Online

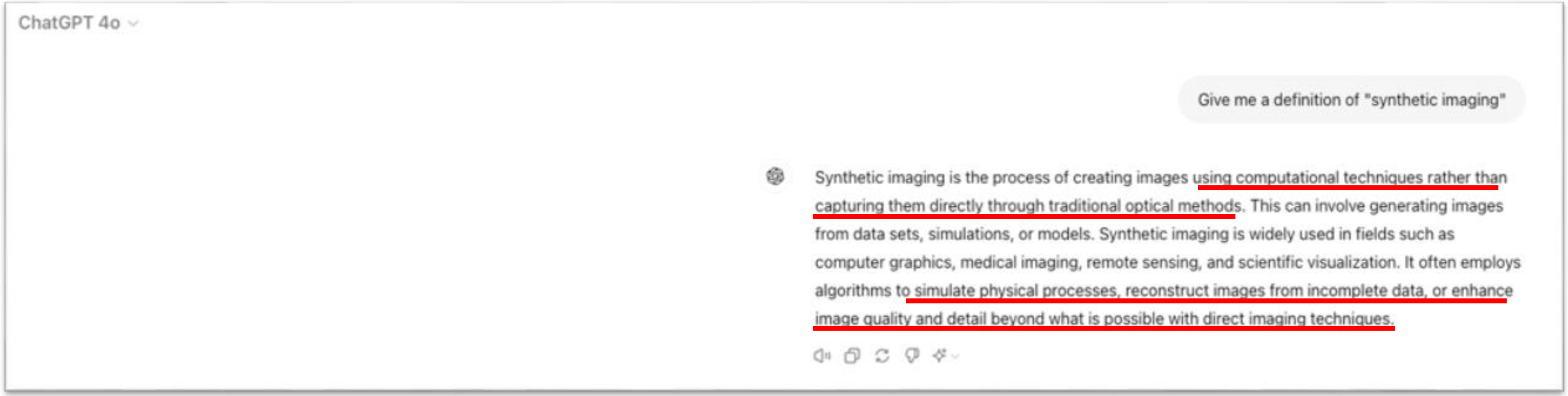
victoregon.staartjes@usz.ch – www.micnlab.com

Where are we?

Types of (useful) ML / AI publications



Synthetic Imaging



Synthetic Imaging

- Image Modality Conversion (MRI to CT, CT to MRI, Sonography to MRI)
- Dimensionality enhancement (Tomographic imaging from radiographic imaging, 3D from 2D)
- Extracting adjuvant information from existing imaging (Waveform inversion imaging, BOLD contrast-free perfusion imaging)

Synthetic Imaging - Primer

- **Image Modality Conversion (MRI to CT, CT to MRI, Sonography to MRI)**

- Dimensionality enhancement (Tomographic imaging from radiographic imaging, 3D from 2D)
- Extracting adjuvant information from existing imaging (Waveform inversion imaging, BOLD contrast-free perfusion imaging)

University of Zurich[UZH] | USZ Universitäts Spital Zürich

Department of Neurosurgery

Klinisches Neurozentrum

MICN Lab

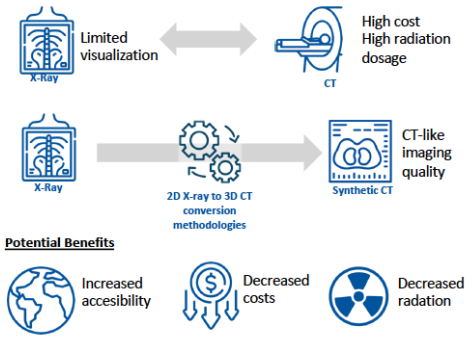
Strategies for generating tomography from radiographs

Graphical Abstract

Strategies for generating synthetic computed tomography-like imaging from radiographs: A scoping review

MICN Lab

Background



Potential Benefits

Increased accesibility

Decreased costs

Decreased radation

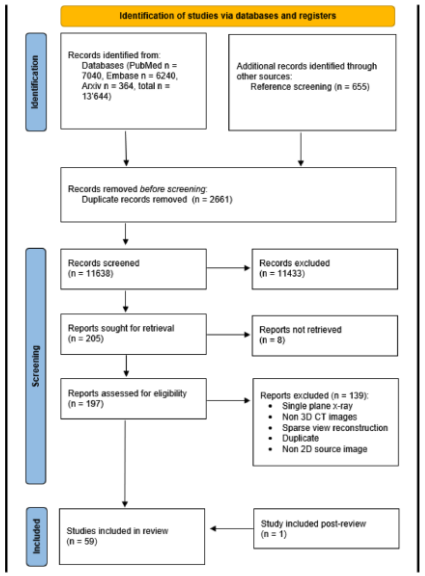
Methods

Databases: PubMed, Embase, arXiv

Search Date: July 6, 2023

Eligibility Criteria: Biplanar or 4-projection x-ray → 3D CT-like imaging

Results



Majority (50.8%) were published between 2010 and 2020

Anatomical areas of interest: Chest (18.6%), spine (15.3%), coronary arteries (11.9%), and head (11.9%), among others. 22% without medical data.

Main methodology: Deep learning

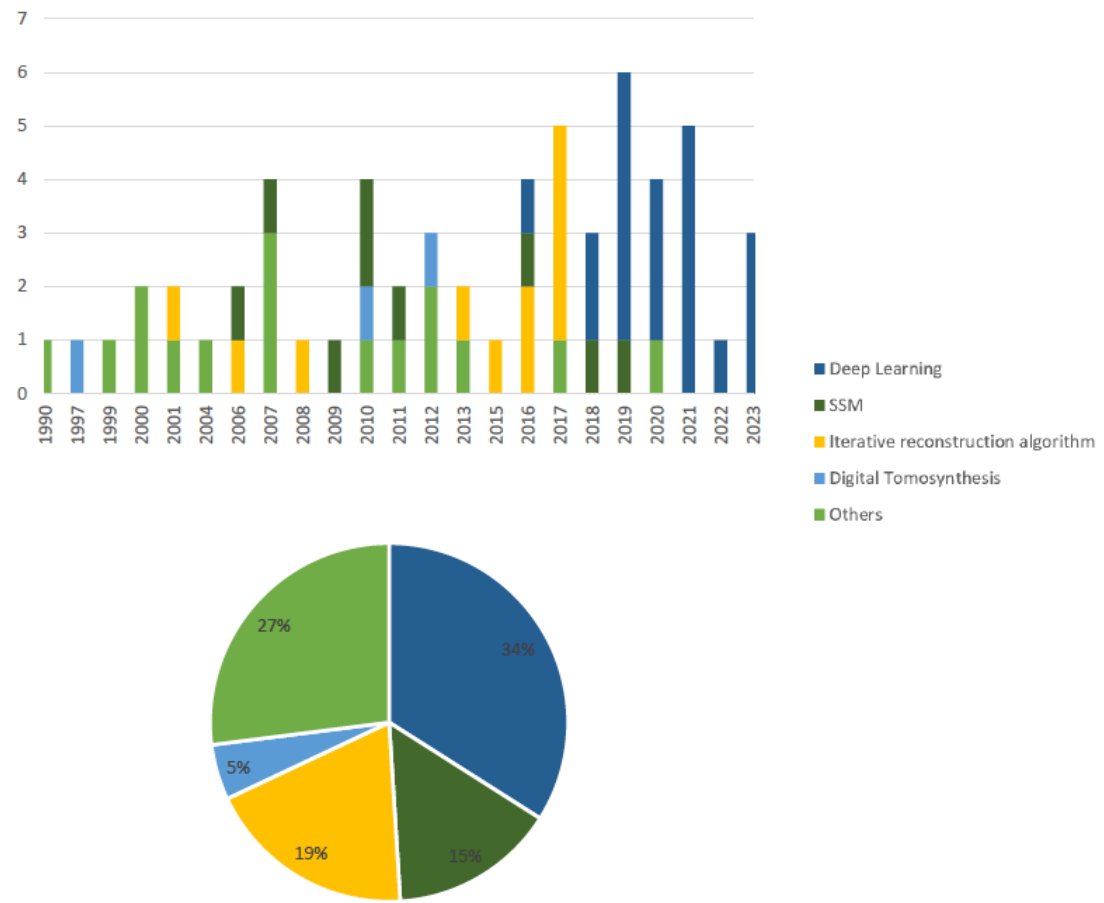
Other methodologies: statistical shape models, iterative reconstruction algorithms, digital tomosynthesis and more.

91.7% of studies reported validation of their methodology: 47.2% validated their methods without using any initial data for training, 24.5% using external datasets, 15.1% using cross-validation, and 15.1% did not specify the validation method.

Conclusion

Despite promising results, the field **encounters challenges.** There are a wide range of differing methodologies that frequently **lack proper validation. Additional research needs** to address these challenges to establish the role of 2D to 3D image conversion in the medical field.

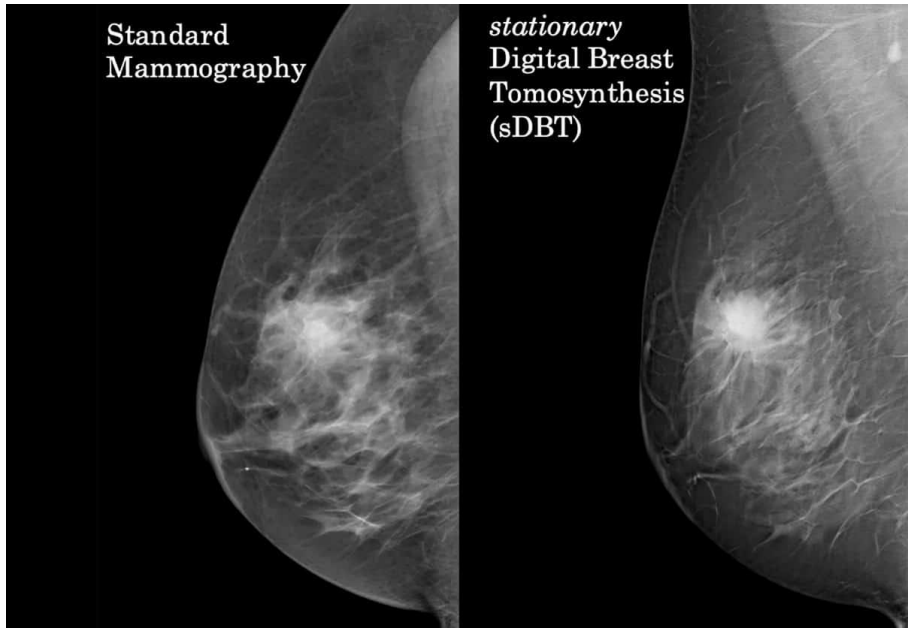
Fig. 2 Overview of studies included based on their date of publication and relative occurrence, grouped by methodology

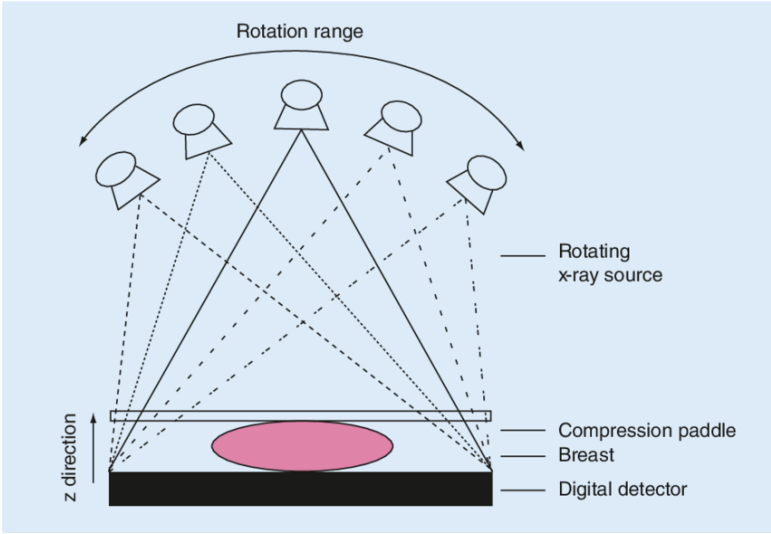


Strategies for generating synthetic computed tomography-like imaging from radiographs: A scoping review
Daniel De Wilde, Olivier Zanier, Raffaele Da Mutten, Michael Jin, Luca Regli, Carlo Serra, Victor E. Staartjes
Under Review

Strategies for generating tomography from radiographs

- Digital Tomosynthesis
  - Mammography
  - Multiple Projections acquired + filtered backprojection or iterative reconstruction

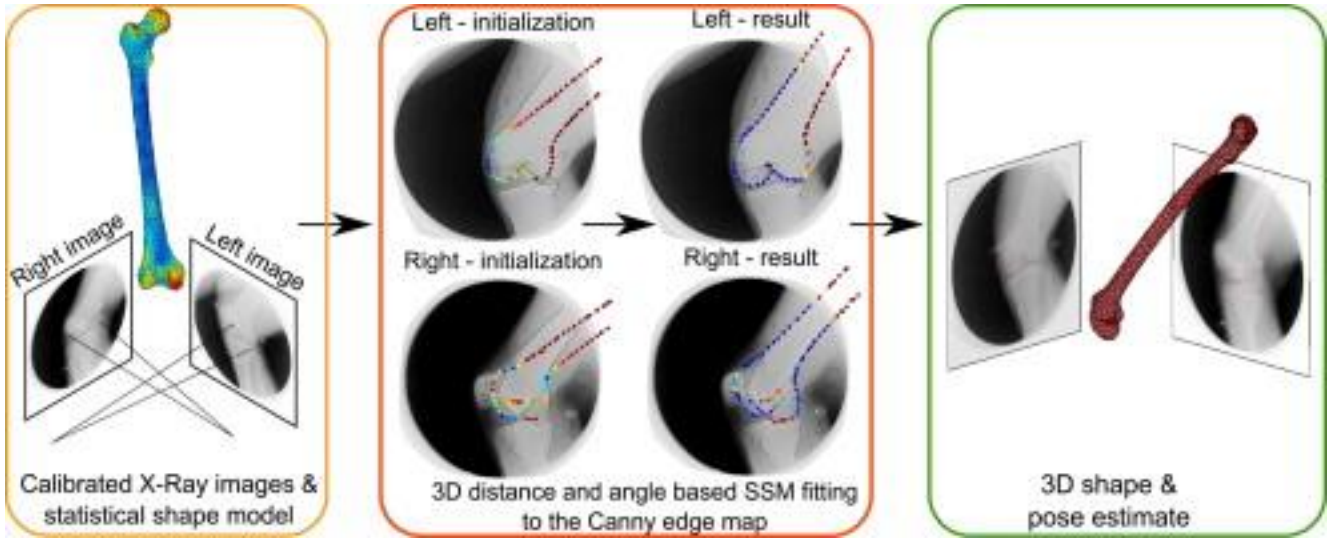


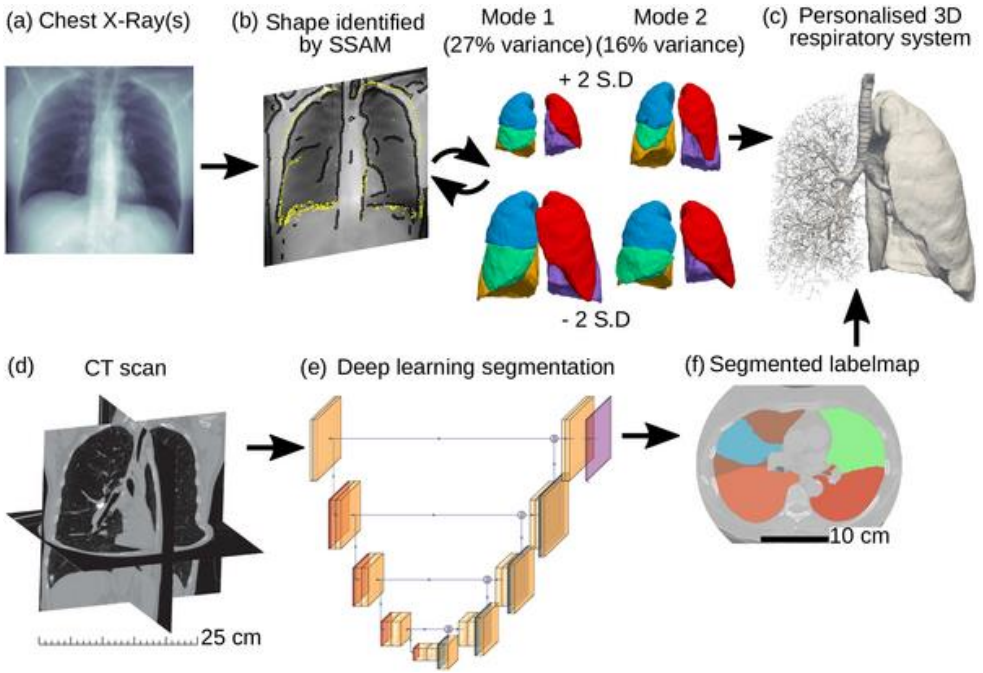


Strategies for generating synthetic computed tomography-like imaging from radiographs: A scoping review
Daniel De Wilde, Olivier Zanier, Raffaele Da Mutten, Michael Jin, Luca Regli, Carlo Serra, Victor E. Staartjes
Under Review

Strategies for generating tomography from radiographs

- Statistical Shape Models (SSM)
  - Bone Imaging, Chest Imaging
  - Usually two projections + manual initialization + estimation of known shapes





Strategies for generating synthetic computed tomography-like imaging from radiographs: A scoping review
Daniel De Wilde, Olivier Zanier, Raffaele Da Mutten, Michael Jin, Luca Regli, Carlo Serra, Victor E. Staartjes
Under Review

Methods based on deep learning:

Project TomoRay

X2CT-GAN: Reconstructing CT from Biplanar X-Rays with Generative Adversarial Networks

Xingde Ying*,†,♭, Heng Guo*,§,♭, Kai Ma♭, Jian Wu†, Zhengxin Weng§, and Yefeng Zheng♭

♭ YouTu Lab, Tencent {kylekma, yefengzheng}@tencent.com
† Zhejiang University {yingxingde, wujian2000}@zju.edu.cn
§ Shanghai Jiao Tong University {ghgood, zxweng}@sjtu.edu.cn

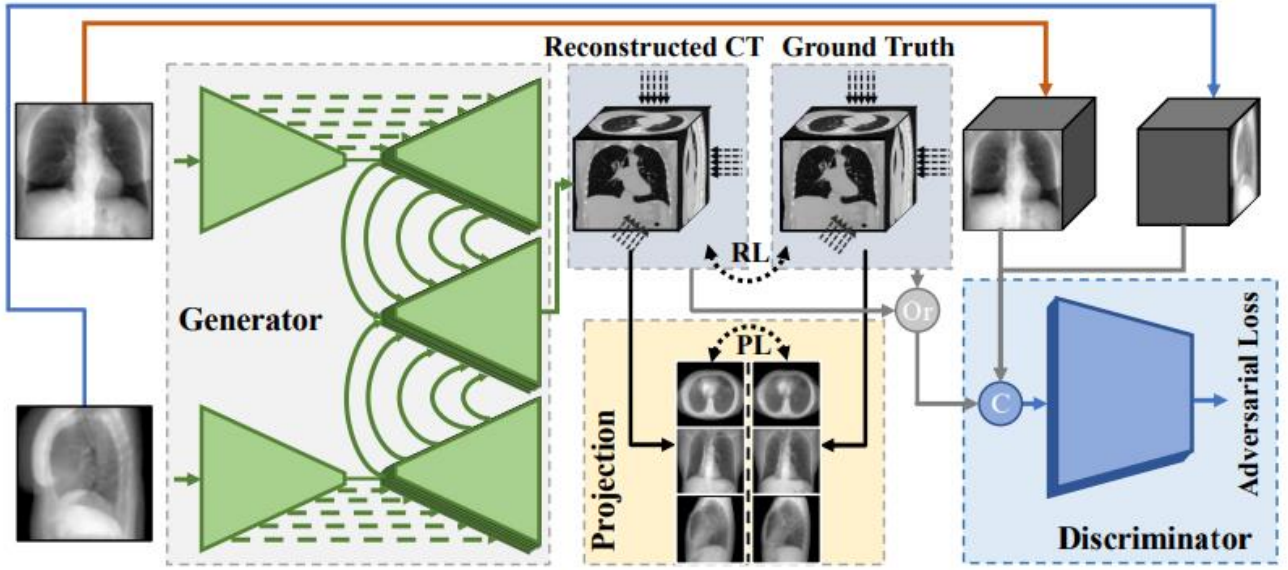
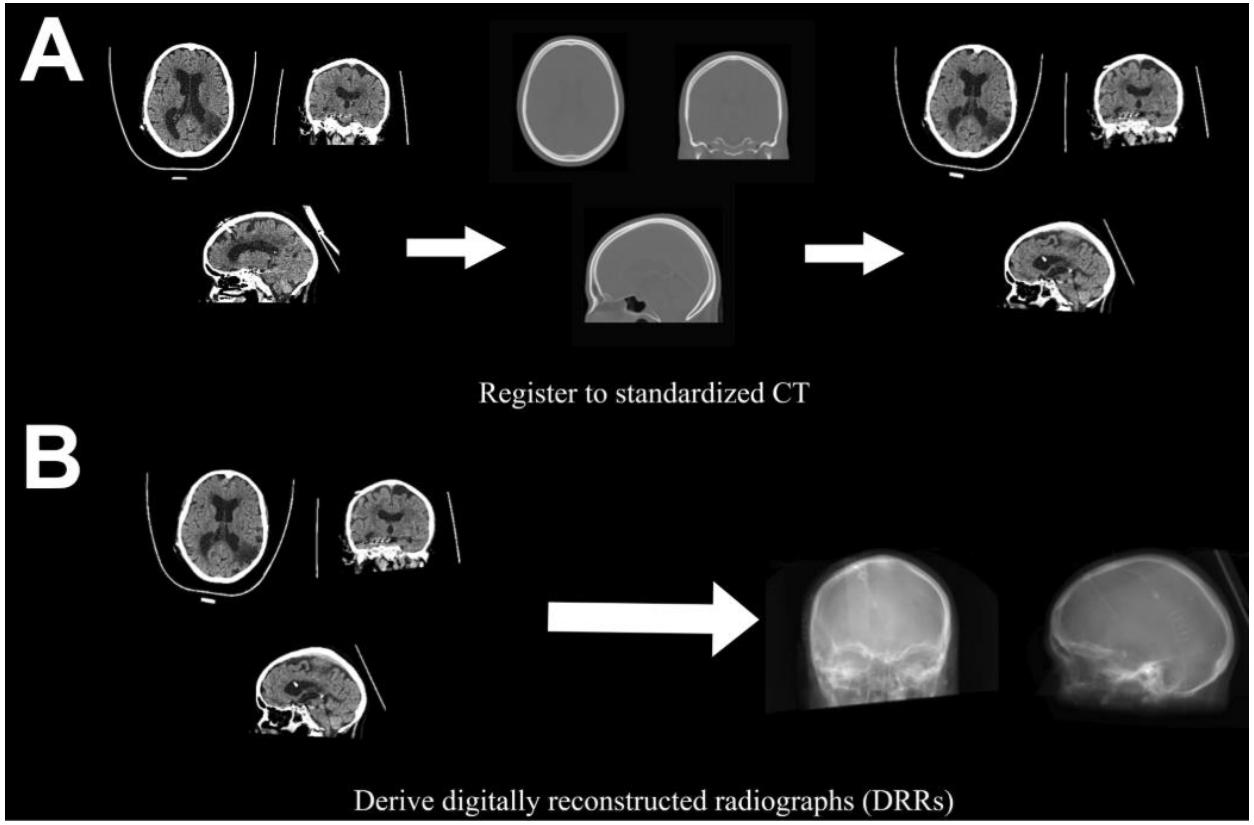
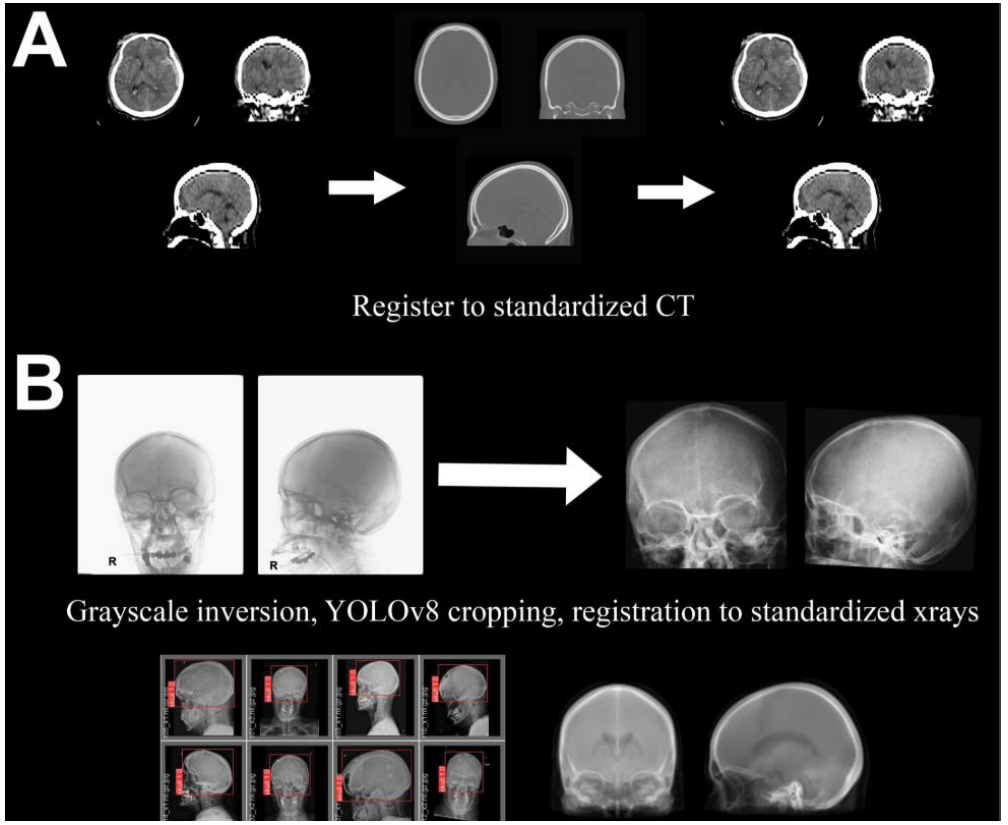


Figure 2. Overview of the X2CT-GAN model. RL and PL are abbreviations of the reconstruction loss and projection loss.

TomoRay: Cranial

– Two models: Based on DRRs (Model 1) or based on x-rays (Model 2)





TomoRay Cranial: Synthesis of Cranial CT Imaging from Biplanar Radiographs using a Generative Adversarial Model
Olivier Zanier, Seung-Jun Ryu, Raffaele Da Mutten, Alessandro Carretta, Giorgio Palandri, Diego Mazzatenta, Luca Regli, Carlo Serra, Victor E. Staartjes
Submitted

TomoRay: Cranial

– Two models: Based on DRRs (Model 1) or based on x-rays (Model 2)

Table 3 Model 1 performance on training and holdout set. Final values reported are the mean and median over the whole dataset.

	PSNR 3D (dB)	SSIM
	Training (n = 150)	
mean ± SD	21.90 ± 1.18	0.873 ± 0.024
median (IQR)	21.87 (21.22 – 22.53)	0.876 (0.855-0.890)
	Internal Validation (n = 35)	
mean ± SD	16.68 ± 1.01	0.799 ± 0.029
median (IQR)	16.39 (15.90 – 17.33)	0.806 (0.778 - 0.820)
	External Validation (n = 50)	
mean ± SD	15.61 ± 1.02	0.782 ± 0.030
median (IQR)	15.63 (14.92 – 16.32)	0.778 (0.766 – 0.804)

SD, standard deviation; IQR, interquartile range

Table 5 Model 2 performance on training and holdout set. Final values reported are the mean and median over the whole dataset.

	PSNR 3D (dB)	SSIM
	Training (n = 1200)	
mean ± SD	17.13 ± 1.05	0.803 ± 0.033,
median (IQR)	17.06 (16.42-17.69)	0.806 (0.782-0.821)
	Internal Validation (n = 123)	
mean ± SD	14.69 ± 0.85	0.717 ± 0.037
median (IQR)	14.77 (14.25 – 15.23)	0.720 (0.696 – 0.741)

SD, standard deviation; IQR, interquartile range

TomoRay Cranial: Synthesis of Cranial CT Imaging from Biplanar Radiographs using a Generative Adversarial Model
Olivier Zanier, Seung-Jun Ryu, Raffaele Da Mutten, Alessandro Carretta, Giorgio Palandri, Diego Mazzatenta, Luca Regli, Carlo Serra, Victor E. Staartjes
Submitted

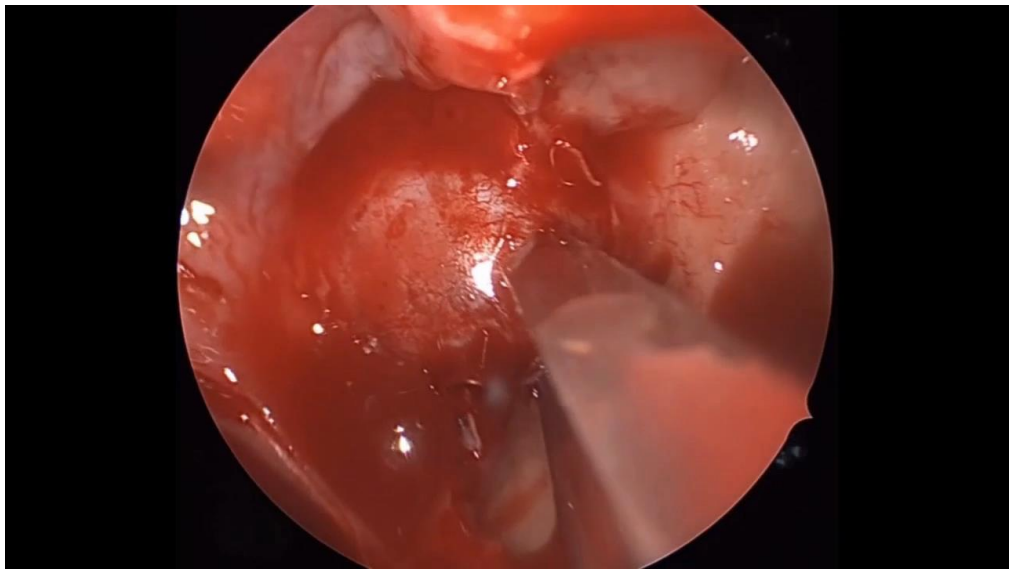
What if we could grant stereoscopic vision in endoscopy?

→ with a 3D endoscope or akin to DaVinci steroscopic camera system

→ requires hardware, costs, etc.

5. Conclusion

When examining the fairly standardized surgical procedure of transphenoidal endoscopic pituitary adenomectomy, this prospective cohort study does not show obvious advantages of 3-D endoscopy for patients using basic outcome parameters including post-operative QoL. To our knowledge this is the first prospective cohort study on the matter, thus corroborating previous retrospective studies of 3-D neuroendoscopy and 3-D endoscopy in general, indicating that the clinical advantage of 3-D endoscopy is less than that reported in preclinical simulator environments. The main advantage of increased depth perception is more likely found in more complex extended transphenoidal skull base procedures though this remains to be proven.



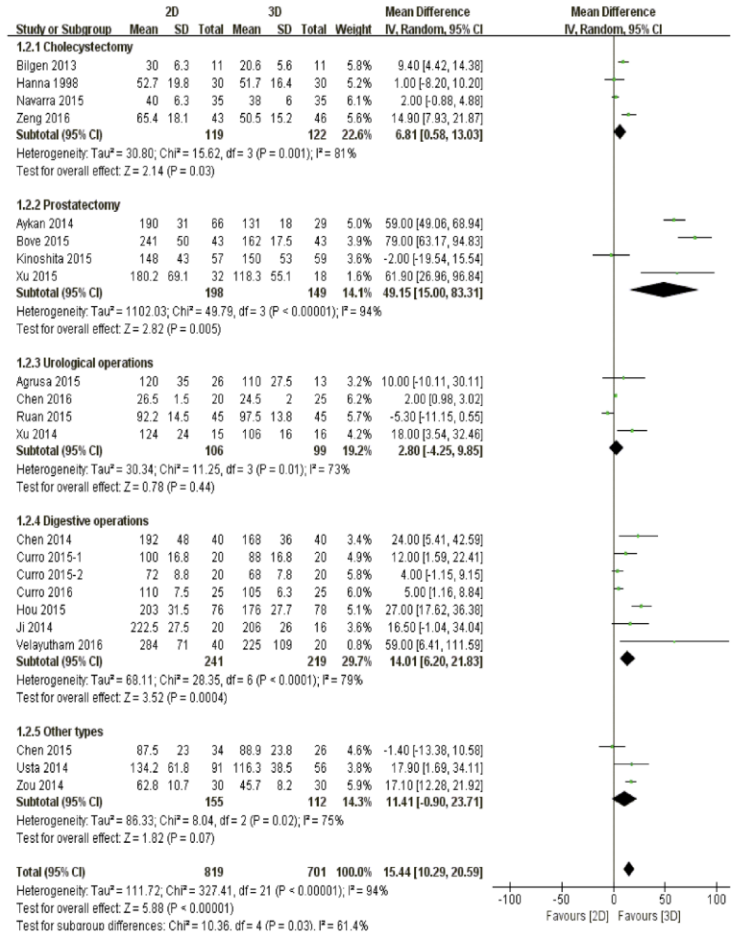
→ **Advantages probably hard to show "evidence-based" as improved outcomes**

→ **Added information: The more you know, the more you see & the more you see, the more you know**

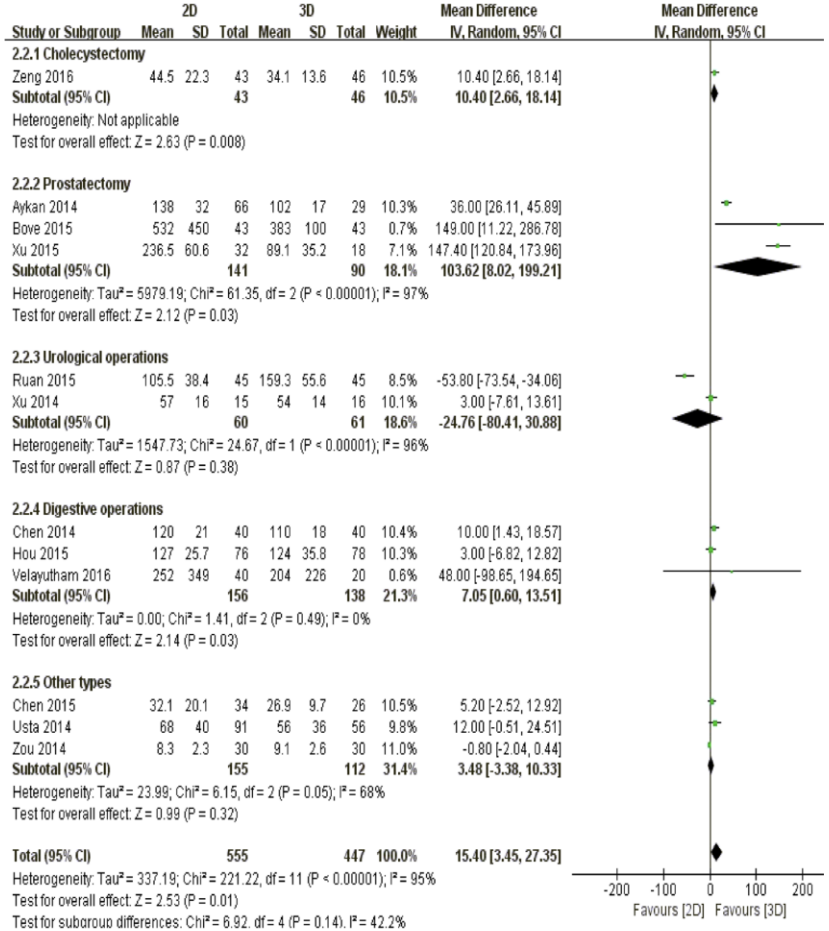
Uvelius et al. *3-D endoscopy in surgery of pituitary adenomas, prospective evaluation of patient gain using basic outcome parameters*, J Clin Neuroscience 2020

Some evidence of 3D endoscopy advantages … from general surgery

Decrease in surgical time



Decrease in blood loss



Uvelius et al. *Two-dimensional versus three-dimensional laparoscopy in surgical efficacy: a systematic review and meta-analysis*, Oncotarget 2016

Comparison of methods

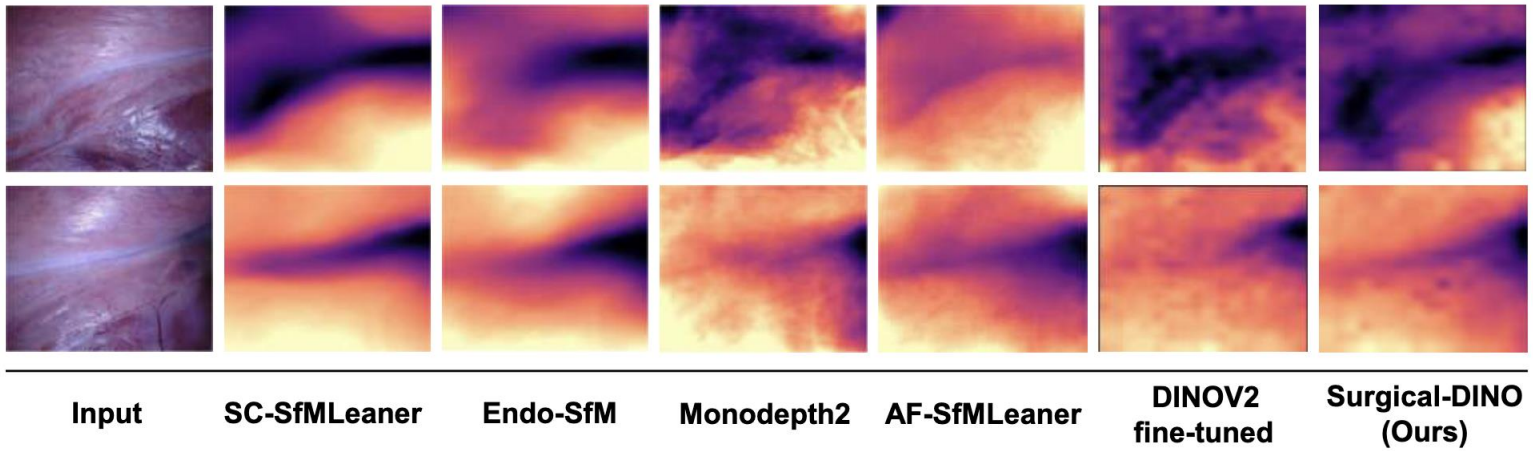


Fig. 3: Qualitative depth comparison on the SCARED dataset.

Table 1: Quantitative depth comparison on the SCARED dataset of SOTA depth estimation methods. The best results are in bold. The second-best results are underlined.

Method	Abs Rel ↓	Sq Rel ↓	RMSE ↓	RMSE log ↓	δ ↑
SfMLearner [20]	0.079	0.879	6.896	0.110	0.947
Fang et al. [21]	0.078	0.794	6.794	0.109	0.946
Defeat-Net [22]	0.077	0.792	6.688	0.108	0.941
SC-SfMLearner [23]	0.068	0.645	5.988	0.097	0.957
Monodepth2 [24]	0.071	0.590	5.606	0.094	0.953
Endo-SfM [25]	0.062	0.606	5.726	0.093	0.957
AF-SfMLearner [17]	<u>0.059</u>	0.435	4.925	0.082	<u>0.974</u>
DINOv2 [6] (zero-shot)	0.088	0.963	7.447	0.120	0.933
DINOv2 [6] (fine-tuned)	0.060	0.459	<u>4.692</u>	<u>0.081</u>	0.963
Surgical-DINO SSL (Ours)	<u>0.059</u>	<u>0.427</u>	4.904	<u>0.081</u>	0.974
Surgical-DINO (Ours)	**0.053**	**0.377**	**4.296**	**0.074**	**0.975**

Cui et al. *Surgical-DINO: Adapter Learning of Foundation Models for Depth Estimation in Endoscopic Surgery*, arXiv 2024

How does DinoV2 work?

- Open-source base from Facebook AI (released 2023)
- **Foundation model:** learn **generalized representations**
- Self-supervised patch-based **vision transformer** – 142 million images w/o labels
- Huge model with 1 billion parameters «distilled» into several specialized models
- DinoV2-Depth: Base model (unchanged) as encoder + light decoder network

- **Intuition:**
  - Experienced master surgeon who has not done *this* particular procedure before
  - Learns to pick up cues in lighting, shading, movement, etc.

Quantitative Performance?

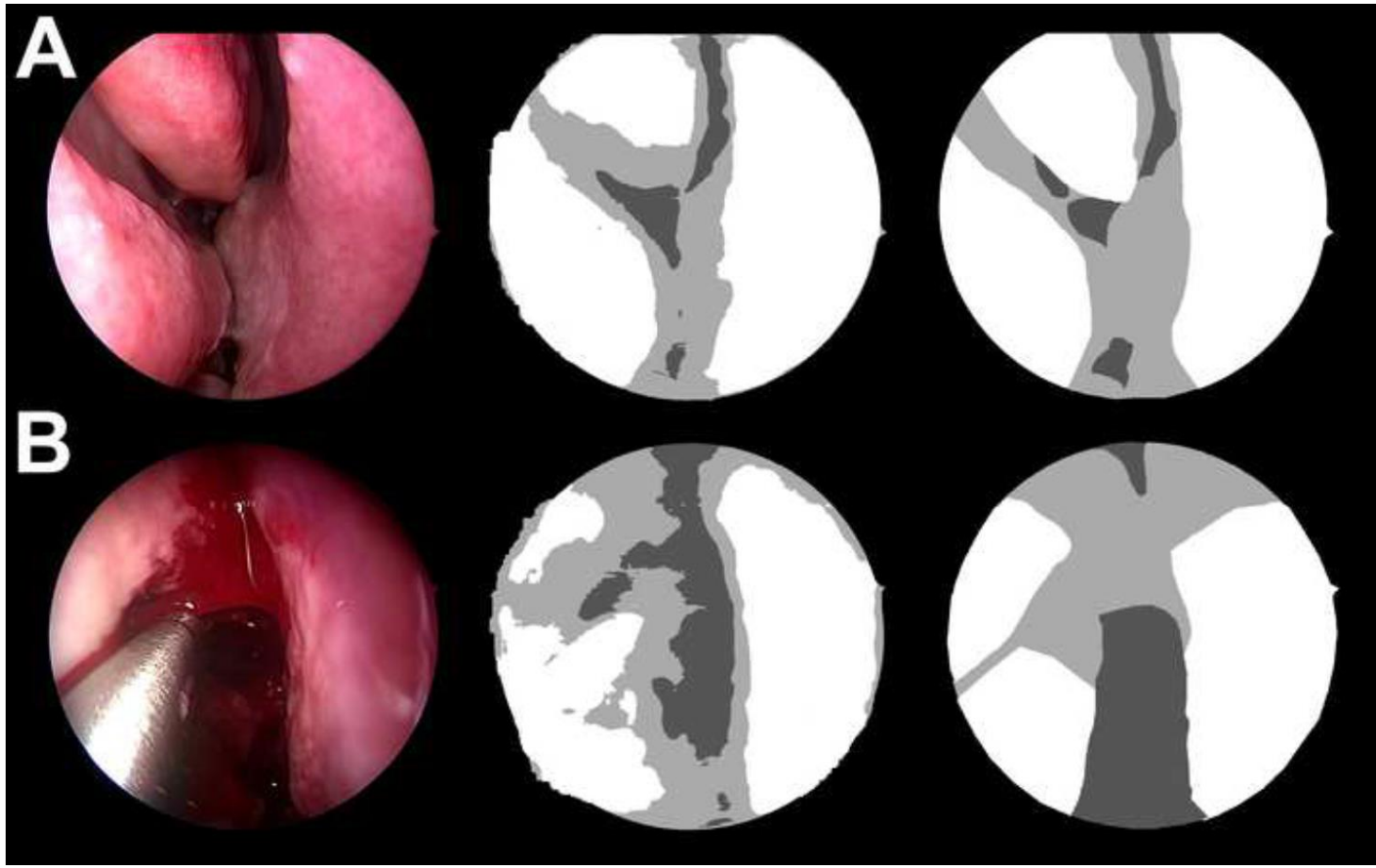


Figure 2: Illustration of the semi-quantitative validation process.

From left to right: Input image, model depth estimation in thresholded format, manual ground truth.

A: well-performing example, B: inferior example

Zanier, Serra, Staartjes; *Real-time intraoperative depth estimation in transsphenoidal surgery using deep learning: a pilot study*, submitted (2025)

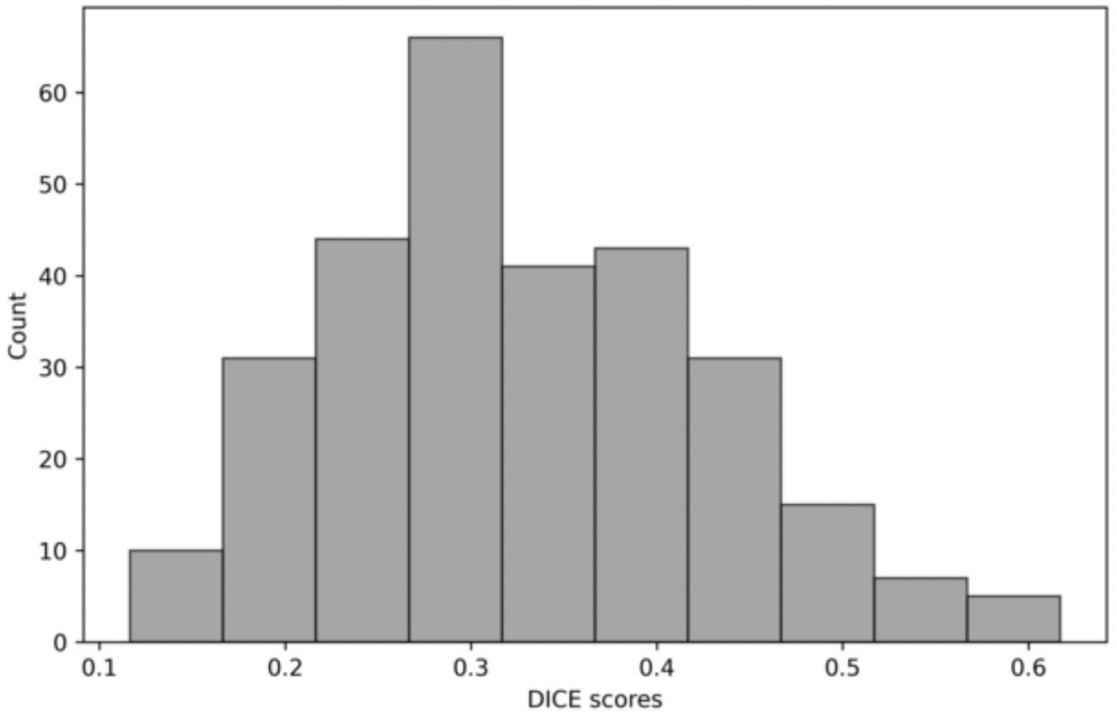
Quantitative Performance?

Table 1: Semi-quantitative performance on evaluation dataset

	Validation Set (n = 293)
mean ± SD	0.33 ± 0.10
median (IQR)	0.31 (0.25 – 0.39)
maximum	0.62
minimum	0.12

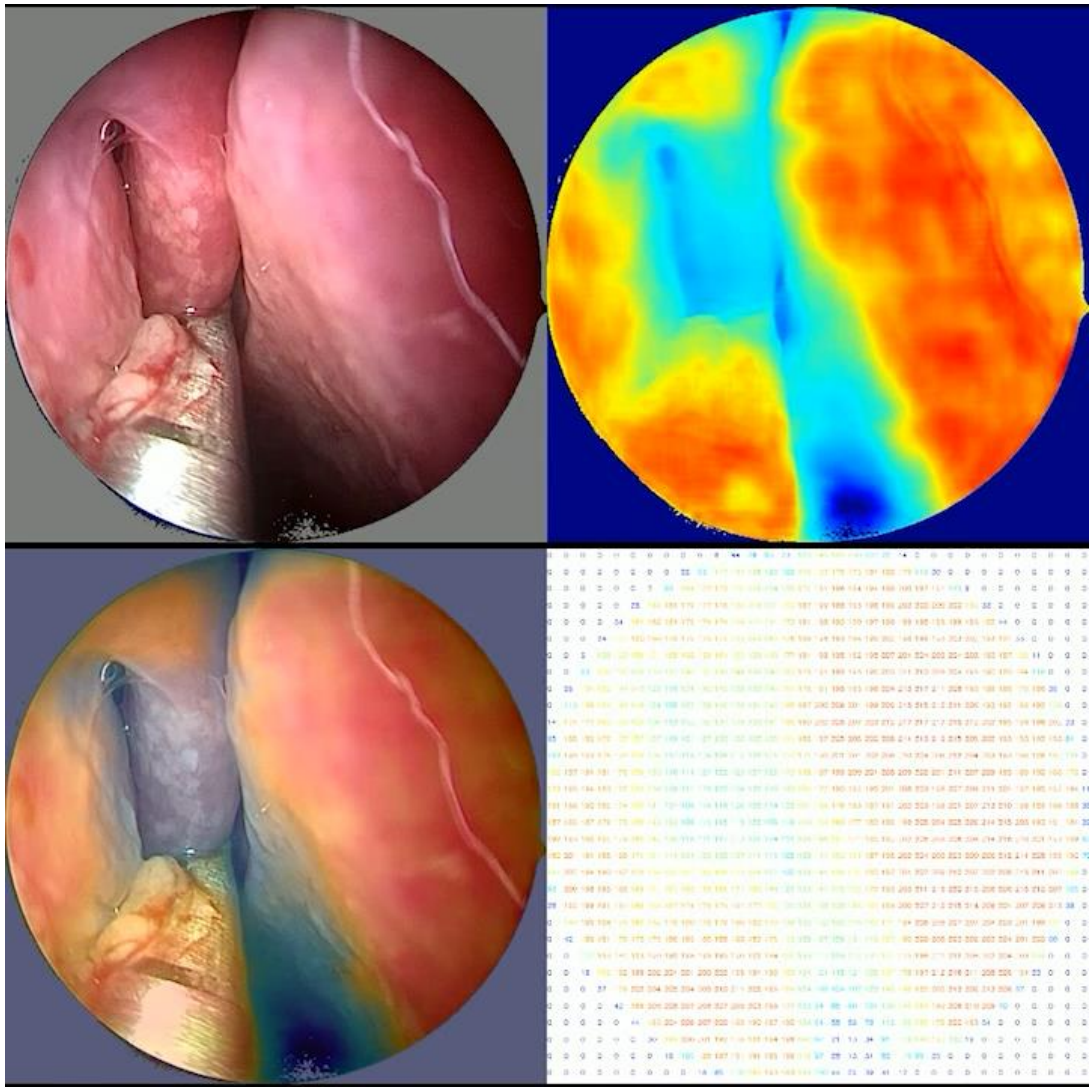
SD: standard deviation, IQR: interquartile range

Figure 3: Distribution of DICE scores over the validation dataset.



Zanier, Serra, Staartjes; *Real-time intraoperative depth estimation in transsphenoidal surgery using deep learning: a pilot study*, submitted (2025)

How to get from a synthetic 3D image to binocular vision?
DinoV2 in action



Zanier, Serra, Staartjes; *Real-time intraoperative depth estimation in transsphenoidal surgery using deep learning: a pilot study*, submitted (2025)