

PENGWIN 2026

MICCAI2026

Peripelvic Fracture Segmentation and Reduction Planning

ROSSUM ROBOT · BEIHANG UNIVERSITY · KIT Karlsruher Institut für Technologie · JOHNS HOPKINS UNIVERSITY · JISHUITAN HOSPITAL · 香港中文大學 CUHK

Task 3 - Reduction Planning

Residual Coordinate Assembly Transformer

sano · AGH · MultiGeoMed

Research Team

MultiGeoMed



Wojciech Szymanski
PhD Student



Gniewosz Drwiega
Postdoctoral Researcher



Marek Wodzinski
Associate Professor

Affiliations

sano
Research Center

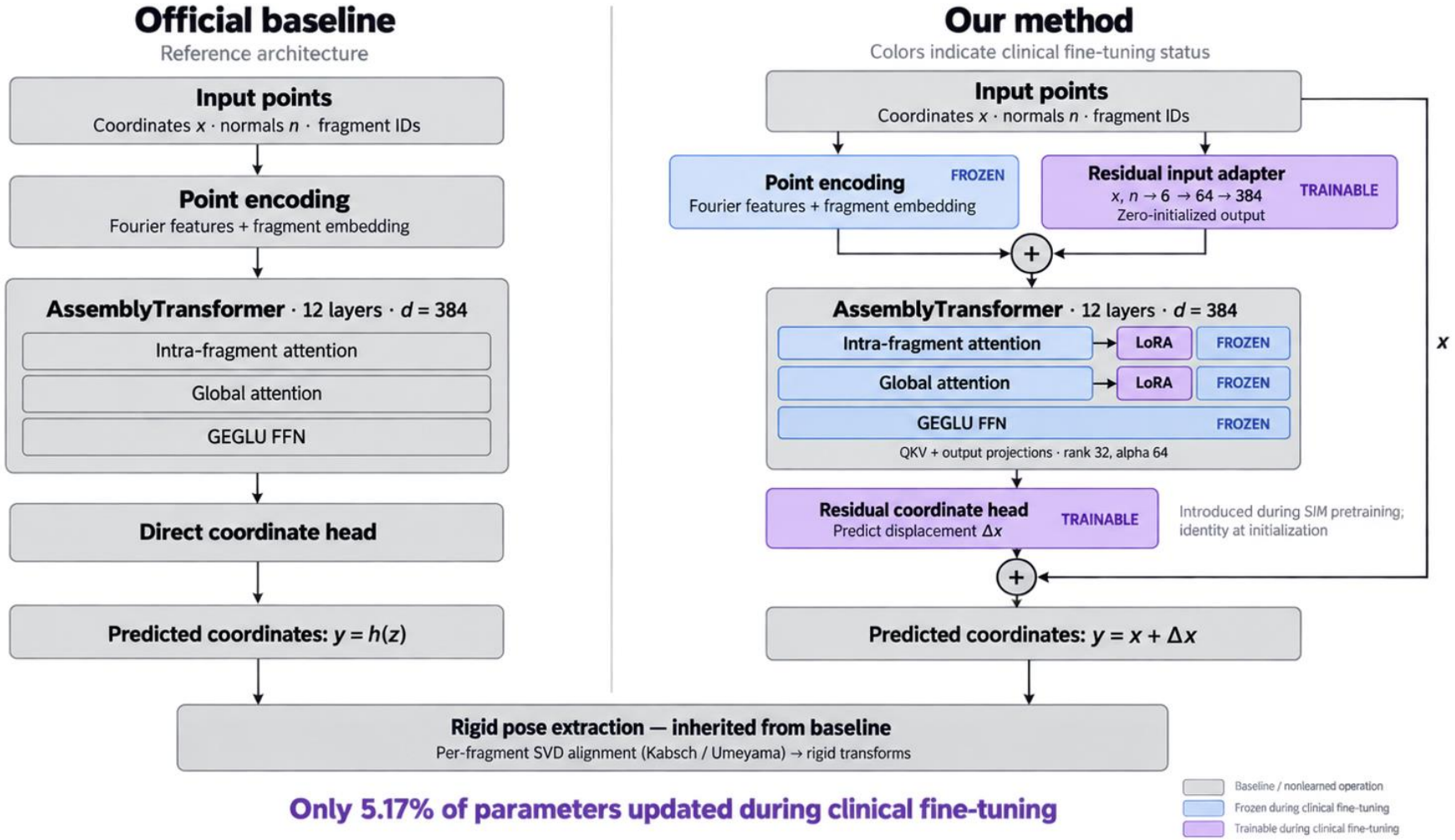
AGH
University of Krakow

Acknowledgments

CYFRONET
HPC Infrastructure

NATIONAL SCIENCE CENTRE POLAND
Research Funding

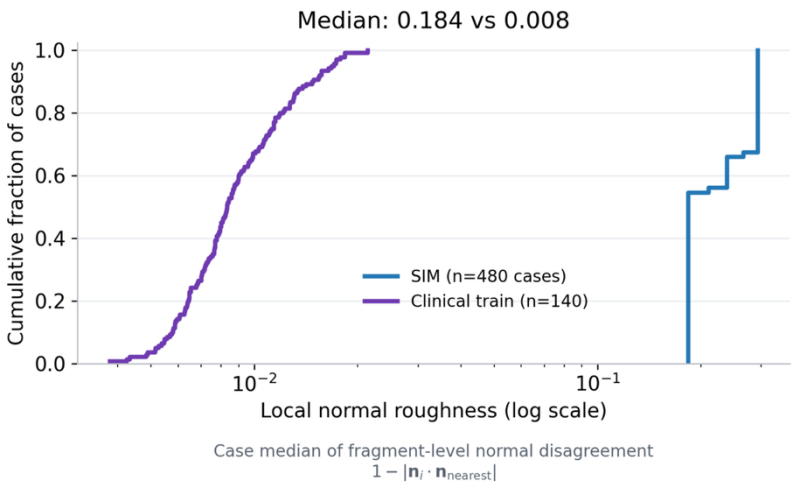
From Assembly Transformer to Clinical Adaptation



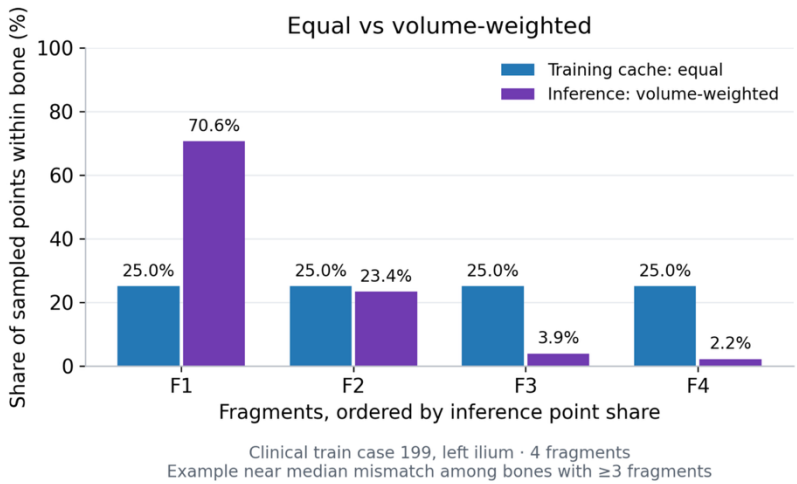
sano AGH MultiGeoMed

Understanding the Simulation-Clinical Geometry Gap

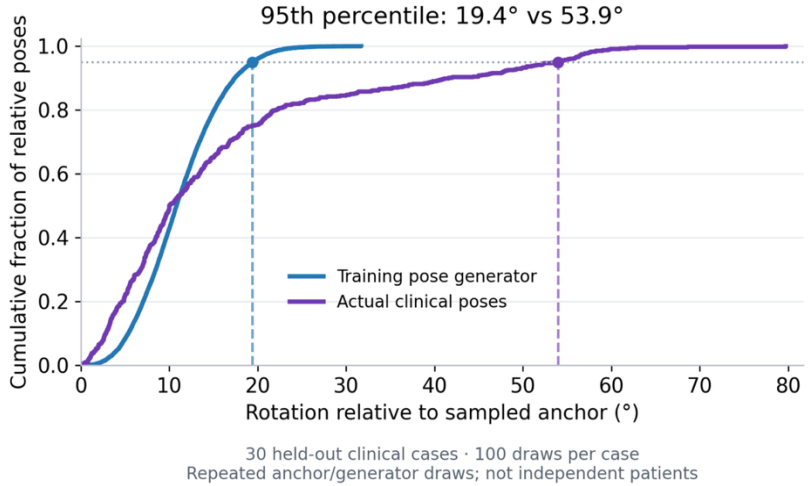
Surface Roughness



Fragment Representation



Fragment Displacement



Key Observations:

- ✓ **Surface Roughness:** Clinical surfaces have substantially smoother normals than simulated surfaces.
- ✓ **Fragment Representation:** Small fragments are overrepresented during training compared with inference.
- ✓ **Fragment Displacement:** Clinical cases exhibit a much broader rotation range than the training augmentation covers.

sano AGH MultiGeoMed

What We Tested and What We Retained

Experiment Area	Variants Explored	Final Method
Backbone	Pure PTv3 PTv3 + Transformer hybrids	AssemblyTransformer + Residual coordinate head
Clinical adaptation	Full / partial fine-tuning LoRA ranks, FFN LoRA Bottleneck / input adapters	Attention LoRA Rank 32 · alpha 64 + Residual input adapter
Training objective	Fragment-balanced loss Pose auxiliary losses Sampling consistency	Point-wise coordinate MSE
Data augmentation	Pose mixtures / curriculum Clinical geometry resimulation Normal corruption / erosion	Baseline augmentations SIM fragment dropout: 0.10 → 0.15 No real-pose replay

sano AGH MultiGeoMed

Training, Epoch Selection, and Full-Data Refit

1

SIM Pretraining
480 source anatomies with simulated fractures
Residual AssemblyTransformer
1,000-epoch run | Select checkpoint at epoch 914

2

Method Development
170 clinical cases: 126 train / 14 validation / 30 held-out test
Test selection matched dataset-level fragment counts across SA / LI / RI
Develop and compare fine-tuning recipes

3

Epoch Selection
Re-split all 170 clinical cases: 153 train / 17 validation
Validation: 17 clinical + 17 SIM cases
Select epoch 182 for the final refit

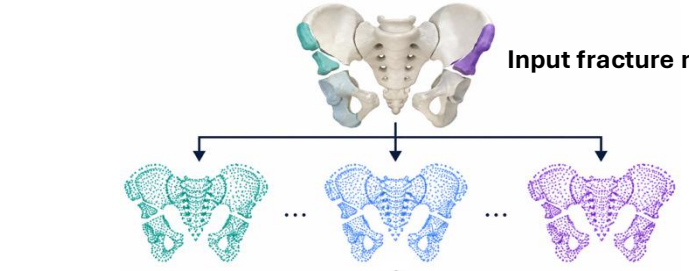
4

Full-Data Refit
Train on all 170 clinical cases, mixed with simulated data
Fresh initialization from SIM checkpoint 914
Save the preselected epoch 182, no checkpoint selection on overlapping validation data

Shared settings for stages 2-4
Each fine-tuning run starts from SIM checkpoint 914 | 50% SIM / 50% clinical training mixture | 300-epoch OneCycle schedule

sano AGH MultiGeoMed

Sampling Ensemble for Robust Inference



Input fracture mesh

x25 independent surface samplings

One shared model checkpoint
Iterative rigid reduction up to 10 steps per sampling

Common reference frame
All predictions relative to the same reference fragment

Per-fragment pose aggregation

Transformed centroid
Coordinate-wise median

Rotation
Quaternion mean

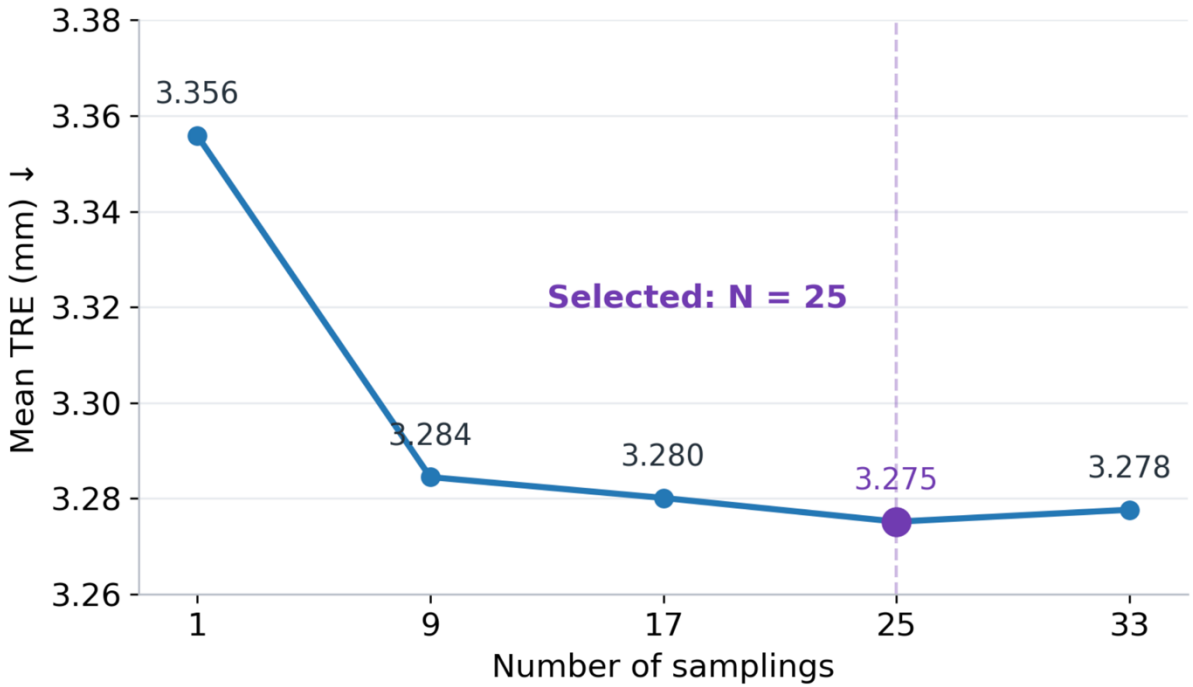
Reconstruct translation from rotation and centroid



Final rigid transform

Effect of Sampling Ensemble

Same checkpoint · Same 30 held-out cases



N = 1 → 25: 2.4% lower mean TRE

Development checkpoint: epoch 180 · One evaluation run
Local evaluation; not the final full-data refit

One checkpoint | Mulitple surface samplings | Robust pose aggregation

sano AGH MultiGeoMed

Training and Deployment Setup

Simulation Pretraining	Clinical Fine-Tuning	Challenge Inference
4 x NVIDIA GH200 GPU	2 x NVIDIA GH200 GPU	1 x NVIDIA T4 GPU
48-hour job	2-hour job	10-min limit per case
1,000-epoch schedule	300-epoch schedule	80 final test cases
Batch: 8/GPU Effective: 64	Batch: 8/GPU Effective: 32	1 case/invocation 25 samplings



Experimental campaign:

- ✓ ≈ **150** SLURM jobs
- ✓ ≈ **3,000** GPU-hours
- ✓ HPC infrastructure: **Cyfronet / Helios**

sano AGH MultiGeoMed

Challenge Results

4th

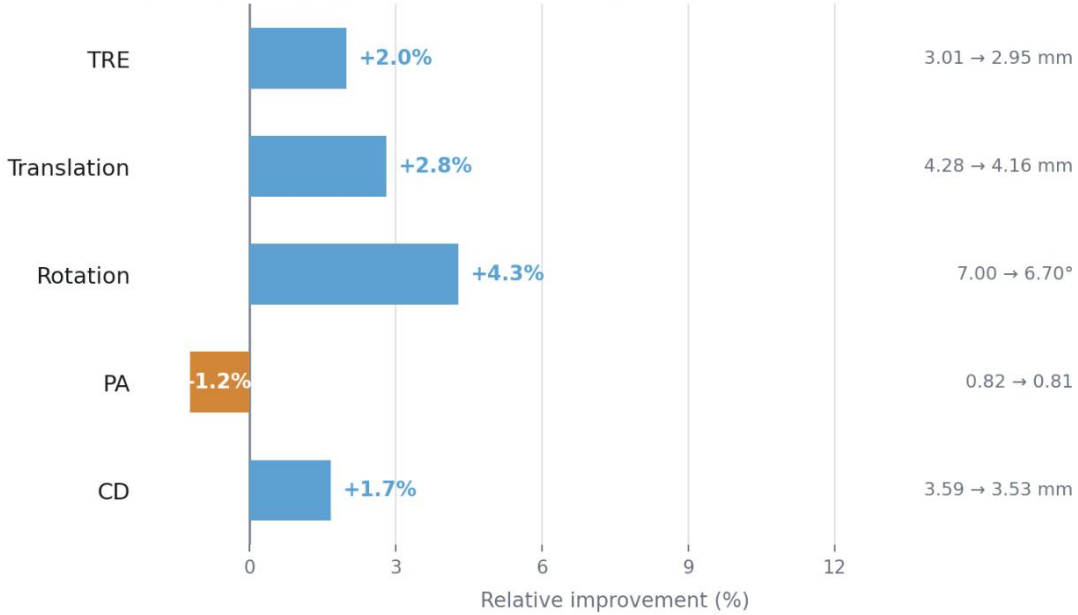
FINAL LEADERBOARD

80 hidden test cases

2.95 mm	4.16 mm	6.70°	0.81	3.53 mm
Target Registration Error	Translation Error	Rotation Error	Part Accuracy	Chamfer Distance

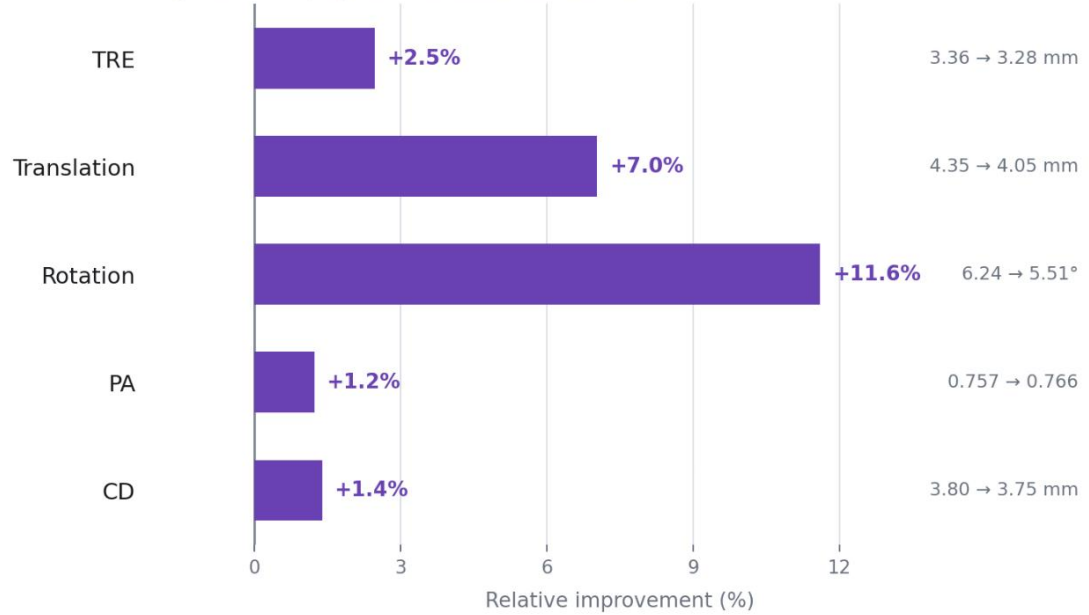
Full-Data Refit

Two official submissions on the same hidden final test set



25-Sample Aggregation

Single vs 25 samplings on 30 held-out clinical cases



Positive values indicate improvement. PA uses higher-is-better orientation; all other metrics use lower-is-better.

sano AGH MultiGeoMed

Key Takeaways

1

Data-driven training and inference

A zero-initialized residual coordinate head provided a stable starting point for simulation pretraining.

2

Targeted clinical adaptation

Attention LoRA and a residual input adapter improved domain transfer while keeping the pretrained backbone frozen.
More complex PTv3 hybrids brought no measurable benefit.

3

Data-driven training and inference

Geometry analysis guided clinical adaptation. Held-out epoch selection, fixed-epoch full-data refitting, and 25-sample aggregation produced the final submission.

Code & Reproducibility

https://github.com/wojtek-szymanski/**PENGWIN2026_Task3_Reduction**

Wojciech Szymanski
Sano Centre for Computational Medicine
AGH University of Krakow

szymanski@agh.edu.pl

sano AGH MultiGeoMed