A Patient-Specific Self-supervised Model for Automatic X-ray/CT Registration

Baochang Zhang[1,2], Shahrooz Faghihroohi[1], Mohammad Farid Azampour[1], Shuting Liu[1], Reza Ghotbi[3], Heribert Schunkert[2,4], and Nassir Navab[1]

[1] Computer Aided Medical Procedures, Technical University of Munich, Munich, Germany
[2] German Heart Center Munich, Munich, Germany
[3] HELIOS Hospital west of Munich, Munich, Germany
[4] German Centre for Cardiovascular Research, Germany

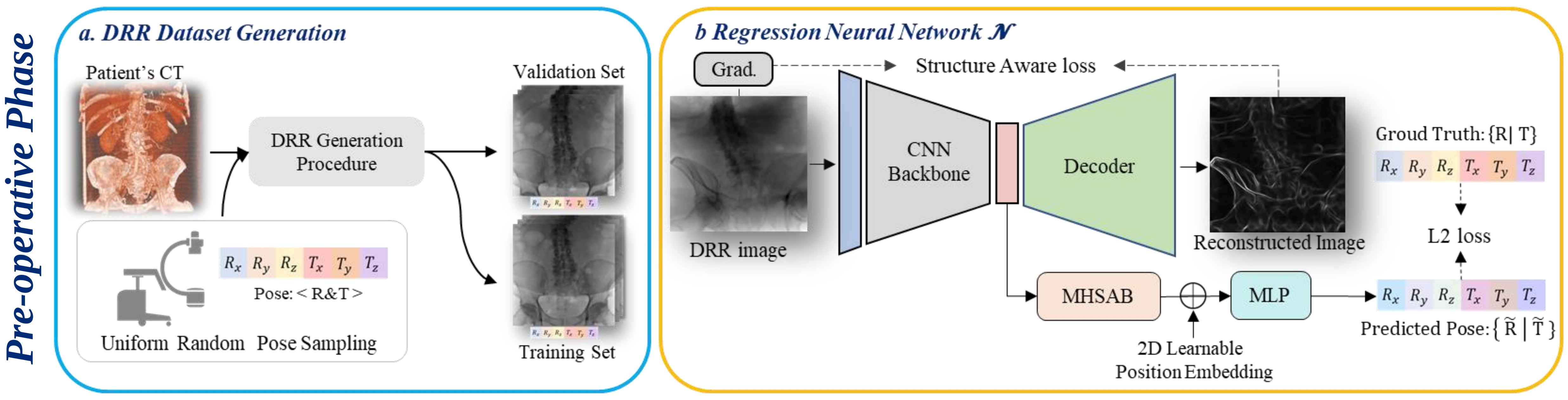
Technische Universität München

MICCAI 2023 Vancouver

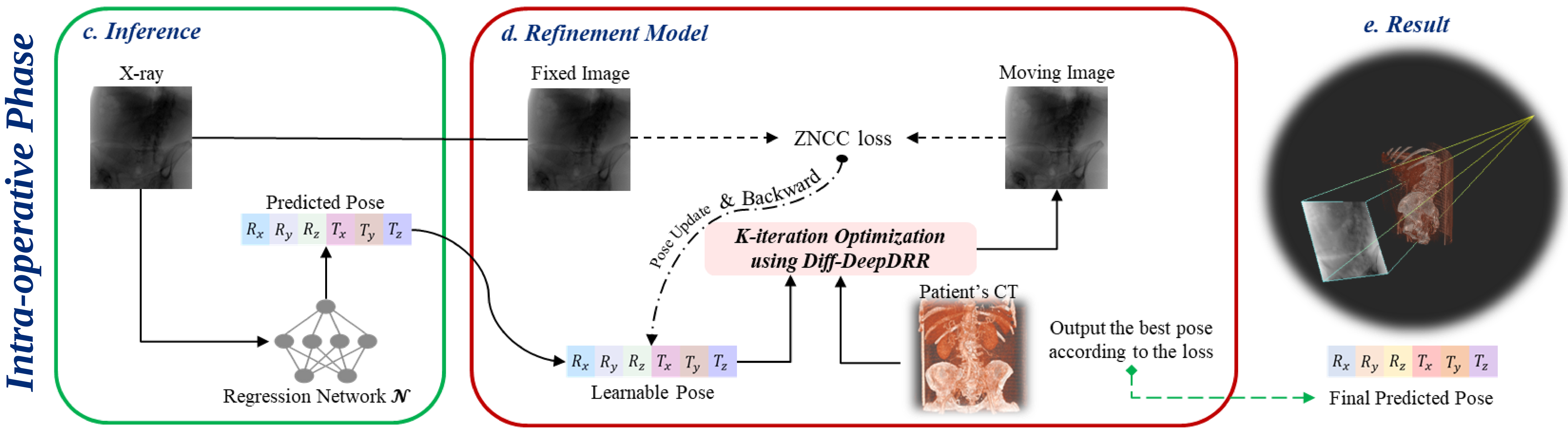
INTRODUCTION

- Motivation
  - Augmentation of intra-operative X-ray images using the pre-operative data has potential to reduce procedure time and improve patient outcomes in minimally invasive procedures.
- Limitations of Existing methods:
  - Manual initialization [1].
  - Requirement of manual annotation for training [2].
  - Reliance on conventional intensity-based refinement method [2].
- Goals
  - Make full use of DRR generation.
  - Patient-specific pose regression model for initialization.
  - Diff-DeepDRR for refinement.
  - Achieving high performance on both DRR and Xray.

METHODOLOGY



- **Proposed pose regression network**
  - Regularized Autoencoder constrains the features extracted by CNN backbone to contain necessary structure information
  - Multi-Head Self-Attention Block captures patient-specific salient information automatically
  - 2D Positional Embedding helps to be more sensitive to spatial transformation
- **Incremental Learning strategy for training**
  - Avoid over-fitting
  - Promote network performance.
- **Training phase can be done within 1 hour using an NVIDIA GPU (Quadro RTX A6000)**



- **Diff-DeepDRR**

$$D_i = \int \overbrace{P_0(E)}^{\text{Xray Spectrum}} \cdot \exp\left\{-\sum_{m\in M} \underbrace{\delta(m, M(V))}_{\text{Materials Mask}} \cdot \overbrace{\varphi_m(E)}^{\text{Mass Attenuation Coefficient}} \cdot \int \underbrace{\rho(V)}_{\text{Mass density}} dl_i\right\} dE$$

  - Follows physics
  - Differentiable
  - High speed (256x256 image in 15.6 ms)
- **The proposed method achieves an average run-time of around *2.5s***

EVALUATION

Comparison Results

Table 1. Comparison Results on DRR dataset

-	mTRE(mm)	SR(%)
[3]	7.68	67.13
[2]	4.21	91.9
Ours	**2.67**	**100**

Table 2. Comparison Results of refinement methods with same initialization on Xray dataset.

-	NCC	SSIM	CSS
Method in [2]	0.9773	0.9342	0.9375
Ours	**0.9880**	**0.9469**	**0.9503**

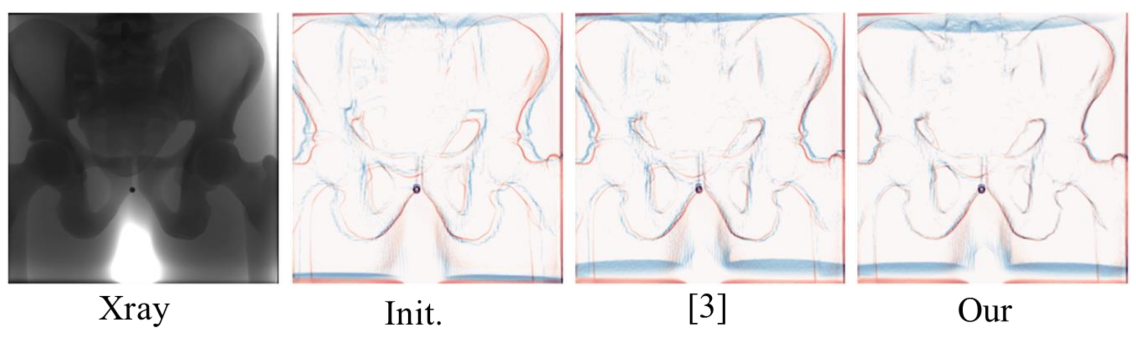


Fig 1. Results of refinement methods with same initialization on Xray Image.

More Results of Our Method

Table 3. The results of ablation study & our method's results on 6 DRR datasets. # indicates previous surgical implants.

–	mTRE↓ (mm)		mPD↓ (mm)		Rx↓ (degree)		Ry↓ (degree)		Rz↓ (degree)		Tx↓ (mm)		Ty↓ (mm)		Tz↓ (mm)		SR↑ (%)
	mean	std	mean	std	mean	std	mean	std	mean	std	mean	std	mean	std	mean	std	
A	6.71	2.75	6.83	2.05	0.98	0.78	0.93	0.76	0.71	0.56	2.44	2.05	2.57	1.99	3.92	3.05	88.66
A+B	6.13	2.50	6.39	2.07	0.93	0.78	0.99	0.79	0.62	0.52	2.10	1.82	2.33	1.91	3.49	2.73	90.97
A+C	6.17	2.65	6.12	1.98	1.01	0.80	0.85	0.68	0.70	0.58	1.99	1.75	2.05	1.74	3.72	2.88	91.34
A+B+C	5.92	2.57	6.01	2.02	1.02	0.80	0.83	0.65	0.64	0.54	1.92	1.75	2.05	1.78	3.39	2.71	93.31
[A+B+C]*	5.43	2.18	5.80	2.05	0.90	0.73	0.82	0.65	0.57	0.47	1.80	1.53	1.80	1.50	2.89	2.26	95.25
[A+B+C]*+D proposed	**2.85**	**1.91**	**2.92**	**1.93**	**0.41**	**0.53**	**0.31**	**0.39**	**0.24**	**0.35**	**1.00**	**0.83**	**0.90**	**0.84**	**1.55**	**1.35**	**100.00**
dataset1	2.85	1.91	2.92	1.93	0.41	0.53	0.31	0.39	0.24	0.35	1.00	0.83	0.90	0.84	1.55	1.35	100.00
dataset2	2.54	1.28	3.08	1.72	0.38	0.55	0.35	0.38	0.22	0.29	0.65	0.60	0.90	0.51	1.38	0.88	100.00
dataset3	2.71	1.14	3.30	1.45	0.26	0.31	0.33	0.40	0.23	0.26	0.84	0.64	1.04	0.71	1.51	1.09	100.00
dataset4	2.72	1.32	3.12	1.53	0.27	0.48	0.26	0.33	0.23	0.32	0.90	0.86	1.04	0.66	1.58	1.10	100.00
dataset5	2.73	1.03	3.09	1.22	0.35	0.35	0.35	0.33	0.24	0.22	0.74	0.69	0.97	0.58	1.52	1.09	100.00
dataset6#	2.50	1.48	2.57	1.67	0.36	0.43	0.27	0.39	0.20	0.27	0.75	0.64	0.86	1.00	1.44	1.20	100.00

A: Efficientnet-b0 as backbone; **B:** Regularized Autoencoder; **C:** MHSAB & Positional Embedding; **D:** Refinement Model; ***:** Trained via Incremental Learning Strategy.

Table 4. The results of our method on X-ray Cases

-	*Case1*	*Case2*	*Case3*	*Case4*	*Case5*	*Case6*	*Case7*	*Case8*	*Case9*	*Case10*	*Mean*	*Std*
PD↓	0.6762	1.8400	1.4251	2.8431	0.6770	1.3244	1.2777	1.4181	1.5259	2.5655	1.5573	0.6687
NCC↑	0.9885	0.9752	0.9867	0.9941	0.9858	0.9870	0.9888	0.9943	0.9913	0.9880	0.9880	0.0051
SSIM↑	0.9395	0.9220	0.9348	0.9736	0.9424	0.9456	0.9436	0.9616	0.9649	0.9412	0.9469	0.0146
CSS↑	0.9427	0.9321	0.9392	0.9750	0.9453	0.9481	0.9463	0.9630	0.9664	0.9448	0.9503	0.0127

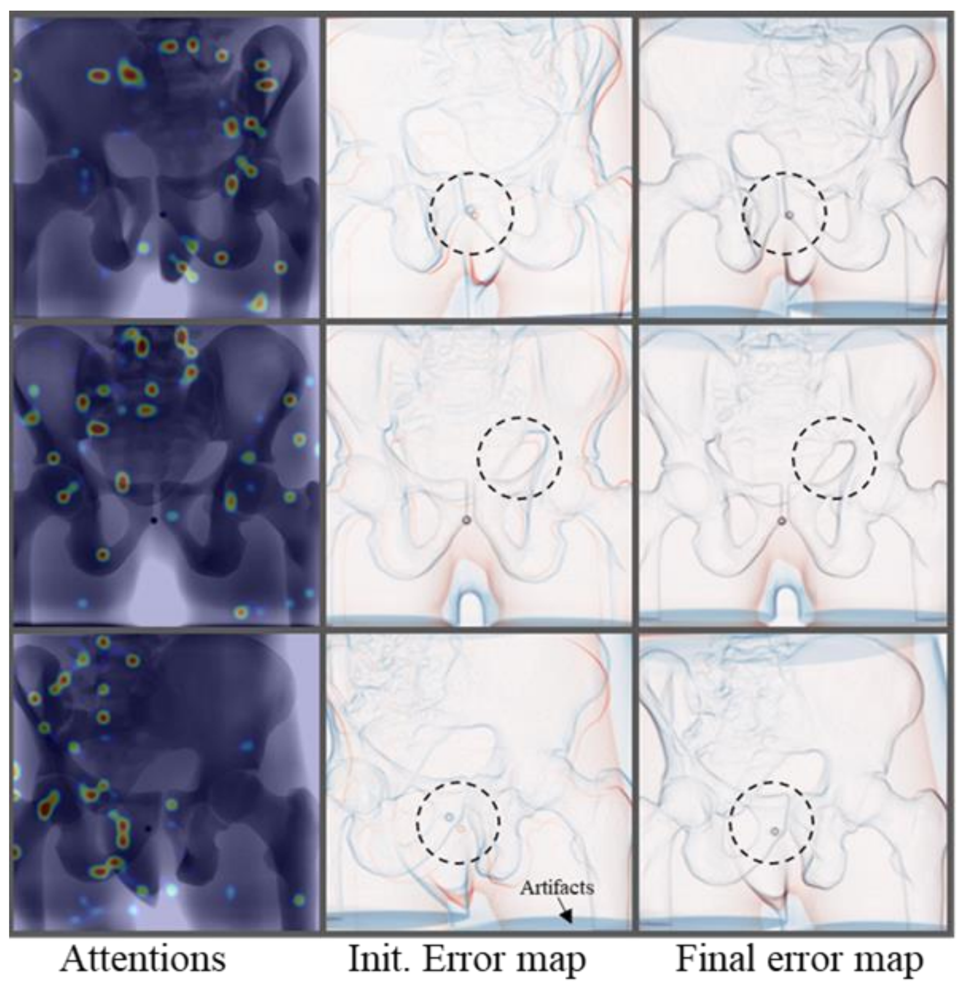


Fig 2. Qualitative Results on X-ray cases.

- Attention maps come from MHSAB;
- Init. Error maps means the result of pose regression model;
- Final Error maps means the result after refinement

REFERENCES

1. Meng, C., Wang, Q., Guan, S., Sun, K., Liu, B.: 2d-3d registration with weighted local mutual information in vascular interventions. IEEE Access 7, 162629–162638(2018)
2. Grimm, M., Esteban, J., Unberath, M., Navab, N.: Pose-dependent weights and domain randomization for fully automatic x-ray to ct registration. IEEE Transactions on Medical Imaging 40(9), 2221–2232 (2021)
3. Salehi, S.S.M., Khan, S., Erdogmus, D., Gholipour, A.: Real-time deep pose estimation with geodesic loss for image-to-template rigid registration. IEEE transactions on medical imaging 38(2), 470–481 (2018)

CONCLUSION

- we present a patient-specific and self-supervised end-to-end approach for automatic X-ray/CT rigid registration.
- Our method effectively addresses the primary limitations of existing methods, such as requirement of manual annotation, dependency on conventional derivative-free optimization, and patient-specific concerns.
- The evaluation results of our proposed method illustrates its superiority and its ability to generalize to X-rays even when trained solely on DRRs.