Intraoperative Registration by Cross-Modal Inverse Neural Rendering

No Author Given

No Institute Given

Abstract. We present in this paper a novel approach for 3D/2D intraoperative registration during neurosurgery via cross-modal inverse neural rendering. Our approach separates implicit neural representation into two components, handling anatomical structure preoperatively and appearance intraoperatively. This disentanglement is achieved by controlling a Neural Radiance Field's appearance with a multi-style hypernetwork. Once trained, the implicit neural representation serves as a differentiable rendering engine, which can be used to estimate the surgical camera pose by minimizing the dissimilarity between its rendered images and the target intraoperative image. We tested our method on retrospective patients' data from clinical cases, showing that our method outperforms state-of-the-art while meeting current clinical standards for registration.

1 Introduction

The use of surgical navigation techniques through patient-to-image registration has become a standard practice in neurosurgery [14]. It allows neurosurgeons to visualize preoperative imaging during the operation, enabling them to achieve a maximal safe tumor resection that is highly correlated with patients' chances of survival [21] and has been shown to reduce risks of postoperative neurological deficits [2]. In this paper, we address patient-to-image registration in neurosurgery as a 6-degrees-of-freedom (DoF) pose estimation problem. This process involves aligning preoperative Magnetic Resonance (MR) images with intraoperative surgical views of the brain surface revealed after a craniotomy and acquired using a camera. Different from previous approaches [10, 20, 15, 6, 17, 5], our proposed method rely solely on imaging already available in the operating room, eliminating the need of cumbersome and time-consuming imaging acquisitions or optical tracking systems.

In most cases, tackling 3D/2D registration in surgery involves bridging the preoperative to intraoperative modality gap and resolving 3D to 2D projection ambiguities. 3D shape reconstruction of surgical scenes has been proposed as a way to tackle both issues at once [22]. They provide a modality-agnostic 3D surface representation of the surgical scenes (3D point clouds or meshes), and re-cast the 3D/2D registration as a 3D-3D point set registration problem where robust methods exist [22]. Other approaches rely on landmark-based matching [13, 3, 9], where anatomical landmarks, extracted from both modalities can act as image

abstractions. Iterative registration methods can then be applied with a closed-form when landmarks are paired which often involve surgical pointers, tracking systems, or *in-vivo* markers [13, 3, 9].

During neurosurgery, the brain surface is revealed and viewed using a surgical camera (microscope), and although the field of view is limited w.r.t other organs where a larger part of the organ is visible, it has the advantage of having visible vessels at the cortical level. These vessels have been used as salient sources of information to drive 3D/2D registration. In [4], segmentations of cortical vessels are used to drive a 3D/2D non-rigid registration. Instead of using segmentations, the authors in [13] proposed to manually trace vessels at the brain surface that match preoperative scans. This method, however, involves a tracked pointer. Other methods proposed to pre-compute the set of plausible transformations preoperatively, using atlas-based approaches [23] or by learning to estimate poses and appearances [3]. However, they are trained in a patient-specific manner on a pre-defined set of transformations and may fail with out-of-distribution transformations. Outside of surgical applications, the field of computer vision has seen the emergence of differentiable rendering [7] and Neural Radiance Fields (NeRFs) [16] making approaches for 6-DoF pose estimation more robust. For instance, methods such as iNeRF [26], PI-NeRF [11], and Parallel Inversion [12] showed that implicit neural representations outperform conventional regression-based methods. These methods, however, are not designed for multimodal registration, where the appearance of the learned representations differs from the one of the target images, as is the case for intraoperative registration. Although NeRFs have recently been used for 3D reconstruction of endoscopic scenes demonstrating remarkable performances [25, 28], their utilization for multimodal registration remains unexplored.

Contribution In this work, we propose a novel 3D/2D registration approach for single-view neurosurgical registration using implicit neural representations. We introduce a new formulation that separates NeRFs into structural and appearance representation, where the anatomical structure is learned preoperatively and appearance is adapted intraoperatively. This is achieved by training a hypernetwork that controls the appearance of the NeRF while leaving its learned representation of the anatomy untouched. Given a single intraoperative image, the hypernetwork crosses the modality gap and enables the NeRF to solve the 6-DoF pose estimation problem. Experiments on synthetic and real data demonstrate the effectiveness of the proposed approach, outperforming the state-of-the-art methods.

2 Methods

2.1 Problem Formulation & Overview

Given a preoperative surface mesh $\mathbf{M}$ of the craniotomy area and an intraoperative image $\mathbf{I}$ obtained from a surgical microscope, we seek to determine the pose

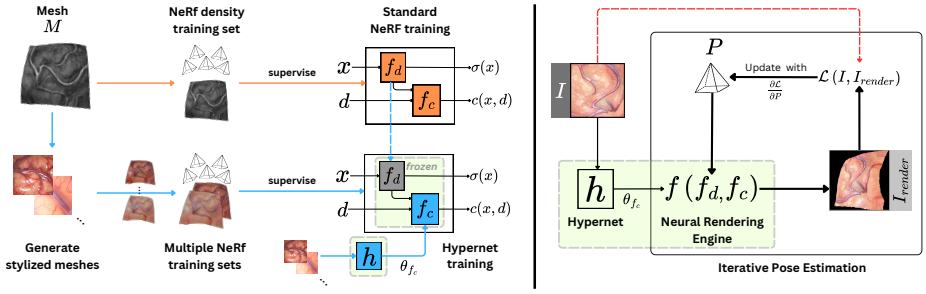


Fig. 1: Left (preoperative): We use a NeRF to first learn density/anatomy (orange) from a mesh $\mathbf{M}$ extracted from an MR scan, then learn single-shot style adaptation (blue) through a hypernetwork h while freezing the rest of the NeRF, keeping the previously learned density f_d fixed. Right (intraoperative): Iterative pose estimation on target $\mathbf{I}$. The trained NeRF and hypernet (green highlights) are used as style-conditioned neural rendering engine using ray casting, with f adapted to the appearance of the intraoperative registration target $\mathbf{I}$ through the hypernetwork h.

$\mathbf{P} \in SE(3)$. This pose minimizes a loss function $\mathcal{L}(\mathbf{P}|\mathbf{I}, \mathbf{M})$, quantifying the discrepancy between the observed intraoperative image and the preoperative mesh $\mathbf{M}$ when positioned and oriented according to the pose $\mathbf{P}$.

We approach the problem as an optimization in 2D image space. Our method minimizes the loss between $\mathbf{I}$ and images rendered from a continuous and differentiable neural representation f of the craniotomy area. The density of f is learned preoperatively from $\mathbf{M}$ and its appearance is controlled intraoperatively by $\mathbf{I}$. This approach effectively addresses the 3D-to-2D registration problem by using f to render 2D images based on the anatomy learned from $\mathbf{M}$. It also bridges the modality gap by incorporating intraoperative appearance conditioning on $\mathbf{I}$. We therefore reformulate the optimization problem with our neural representation f and an image-based loss $\mathcal{L}_{rgb}$, yielding:

$$\hat{\mathbf{P}} = \underset{\mathbf{P} \in SE(3)}{\operatorname{argmin}} \; \mathcal{L}_{\text{rgb}}(\mathbf{I}, f(\mathbf{P}|\mathbf{I}, \mathbf{M})) \tag{1}$$

Given that f is differentiable with respect to $\mathbf{P}$, we can find $\hat{\mathbf{P}}$ through iteratively rendering $f(\mathbf{P}|\mathbf{I}, \mathbf{M})$ and optimizing $\mathbf{P}$.

2.2 Bridging the Domain Gap via Hypernetwork Multi-Style NeRF

NeRF Background. A Neural Radiance Field (NeRF) [16] represents a continuous neural representation of a 3D scene. For our method, we follow Instant-NGP [18], a fast NeRF based on hash-grid encodings. The NeRF consists of a density (structure) and a color (appearance) component, evaluating a single

point and viewing direction in space. They are mathematically described by the following equations:

$$\sigma(\mathbf{x}), z(\mathbf{x}) = f_{\mathrm{d}}(\mathbf{x}; \theta_{f_d}), \tag{2}$$

where $\sigma(\mathbf{x})$ is the density and $z(\mathbf{x})$ is an intermediate representation used as input for the RGB component, and

$$c(\mathbf{x}, \mathbf{d}) = f_{\mathrm{c}}(z(\mathbf{x}), \mathbf{d}; \theta_{f_c}), \tag{3}$$

where the color c depends on $z(\mathbf{x})$ and the viewing direction $\mathbf{d}$. Both components consist of a Multilayer Perceptron (MLP), resulting in two sets of parameters, θ_{f_d} for density and θ_{f_c} for color. In practice, each component has a parametrized input encoding function, whose parameters are omitted here for clarity.

The continuous neural representation allows for rendering images by casting rays, evaluating the NeRF along each ray's path, and accumulating RGB values according to densities. For our method, NeRF serves as a neural renderer encoding our 3D mesh $\mathbf{M}$. This differs from traditional mesh representations since it is fully differentiable and has learnable disentangled components for structure f_d and appearance f_c. Both aspects are key to our method, the former for iterative pose estimation, the latter to bridge the domain gap to $\mathbf{I}$.

Training the Hypernetwork Multi-Style NeRF. First, we train a NeRF following Eq. 2 and Eq. 3 on a dataset from the preoperative MR-derived surface mesh $\mathbf{M}$. The objective is to capture the anatomical structure with high fidelity, resulting in a NeRF that faithfully replicates the brain surface from the MRI.

We introduce a hypernetwork that enables adapting to the intra-operative image appearance in real-time using only one single image. The hypernetwork, that we denote h, takes the form of a multi-head MLP and is trained to set the parameters θ_{f_c} of f_c based on the appearance of $\mathbf{I}$, leaving the structure, encoded in f_d, untouched. To train h, we use Neural Style Transfer (NST) to generate multiple training datasets by stylizing $\mathbf{M}$ using a set N of style images $\{\mathbf{J}_i\}_{i=1}^{N}$ (obtained from other surgeries). We combine two NST approaches, WCT^2 [27] for its strong preservation of semantic, and STROTSS [8] for a more photorealistic style. This changes the NeRF formulation to the following.

$$\sigma(\mathbf{x}), z(\mathbf{x}) = f_d(\mathbf{x}; \hat{\theta}_{f_d}), \tag{4}$$

$$c(\mathbf{x}, \mathbf{d}) = f_c^i(z(\mathbf{x}), \mathbf{d}; h(\mathbf{J}_i; \theta_h)), \tag{5}$$

where only h is trained to learn θ_h, while $\hat{\theta}_{f_d}$ has already been learned in the previous step and remains fixed. The whole pipeline is shown in Fig. 1. In practice, h uses a binned histogram of $\mathbf{J}_i$ instead of the whole image, since the structural information is already provided by f_d, allowing for simple low-dimensional color-based features. All training is done preoperatively, implemented in the Nerfstudio environment [24] following their Pytorch [19] implementation of Instant-NGP.

2.3 Intraoperative Registration

During surgery, assuming that camera intrinsics are known, we use f_c and f_d for a rendering engine f that renders images via ray casting. Thus we can reformulate the optimization problem in Eq. 1 as:

$$\hat{\mathbf{P}} = \underset{\mathbf{P} \in SE(3)}{\operatorname{argmin}} \, \mathcal{L}_{\text{rgb}}\big(\mathbf{I}, f(\mathbf{P}; \hat{\theta}_{f_d}, h(\mathbf{I}; \hat{\theta}_h))\big) \tag{6}$$

where $\hat{\mathbf{P}}$ represents the optimal camera pose, and $\mathcal{L}_{\text{rgb}}$ denotes the loss function comparing the intraoperative image $\mathbf{I}$ with the rendered image. $\mathbf{I}$ appears twice, as the target image, and as the input to h to condition f to approximate the appearance of $\mathbf{I}$. This is key in our method since it allows us to bridge the modality gap and express $\mathcal{L}_{\text{rgb}}$ as a conventional RGB image loss that takes the form of a relative L_2 loss [12].

The differentiable nature of f allows computing $\frac{\partial \mathcal{L}}{\partial \mathbf{P}}$, thereby enabling iterative pose refinement via gradient descent. Multiple recent works propose efficient methods to solve Eq. 6 [11][12][26]. We choose Parallel Inversion [12] that achieves state-of-art pose accuracy.

3 Experiments and Results

Datasets. We test our method on 5 clinical cases, each with its preoperative T1 MRI scan and corresponding surgical microscope image, except for case 5 with only a T1 MRI scan. To stylize meshes, we rely on a small dataset of $N = 15$ surgical microscope images from different cases, including the 4 corresponding to our clinical test cases that have a surgical microscope image. We build one dataset per case. To train the NeRF, we generate 100 images per style with their respective camera poses. For each case, this results in 1500 images that show the stylized brain surface mesh in 15 styles with 100 images and poses for each style. Additionally for case 5, for each of the 4 styles that correspond to the other clinical cases, we generate 50 random images and poses as test targets for registration.

View Synthesis. To demonstrate that our hypernetwork produces plausible appearances, we synthesize 3 views from different poses on one of the clinical cases, as illustrated in Fig. 2. These syntheses are obtained by training on the dataset for this case while omitting its style. Our method shows qualitatively photorealistic results that respect the anatomy of the case while being similar in appearance.

Pose Estimation on Synthetic Targets. To evaluate our method on a larger number of registration targets, we use case 5 and generate 50 targets for 4 different styles (corresponding to the clinical test cases), yielding 200 targets in total. The hypernetwork is trained on the remaining 11 styles. For each style,

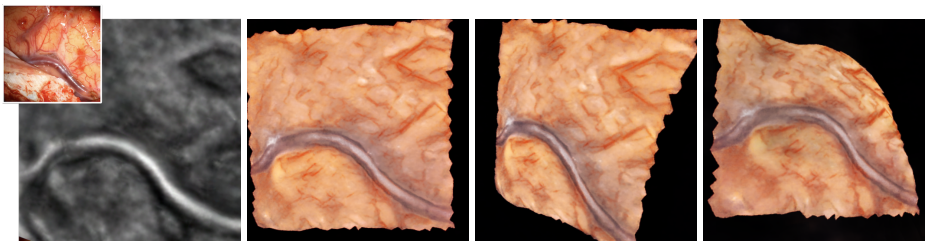
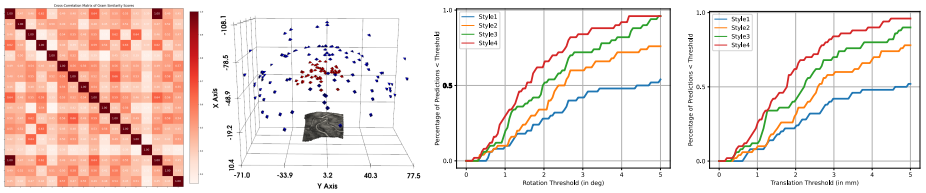


Fig. 2: An example of synthesis from 3 different poses on one of the clinical cases. First: image obtained from the MRI with surgical microscope image $\mathbf{I}$. Remaining images: synthesis with f, style inferred by hypernetwork h on $\mathbf{I}$.

the hypernetwork conducts a singular inference to determine the style-specific appearance, which is then utilized in the pose estimation for all 50 targets within that style. The results are shown as accuracy-threshold curves (Fig. 3), which indicate the proportion of predicted poses that fall within a given error threshold.



Style variance Pose sampling Rotations (deg) Translations (mm)

Fig. 3: Evaluation on synthetic targets (from left to right): cross-correlation matrix of Gram-similarity score of all styles showing pairwise style similarity and dissimilarity; pose distribution (blue: training set, red: test set); and accuracy-threshold curves for rotation and translation.

Fig. 3 shows that rotation and translation errors vary depending on the style of the targets. While 96% of poses estimated for Style 3 and Style 4 targets have a rotation error below 5°, it is 76% for targets of Style 2. Similarly, 96% of poses estimated for Style 4 and 90% of poses estimated for Style 3 have a translation error below 5 mm, whereas the pose estimations on targets of Style 2 reach 78%. Style 1 on the other hand has 52% of poses with translation error below 5 mm and 54% of poses with rotation error below 5°. This can be explained by the small number of style images in the hypernetwork dataset ($N = 11$).

Table 1: Comparative Registration Error

Metrics / Methods	Style1	Style2	Style3	Style4	**Ours**	MR-NeRF	Reg.	Reg. ID
ATE (mm)	4.17	3.24	2.56	1.96	**3.12**	10.12	13.31	6.29
ART (deg)	4.53	3.25	2.41	1.84	**3.01**	11.30	8.78	5.70
Outliers (%)	22	6	0	0	7	46.50	58	4.52

Comparison with Baseline and State-of-the-art. We evaluate our method **Ours** against a baseline and the state-of-the-art method. The baseline is our NeRF trained exclusively on images and poses from the MRI visualized as volume rendering, referred to as **MR-NeRF**. The state-of-the-art method is a style-invariant regressor [3]. We used case 5 data with 11 styles of 100 images and poses each, and evaluated on a test set of 4 styles of 50 images and ground-truth poses each. We trained and tested two versions of the regressor, **Reg.** with the described data also used for **MR-NeRF** and **Ours**, and **Reg. ID**, with a modified training and test split with poses in-distribution, following the method guidelines. The pose distribution is visualized in Fig 3, where red poses correspond to the test set and blue ones to the training set. The style distribution is also shown in Fig 3 with the cross-correlation matrix of Gram-similarity scores, a common similarity measure used in Neural Style Transfer.

For all methods, we report the Average Translation Error (ATE) and Average Rotation Error (ART) in Tab. 1. We also define outliers as poses with a rotation error larger than 20°. Outliers are excluded from ATE and ART and reported separately as a percentage of the total number of test poses. For **Ours**, we additionally report results on all styles individually.

Our method outperforms both baseline and state-of-the-art and achieves ATE and ART of 3.12mm and 3.01° that meet clinical needs [1]. We also achieve style-invariance for Style 2, Style 3, and Style 4. However, we do not achieve style-invariance for Style 1 with a high number of outliers, in line with the Accuracy-threshold curve experiment. Given that the underlying anatomy and target poses are the same across all styles, this indicates that the appearance approximated by our hypernetwork is not sufficiently similar enough to Style 1 to allow for a robust pose estimation with gradient descent. The **MR-NeRF** is not able to cross the modality gap to the target images while **Reg.** does not generalize to out-of-distribution target poses. Only when withholding and evaluating part of the more homogeneous training set does the performance of the regressor improve with **Reg. ID** as shown in Tab. 1.

Tests on Clinical Cases. We evaluate each case separately and exclude its style from the training set. The cases represent different craniotomy openings with varied appearances and anatomies. We provide qualitative results with a visual assessment of each case in Fig. 4. Except for the case in row 3 where we observed a very early local minimum in the optimization, all other pose estimations reach a visually correct registration.

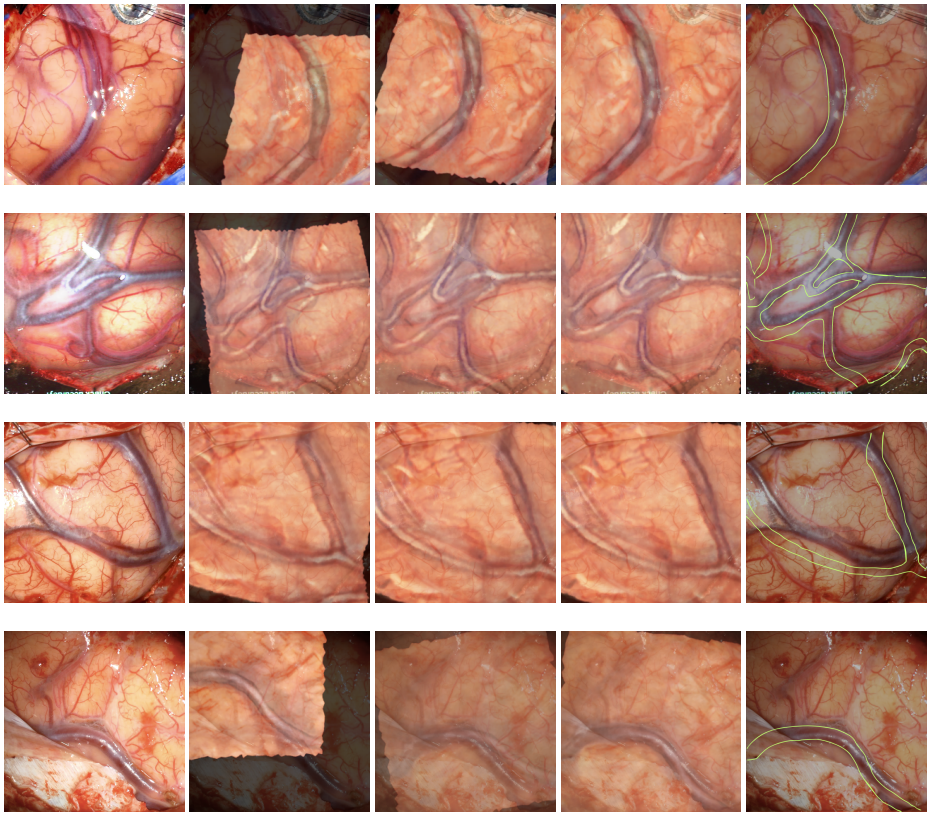


Fig. 4: Tests on real cases, one case per row. Left column: target images from the surgical microscope. Middle columns: 3 optimization steps: early-optimization, mid-optimization, and final pose estimation. Right column: Intraoperative image with vessel-overlay of our estimated pose.

4 Conclusion

In this paper, we presented a novel 3D/2D intraoperative registration approach for neurosurgery by introducing a cross-modal inverse neural rendering that disentangles NeRF representation into structure and appearance adapting thereby NeRFs to the preoperative and intraoperative settings. We presented experiments with qualitative and quantitative results on synthetic and retrospective real patient data, showing that our method outperforms the state-of-the-art, performs well in real conditions, and meets clinical needs. Future work will extend this representation to handle deformation. This can be achieved by modifying the density component of our implicit neural representation to be robust to non-rigid transformations.

References

1. Frisken, S., Luo, M., Juvekar, P., Bunevicius, A., Machado, I., Unadkat, P., Bertotti, M., Toews, M., Wells, W., Miga, M., Golby, A.: A comparison of thin-plate spline deformation and finite element modeling to compensate for brain shift during tumor resection. International Journal of Computer Assisted Radiology and Surgery **15** (08 2019)
2. Gonzalez-Darder, J.M.: State of the Art of the Craniotomy in the Early Twenty-First Century and Future Development, pp. 421–427. Springer International Publishing, Cham (2019)
3. Haouchine, N., Dorent, R., Juvekar, P., Torio, E., Wells III, W.M., Kapur, T., Golby, A.J., Frisken, S.: Learning expected appearances for intraoperative registration during neurosurgery. In: International Conference on Medical Image Computing and Computer-Assisted Intervention. pp. 227–237. Springer (2023)
4. Haouchine, N., Juvekar, P., Nercessian, M., Wells III, W.M., Golby, A., Frisken, S.: Pose estimation and non-rigid registration for augmented reality during neurosurgery. IEEE Transactions on Biomedical Engineering **69**(4), 1310–1317 (2022)
5. Ji, S., Fan, X., Roberts, D.W., Hartov, A., Paulsen, K.D.: Cortical surface shift estimation using stereovision and optical flow motion tracking via projection image registration. Medical Image Analysis **18**(7), 1169 – 1183 (2014)
6. Jiang, J., Nakajima, Y., Sohma, Y., Saito, T., Kin, T., Oyama, H., Saito, N.: Marker-less tracking of brain surface deformations by non-rigid registration integrating surface and vessel/sulci features. International journal of computer assisted radiology and surgery **11** (03 2016)
7. Kato, H., Ushiku, Y., Harada, T.: Neural 3d mesh renderer. In: Proceedings of the IEEE conference on computer vision and pattern recognition. pp. 3907–3916 (2018)
8. Kolkin, N., Salavon, J., Shakhnarovich, G.: Style transfer by relaxed optimal transport and self-similarity. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. pp. 10051–10060 (2019)
9. Koo, B., Robu, M.R., Allam, M., Pfeiffer, M., Thompson, S., Gurusamy, K., Davidson, B., Speidel, S., Hawkes, D., Stoyanov, D., et al.: Automatic, global registration in laparoscopic liver surgery. International Journal of Computer Assisted Radiology and Surgery pp. 1–10 (2022)
10. Kuhnt, D., Bauer, M.H.A., Nimsky, C.: Brain shift compensation and neurosurgical image fusion using intraoperative mri: Current status and future challenges. Critical Reviews and trade in Biomedical Engineering **40**(3), 175–185 (2012)
11. Li, Z., Fu, K., Wang, H., Wang, M.: Pi-nerf: A partial-invertible neural radiance fields for pose estimation. In: Proceedings of the 31st ACM International Conference on Multimedia. pp. 7826–7836 (2023)
12. Lin, Y., Müller, T., Tremblay, J., Wen, B., Tyree, S., Evans, A., Vela, P.A., Birchfield, S.: Parallel inversion of neural radiance fields for robust pose estimation. In: 2023 IEEE International Conference on Robotics and Automation (ICRA). pp. 9377–9384. IEEE (2023)
13. Luo, M., Larson, P.S., Martin, A.J., Konrad, P.E., Miga, M.I.: An integrated multiphysics finite element modeling framework for deep brain stimulation: Preliminary study on impact of brain shift on neuronal pathways. In: MICCAI 2019. pp. 682–690. Springer International Publishing (2019)
14. Marcus, H.J., Pratt, P., Hughes-Hallett, A., Cundy, T.P., Marcus, A.P., Yang, G.Z., Darzi, A., Nandi, D.: Comparative effectiveness and safety of image guidance

systems in neurosurgery: a preclinical randomized study. Journal of neurosurgery **123**(2), 307–313 (2015)
15. Marreiros, F.M.M., Rossitti, S., Wang, C., Smedby, Ö.: Non-rigid deformation pipeline for compensation of superficial brain shift. In: MICCAI 2013. pp. 141–148. Springer Berlin Heidelberg, Berlin, Heidelberg (2013)
16. Mildenhall, B., Srinivasan, P.P., Tancik, M., Barron, J.T., Ramamoorthi, R., Ng, R.: Nerf: Representing scenes as neural radiance fields for view synthesis. Communications of the ACM **65**(1), 99–106 (2021)
17. Mohammadi, A., Ahmadian, A., Azar, A.D., Sheykh, A.D., Amiri, F., Alirezaie, J.: Estimation of intraoperative brain shift by combination of stereovision and doppler ultrasound: phantom and animal model study. International Journal of Computer Assisted Radiology and Surgery **10**(11), 1753–1764 (Nov 2015)
18. Müller, T., Evans, A., Schied, C., Keller, A.: Instant neural graphics primitives with a multiresolution hash encoding. ACM Transactions on Graphics (ToG) **41**(4), 1–15 (2022)
19. Paszke, A., Gross, S., Chintala, S., Chanan, G., Yang, E., DeVito, Z., Lin, Z., Desmaison, A., Antiga, L., Lerer, A.: Automatic differentiation in pytorch (2017)
20. Pereira, V.M., Smit-Ockeloen, I., Brina, O., Babic, D., Breeuwer, M., Schaller, K., Lovblad, K.O., Ruijters, D.: Volumetric Measurements of Brain Shift Using Intraoperative Cone-Beam Computed Tomography: Preliminary Study. Operative Neurosurgery **12**(1), 4–13 (08 2015)
21. Sanai, N., Polley, M.Y., McDermott, M.W., Parsa, A.T., Berger, M.S.: An extent of resection threshold for newly diagnosed glioblastomas: Clinical article. Journal of Neurosurgery JNS **115**(1), 3–8 (2011)
22. Stoyanov, D.: Surgical vision. Ann Biomed Eng **40**, 332–345 (2012)
23. Sun, K., Pheiffer, T., Simpson, A., Weis, J., Thompson, R., Miga, M.: Near real-time computer assisted surgery for brain shift correction using biomechanical models. IEEE Translational Engineering in Health and Medicine **2**, 1–13 (2014)
24. Tancik, M., Weber, E., Ng, E., Li, R., Yi, B., Wang, T., Kristoffersen, A., Austin, J., Salahi, K., Ahuja, A., et al.: Nerfstudio: A modular framework for neural radiance field development. In: ACM SIGGRAPH 2023 Conference Proceedings. pp. 1–12 (2023)
25. Wang, Y., Long, Y., Fan, S.H., Dou, Q.: Neural rendering for stereo 3d reconstruction of deformable tissues in robotic surgery. In: International Conference on Medical Image Computing and Computer-Assisted Intervention. pp. 431–441. Springer (2022)
26. Yen-Chen, L., Florence, P., Barron, J.T., Rodriguez, A., Isola, P., Lin, T.Y.: inerf: Inverting neural radiance fields for pose estimation. In: 2021 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS). pp. 1323–1330. IEEE (2021)
27. Yoo, J., Uh, Y., Chun, S., Kang, B., Ha, J.W.: Photorealistic style transfer via wavelet transforms. In: Proceedings of the IEEE/CVF International Conference on Computer Vision. pp. 9036–9045 (2019)
28. Zha, R., Cheng, X., Li, H., Harandi, M., Ge, Z.: Endosurf: Neural surface reconstruction of deformable tissues with stereo endoscope videos. In: International Conference on Medical Image Computing and Computer-Assisted Intervention. pp. 13–23. Springer (2023)