SRPAN: Implicit Neural Slice-to-Volume Reconstruction for Super-Resolution Pancreatic MRI

Fredrik Olofsson[1(✉)⋆], Koen J. H. Wijsman[2⋆] Megan M. L. Engels[2], Bas Boekestijn[1], Shirin Shahbazi Feshtali[1], Monique E. van Leerdam[2,3], Alexander Broersen[1], Jeanin E. van Hooft[2], and Jouke Dijkstra[1]

[1] Department of Radiology, Leiden University Medical Center, Leiden, The Netherlands
[2] Department of Gastroenterology and Hepatology, Leiden University Medical Center, Leiden, The Netherlands
[3] Department of Gastrointestinal Oncology, Netherlands Cancer Institute, Amsterdam, The Netherlands
{k.f.s.olofsson, k.j.h.wijsman, j.e.van_hooft, j.dijkstra}@lumc.nl

Abstract. Clinical pancreatic MRI is central to surveillance for pancreatic ductal adenocarcinoma, yet most clinical scan protocols provide only two approximately orthogonal thick-slice stacks. This yields anisotropic volumes with limited through-plane resolution, constraining 3D assessment and the visibility of small lesions. Existing slice-to-volume reconstruction (SVR) methods, largely developed for fetal brain MRI, assume multi-stack redundancy and use slice-level outlier rejection, assumptions that do not hold in the low-redundancy two-stack setting where each slice contributes essential information. To address localized cross-view inconsistencies without discarding slices, we propose **SRPAN**, which introduces pixel-wise reliability weighting via cross-view gating. SRPAN is an implicit neural representation framework that jointly optimizes per-slice rigid registration and a point spread function (PSF)-aware continuous reconstruction under heteroscedastic uncertainty. In a clinical cohort of 40 high-risk individuals undergoing pancreatic MRI surveillance, SRPAN achieves the highest overall reprojection fidelity (PSNR 25.66, SSIM 0.81, NCC 0.96), leads in 8/9 metrics, and produces the most visually consistent 1 mm isotropic reconstructions across planes. Qualitative results show that SRPAN suppresses artifacts in regions of cross-view disagreement where symmetric L_1 objectives can imprint corrupted structure into the final volume. Code is available at: https://github.com/OlofssonFredrik/SRPAN_MICCAI

Keywords: Slice-to-Volume Reconstruction · Implicit Neural Representation · Pancreatic Imaging · MRI

⋆ Equal contribution.

1 Introduction

Pancreatic ductal adenocarcinoma (PDAC) is among the deadliest malignancies, with a 5-year survival rate of 13% [1]. This poor prognosis is largely attributable to late-stage diagnosis, as early-stage PDAC typically presents with vague and nonspecific symptoms [2]. Although survival outcomes are substantially better when PDAC is detected at an early stage, identifying small lesions remains challenging, even with magnetic resonance imaging (MRI), such as in pancreatic surveillance programs for high-risk individuals (HRIs) [3]. Moreover, when lesions are detected, it is often difficult to distinguish malignant lesions from benign. A standard pancreatic MRI examination comprises multiple image series with high in-plane but limited out-of-plane resolution. For example, the largest publicly available pancreatic MRI dataset released with PanSegNet exhibits a 4–8× lower out-of-plane resolution due to the large inter-slice spacing [4]. Information is combined from complementary series acquired from different views into a single high-resolution three-dimensional (3D) volume. This could, once thoroughly validated, enhance lesion detection and differentiation by radiologists. Additionally, isotropic high-resolution 3D volumes may improve the robustness and accuracy of downstream artificial intelligence applications, including pancreas and tumor segmentation, classification, and risk prediction, ultimately enabling AI-driven early diagnosis of PDAC.

Reconstructing an isotropic high-resolution 3D volume from multi-view thick-slice acquisitions requires sophisticated slice-to-volume reconstruction (SVR). To our knowledge, no pancreas-specific SVR method exists for clinically acquired multi-view pancreatic MRI. We therefore draw inspiration from the closest and most mature related literature, fetal brain MRI, where motion-corrupted thick-slice stacks are reconstructed into isotropic volumes using optimization-based pipelines such as SVRTK [5, 6]. More recently, SVR has been advanced with implicit neural representations (INRs) that jointly estimate motion and reconstruction, e.g., NeSVoR [7] and SSVR [8]. Importantly, both optimization-based and INR-based SVR methods are commonly developed and evaluated in settings with multiple partially redundant stacks, where robustness is often achieved through slice-level outlier handling such as rejection or robust averaging. In contrast, clinically acquired pancreatic MRI commonly provides only two orthogonal stacks, yielding a low-redundancy regime in which every slice contributes essential information and slice rejection becomes untenable. SRPAN builds on INR-based SVR but is tailored to this two-stack pancreatic setting: rather than relying on stack redundancy and slice-level outlier removal, it not only introduces cross-view gated pixel-wise reliability to handle local inconsistencies, but also couples this with joint motion estimation and PSF-aware continuous reconstruction under heteroscedastic uncertainty, preserving all slices for accurate reconstruction.

Contributions.

- We introduce SRPAN, an INR-based SVR reconstruction framework for two clinically acquired orthogonal pancreatic MRIs in a low-redundancy setting.

– We develop a PSF-aware implicit reconstruction model with heteroscedastic uncertainty and pixel-wise reliability weighting to handle local cross-view inconsistencies without slice-level outlier rejection.
– We evaluate SRPAN on a 40-subject clinical cohort, benchmarking against representative baselines and performing ablations on key design choices.

2 Method

We present SRPAN, an end-to-end SVR method for clinically acquired pancreatic MRI. Given two orthogonal anisotropic T2-weighted stacks (axial and coronal), SRPAN reconstructs a single isotropic 3D volume as a continuous implicit field that can be queried at arbitrary coordinates (Fig. 1). We reconstruct a 1 mm isotropic volume as a clean evaluation baseline: this provides 3× super-resolution along the low-resolution axis while remaining close to the native in-plane spacing of the axial (0.66 mm) and coronal (0.88 mm) stacks. SRPAN jointly optimizes slice-wise rigid registration with intensity scaling and a PSF-aware implicit reconstruction with heteroscedastic uncertainty.

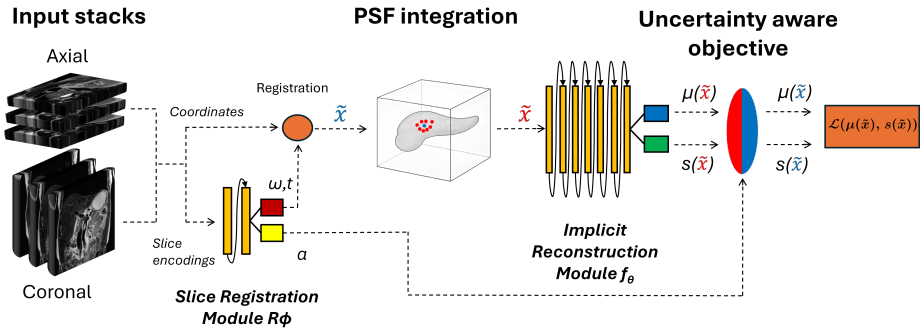

Fig. 1. SRPAN overview. Axial and coronal stacks are jointly aligned via a slice-wise rigid registration module with per-slice intensity scaling. An implicit 3D field is optimized using PSF-based slice formation. A heteroscedastic objective together with cross-view gating down-weights locally inconsistent pixels.

2.1 Slice Registration Module

Because our scanning protocol employs respiratory-triggered and breath-hold acquisitions, pronounced non-rigid deformations within individual slices are not observed. However, consistent slice-wise shifts occur across breathing cycles. To ensure robustness in this first SVR method tailored to the pancreas, we adopt a simplified approach and model motion with per-slice rigid transforms, which proved adequate for our data after visual inspection. The registration module

predicts a rigid transform and an intensity rescaling for each 2D slice. For slice i, we use a stack identifier $s_i \in \{\text{axial}, \text{coronal}\}$ and a normalized slice coordinate $z_i \in [-1, 1]$. A small network R_ϕ maps (s_i, z_i) to motion and an intensity-scale logit:

$$(w_i, t_i, a_i^{\text{scale}}) = R_\phi(s_i, z_i), \tag{1}$$

where $w_i, t_i \in \mathbb{R}^3$ are rotation and translation vectors. Rotations are mapped to $SO(3)$ via the exponential map. The scale logits are mapped to positive slice scales C_i using a softmax followed by unit-mean renormalization. We implement R_ϕ as a 2-layer SIREN [9] (width 64) with two linear heads.

2.2 Implicit Reconstruction Module

We represent a continuous 3D volume shared across both stacks and simulate slice formation via an acquisition point spread function (PSF). We map voxels $x \in \mathbb{R}^3$ to normalized world coordinates $\tilde{x} \in [-1, 1]^3$ via an affine transform. A 7-layer SIREN [9] (width 330) parameterizes the implicit field with mean and log-variance heads yielding a heteroscedastic Gaussian regression model [10, 11]:

$$f_\theta(\tilde{x}) = \big(\mu(\tilde{x}),\, s(\tilde{x})\big), \qquad s(\tilde{x}) = \log \sigma^2(\tilde{x}), \tag{2}$$

We follow the standard SVR acquisition model, where each observed slice is generated from the underlying volume through slice motion and an acquisition PSF [5, 7, 8, 12, 13]. For slice i and a queried pixel center x in world coordinates (after applying the estimated rigid transform), we approximate the PSF integral around x via Monte Carlo sampling: draw $u_{ik} \sim \mathcal{N}(0, \Sigma_i)$, set $q_{ik}(x) = x + u_{ik}$, and evaluate the field at normalized locations. This yields the predicted mean and variance per pixel including the slice-wise intensity scale C_i as

$$\mu_i(x) = C_i \, \frac{1}{K} \sum_{k=1}^{K} \mu\big(\tilde{q}_{ik}(x)\big), \tag{3}$$

$$s_i(x) \;=\; \log \sigma_i^2(x) = \log\left(C_i^2 \, \frac{1}{K} \sum_{k=1}^{K} \exp\big(s(\tilde{q}_{ik}(x))\big) \right). \tag{4}$$

We model the PSF as an anisotropic Gaussian as in [5, 7, 12], with

$$\Sigma_i = \operatorname{diag}\left(\left(\frac{1.2\, r_x}{2.355}\right)^2,\; \left(\frac{1.2\, r_y}{2.355}\right)^2,\; \left(\frac{r_z}{2.355}\right)^2 \right), \tag{5}$$

for in-plane spacings (r_x, r_y) and slice thickness r_z. During training, we clamp log-variance predictions to $s_{\max} = \log(0.35^2)$; at inference, we use only the mean.

2.3 Loss function

Balancing the reconstruction across both stacks is challenging because the acquisitions differ in resolution and artifact profiles. In our protocol, the axial stack

provides a stable depiction of fine structures, whereas the coronal stack occasionally exhibits localized intensity artifacts. If optimized symmetrically, such inconsistencies propagate into the volume. We therefore treat the axial stack as a reference signal and downweight disagreeing coronal pixels. While dynamically selecting the higher-quality stack could benefit edge cases, we fixate the axial stack to maintain a fully automatic pipeline. This is grounded in our dataset: both scans are highly consistent in the coronal view, but the coronal scan frequently exhibits out-of-plane degradations in the axial view, making the axial acquisition the reliable reference.

For each coronal pixel x, we sample the axial input stack at the same location to obtain $Y_{\mathrm{ax}\to\mathrm{co}}(x)$ via trilinear interpolation. We define the cross-view discrepancy $\Delta(x) = \left|Y_{\mathrm{ax}\to\mathrm{co}}(x) - I_i(x)\right|$ and estimate a robust scale τ from the upper tail of $\Delta(x)$. We then compute the cross-view gating weight

$$w(x) = W_{\min}(t) + \frac{1 - W_{\min}(t)}{1 + \left(\dfrac{\Delta(x)}{\tau + \varepsilon}\right)^{P}}, \tag{6}$$

where $P > 1$ controls the transition sharpness and $W_{\min}(t)$ is annealed during training. We use $P = 3$, $\varepsilon = 10^{-6}$, $\tau = \max(Q_{0.90}(\Delta), 0.05)$, and linearly anneal $W_{\min}(t)$ from 0.15 to 0. Given an observed intensity $I_i(x)$ and prediction $\big(\mu_i(x), s_i(x)\big)$, we use the heteroscedastic Gaussian negative log-likelihood (HGNLL) [10, 11]:

$$\mathcal{L}_i(x) = \frac{1}{2}\left[\exp\big(-s_i(x)\big)\big(I_i(x) - \mu_i(x)\big)^2 + s_i(x)\right]. \tag{7}$$

The total loss is the average of $\mathcal{L}_i(x)$ over axial pixels and a gated average over coronal pixels using $w(x)$.

3 Experimental Setup

3.1 Dataset

We evaluated SRPAN on an internal dataset of 40 anonymized MRI examinations, acquired between 2023 and 2025 to reflect current clinical practice as closely as possible. The examinations were obtained from 40 unique subjects randomly selected from a larger clinical cohort of high-risk individuals (HRIs) with a family-inherited CDKN2A genetic mutation. These HRIs have an estimated lifetime risk of approximately 20% of developing PDAC and therefore undergo annual MRI surveillance [3]. Each MRI examination consisted of a standardized abdominal MRI with a pancreatic protocol. For reconstruction, SRPAN used the axial T2-weighted sequence (voxel spacing $0.66 \times 0.66 \times 3.00$ mm, respiratory-triggered) and the coronal T2-weighted sequence (voxel spacing $0.88 \times 3.00 \times 0.88$ mm, breath-hold) as input.

3.2 Training protocol

We train one model per HRI. We use Adam for the INR and registration modules (lr 5×10^{-5} and 5×10^{-4}) with cosine annealing (min lr 2.5×10^{-5}). Each iteration samples a batch of 6000 pixels (axial/coronal 50/50). The total amount of iterations is $\text{iters}_{\max} = \alpha_{\text{iter}} \frac{N_{\text{voxels}}}{\text{batch size}}$ with $\alpha_{\text{iter}} = 500$ (SRPAN) and $\alpha_{\text{iter}} = 125$ (SRPAN-125, used for ablations). The number of Monte Carlo samples K used to approximate the PSF integral is increased quadratically from 1 to $K_{\max} = 64$.

3.3 Evaluation

We compare the results of SRPAN with three SVR baselines: SVRTK, NeSVoR, and SSVR. We use the public implementations of SVRTK [6] and NeSVoR [7]. Since no official SSVR code is available, we re-implement SSVR from [8] and match the reported formulation; hyperparameters are kept aligned with SRPAN where possible.

As no ground-truth isotropic volume is available, we evaluate data fidelity by re-projecting each reconstruction to the acquired slice locations and computing the peak signal-to-noise ratio (PSNR), structural similarity (SSIM), and normalized cross-correlation (NCC) in the native axial and coronal planes. SRPAN and SSVR are queried directly at the slice coordinates, NeSVoR is sampled using `sample-slices` (with `-stack` registration), and SVRTK is resliced via trilinear interpolation. We report metrics separately for the axial and coronal views, as well as *overall* metrics (axial and coronal combined). For the overall metrics, the slice-level metrics for all axial and coronal slices were first averaged within each HRI. Subsequently, all metrics were averaged across HRIs with every HRI contributing equally, independent of the number of slices per HRI.

4 Results

To our knowledge, this is the first evaluation of implicit slice-to-volume reconstruction (SVR) on clinically acquired pancreatic MRI. Across 40 HRIs (Table 1), SRPAN achieves the best overall reprojection fidelity (PSNR 25.66, SSIM 0.81, NCC 0.96) and leads in 8/9 reported metrics. Although SVRTK attains higher axial SSIM (0.90 vs. 0.83), this reflects a clear bias towards the axial input rather than coherent multi-view fusion, consistent with its low coronal SSIM (0.30).

Figure 2 qualitatively compares the 1 mm isotropic reconstructions across methods and against the two input stacks. Because the acquired axial and coronal stacks are not perfectly orthogonal and have slightly different coverage, SVRTK and NeSVoR show a reduced effective field-of-view (FOV) and residual cross-plane inconsistencies. In contrast, SRPAN better preserves the overlap region and yields the most coherent anatomy across axial and coronal views, while apparently providing the sharpest sagittal reformat. Moreover, SRPAN most effectively suppresses the checkerboard acquisition artifacts visible in the low-resolution inputs. In this example, SVRTK and NeSVoR prioritize fidelity to the

axial stack, whereas SSVR favors the coronal stack; SRPAN instead produces the most consistent 3D anatomy across all planes.

The ablation study (Table 2) shows that performance is driven by the combination of implicit representation and slice modeling: SIREN activations and physics-based terms (PSF, registration, slice-wise intensity scaling) are important, and removing variance stabilization causes failure. While replacing the heteroscedastic objective with L_1 yields similar average metrics, Fig. 3 shows that uncertainty modeling is critical under local cross-view disagreement, preventing corrupted regions in one view from imprinting artifacts into the 3D reconstruction.

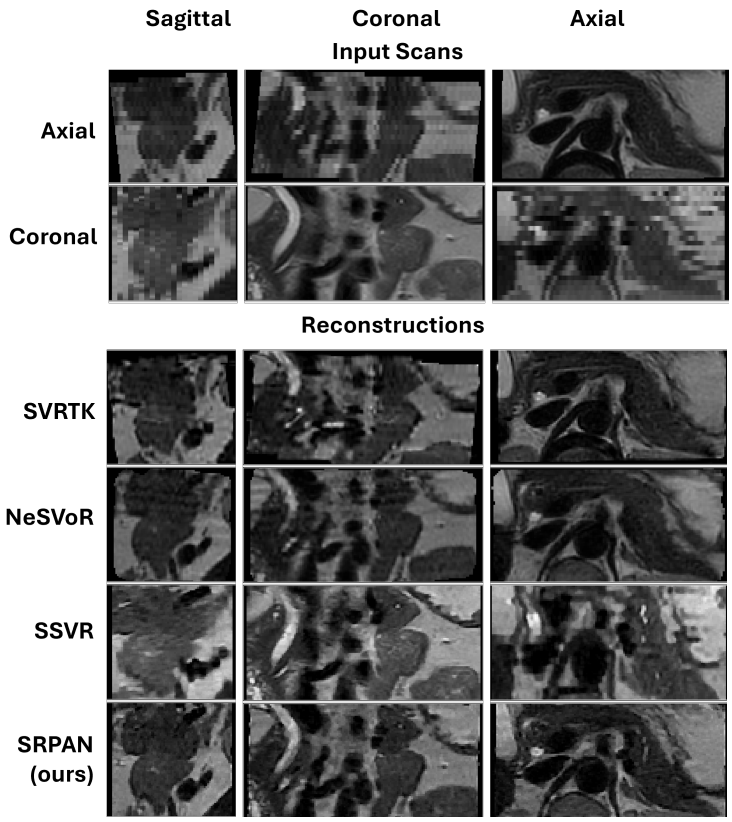

Fig. 2. Isotropic $1.0 \times 1.0 \times 1.0$ mm reconstructions across methods. Top: the axial and coronal input stacks shown in sagittal, coronal, and axial planes. Bottom: reconstructions produced by SVRTK, NeSVoR, SSVR, and SRPAN. The sagittal view is super-resolved by $\times 3$ for all methods. Images are shown on the native pixel grid without any smoothing or interpolation.

Table 1. Quantitative performance, reported as the mean over 40 HRIs, with standard deviation in parentheses. Higher is better (↑). Best score per metric is shown in **bold**.

Method	PSNR ↑	SSIM ↑	NCC ↑
Axial slices			
SVRTK	24.63 (2.17)	**0.90** (0.03)	0.94 (0.03)
NeSVoR	20.15 (2.03)	0.71 (0.10)	0.82 (0.06)
SSVR	21.37 (4.77)	0.56 (0.20)	0.89 (0.23)
SRPAN (ours)	**26.78** (1.73)	0.83 (0.05)	**0.97** (0.02)
Coronal slices			
SVRTK	15.03 (1.20)	0.30 (0.08)	0.59 (0.10)
NeSVoR	15.25 (1.25)	0.34 (0.08)	0.62 (0.09)
SSVR	17.36 (5.33)	0.40 (0.26)	0.66 (0.18)
SRPAN (ours)	**23.87** (1.90)	**0.78** (0.05)	**0.93** (0.03)
Overall slices			
SVRTK	20.84 (1.83)	0.66 (0.06)	0.80 (0.05)
NeSVoR	18.22 (1.60)	0.56 (0.09)	0.74 (0.07)
SSVR	20.15 (1.45)	0.52 (0.07)	0.81 (0.07)
SRPAN (ours)	**25.66** (1.60)	**0.81** (0.04)	**0.96** (0.02)

Table 2. Ablation study on a 20-HRI subset using SRPAN-125, reported as the mean over HRIs, with standard deviation in parentheses. Overall scores are reported.

Setting	PSNR ↑	SSIM ↑	NCC ↑
Baseline (SRPAN-125)	24.95 (1.66)	0.79 (0.04)	0.95 (0.02)
SIREN → ReLU	17.15 (1.07)	0.41 (0.05)	0.74 (0.06)
PSF: None	21.36 (1.14)	0.61 (0.04)	0.90 (0.02)
Registration: None	20.07 (3.16)	0.59 (0.19)	0.80 (0.16)
Slice scaling: None	23.64 (1.60)	0.76 (0.05)	0.94 (0.02)
HGNLL → L1	24.91 (1.85)	0.80 (0.05)	0.95 (0.02)
Gating: None	24.94 (3.13)	0.77 (0.15)	0.93 (0.12)
Variance clamp: None	17.03 (4.20)	0.38 (0.26)	0.64 (0.20)

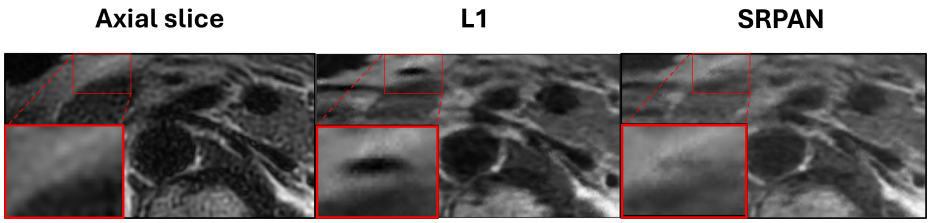

Fig. 3. Comparison between the L1-trained variant (middle) and SRPAN with heteroscedastic uncertainty (right) against the input slice (left). When input views disagree locally, the L1-trained model can introduce implausible structures, whereas SRPAN suppresses these artifacts by down-weighting inconsistent pixels via the proposed objective.

5 Conclusion

We presented SRPAN, an implicit slice-to-volume reconstruction framework tailored to two-stack pancreatic MRI, a low-redundancy regime where slice-level rejection is not viable. SRPAN jointly optimizes per-slice rigid registration and a PSF-aware implicit reconstruction with heteroscedastic uncertainty, introducing pixel-wise cross-view gating to attenuate localized inconsistencies. On a cohort of 40 high-risk individuals undergoing pancreatic surveillance, SRPAN achieves the highest overall reprojection fidelity and yields the most consistent 1 mm isotropic reconstructions. More broadly, this reliability-weighted approach could extend to other sparse multi-view clinical acquisitions. Current limitations include the lack of an isotropic ground truth and reliance on an asymmetric gating strategy assuming a high-quality axial reference. Future work will explore non-rigid deformation modeling and evaluate the impact of these isotropic reconstructions on downstream clinical AI tasks, such as lesion detection and segmentation.

Acknowledgments. This research was supported by a grant of Stichting Hanarth Fonds, The Netherlands (call 2023 #47) and it was performed using the ALICE compute resources provided by Leiden University.

Disclosure of Interests. The authors have no competing interests to declare that are relevant to the content of this article.

References

1. Siegel, R. L., Giaquinto, A. N. & Jemal, A. Cancer statistics, 2024. *CA: a cancer journal for clinicians* **74** (2024).
2. Mannucci, A. & Goel, A. Advances in pancreatic cancer early diagnosis, prevention, and treatment: The past, the present, and the future. *CA: A Cancer Journal for Clinicians* (2025).
3. Klatte, D. C., Wallace, M. B., Löhr, M., Bruno, M. J. & van Leerdam, M. E. Hereditary pancreatic cancer. *Best Practice & Research Clinical Gastroenterology* **58,** 101783 (2022).
4. Zhang, Z. *et al.* Large-scale multi-center CT and MRI segmentation of pancreas with deep learning. *Medical image analysis* **99,** 103382 (2025).
5. Kuklisova-Murgasova, M., Quaghebeur, G., Rutherford, M. A., Hajnal, J. V. & Schnabel, J. A. Reconstruction of fetal brain MRI with intensity matching and complete outlier removal. *Medical image analysis* **16,** 1550–1564 (2012).
6. Uus, A. *et al.* Deformable slice-to-volume registration for motion correction of fetal body and placenta MRI. *IEEE transactions on medical imaging* **39,** 2750–2759 (2020).
7. Xu, J. *et al.* NeSVoR: implicit neural representation for slice-to-volume reconstruction in MRI. *IEEE transactions on medical imaging* **42,** 1707–1719 (2023).

8. Dannecker, M. *et al.* Meta-learning Slice-to-Volume Reconstruction in Fetal Brain MRI using Implicit Neural Representations. *arXiv preprint arXiv: 2505.09565* (2025).
9. Sitzmann, V., Martel, J., Bergman, A., Lindell, D. & Wetzstein, G. Implicit neural representations with periodic activation functions. *Advances in neural information processing systems* **33,** 7462–7473 (2020).
10. Nix, D. A. & Weigend, A. S. *Estimating the mean and variance of the target probability distribution* in *Proceedings of 1994 ieee international conference on neural networks (ICNN'94)* **1** (1994), 55–60.
11. Kendall, A. & Gal, Y. What uncertainties do we need in bayesian deep learning for computer vision? *Advances in neural information processing systems* **30** (2017).
12. Jiang, S. *et al.* MRI of moving subjects using multislice snapshot images with volume reconstruction (SVR): application to fetal, neonatal, and adult brain studies. *IEEE transactions on medical imaging* **26,** 967–980 (2007).
13. Rousseau, F. *et al.* Registration-based approach for reconstruction of high-resolution in utero fetal MR brain images. *Academic radiology* **13,** 1072–1081 (2006).