Representing 3D Ultrasound with Neural Fields

Ang Nan Gu[1] GUANGNAN@ECE.UBC.CA
Purang Abolmaesumi PURANG@ECE.UBC.CA
[1] *Department of Electrical and Computer Engineering, University of British Columbia*
Christina Luong[2] CHRSTINA.LUONG@UBC.CA
[2] *Vancouver General Hospital, Vancouver, British Columbia*
Kwang Moo Yi[3] KMYI@CS.UBC.CA
[3] *Department of Computer Science, University of British Columbia*

Editors: Under Review for MIDL 2022

Abstract

3D Ultrasound (3D-US) is a powerful imaging modality, but the high storage requirement and low spatial resolution challenge wider adoption. Recent advancements in Neural Fields suggest a potential for efficient storage and construction of 3D-US data. In this work, we show how to effectively represent 3D-US data with Neural Fields, where we first learn the 2D slices of the 3D ultrasound data and expand to 3D. This two-stage representation learning improves the quality of 3D-US in terms of Peak Signal-to-Noise Ratio (PSNR) to 31.84dB from 28.7dB, a significant improvement directly noticeable to the human eye.

Keywords: Neural Fields, Ultrasound, 3D Reconstruction, Meta-Learning

1. Introduction

Ultrasound is a versatile medical imaging technology since it is low-cost, non-ionizing, and can display information in real-time. 3D ultrasound (3D-US) provides more comprehensive imaging of a region of interest than its more ubiquitous 2D counterpart. However, 3D volumes require more storage than 2D images, which puts high pressure on hospital data centers, and 3D-US generally have lower spatial resolution than 2D at a similar price point (Muraru et al., 2018).

Neural fields (Xie et al., 2021) provide a potential solution to both storage and spatial resolution problems. In essence, they learn to represent one data-point by deliberately overfitting to it on a voxel level. Reconstruction of the voxels using neural field weights leads to a more compact solution for storage. Furthermore, this neural representation is continuous, allowing for interpolation at arbitrary resolutions and quick slicing of the volume.

However, neural fields can be impractical since they are wholly trained on a single data-points of the entire volume, which is not how 3D-US data is stored or acquired—they come in slices. Here, we show that blindly training the entire 3D volume based on slices provide suboptimal results. Instead, we show that incorporating Meta-Learning (Tancik et al., 2021) to pretrain the network to learn a good global initialization for any 2D-US slice, then learning the entire 3D-US volume with this initialization leads to significantly improved results.

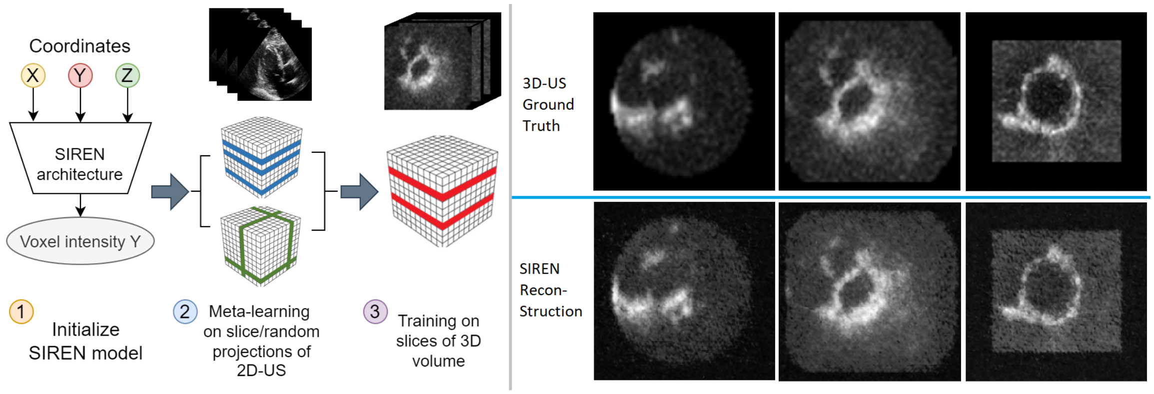

Figure 1: Left: Illustration of the training workflow. We conduct perform meta-learning on an initialized SIREN model before training slice-wise on the 3D volume. Right: Slices taken from ground-truth and learned SIREN representations for 3D cardiac ultrasound volumes.

2. Method and Discussion

We train a SIREN model (Sitzmann et al., 2020) to map coordinates $\mathbf{x} = (x, y, z)$ to grayscale voxel intensities $y = I(\mathbf{x})$. First, we use a meta-learning technique for SIREN which leverages the more commonly used high-resolution 2D-US image to build an intermediate representation. 2D-US images are mapped to (x, y, z) as slices, by projecting to random xy-plane (Meta-slice) and to random planes (Meta-random). We then train the SIREN on xy-plane slices of the volume, since the volume is too large to wholly fit into memory.

The SIREN model uses 3 hidden layers and $\omega = 30$, but we modify $\omega_0 = 240$ for faster training with respect to high-frequency characteristics like ultrasound speckle. We perform meta-learning with 2D-US using Reptile with 10^5 outer loops, 2 inner loops, and meta-learning rate $\beta = 10^{-5}$. For fitting 3D-US slices, we use the Adam optimizer with $lr = 10^{-4}$ and a batch size of 4 over 20000 iterations. 3D-US volumes are obtained from the STACOM dataset (Tobon-Gomez et al., 2013) with 208^3 voxels. 2D-US images are retrieved from an ethics-cleared, anonymous ultrasound dataset acquired by Vancouver General Hospital with *GE Vivid 7, Vivid i, Vivid E9*, and *Philips iE33* transducers. We implement SIREN and Reptile in Pytorch, and train/test on an NVIDIA Titan V GPU. The code for this project is available at https://github.com/an-michaelg/NeuralField3DUS, the supplementary materials within contain videos showing the SIREN's performance over every slice, with hyperparameter and meta-learning changes.

Our method successfully replicates high-level features of the cardiac ultrasound, such as bright regions indicating presence of the myocardium (heart muscle), heart valves, and the general "grainy" texture from speckle (Figure 1). Table 1 shows the average PSNR over xy-plane slices. Meta-learning improved the final PSNR since the 2D-US helped the initialization develop domain-specific properties. *Meta-slice* enhanced subsequent training more than *Meta-random* possibly due to a more comprehensive meta-learning of the volume.

Compared to storing the 3D volume as an `.npy` array which uses 8.9MB, the SIREN model can be stored as a 0.5MB Pytorch `state dict`, which is a compression ratio of around 18:1.

Method	Average slice PSNR
Glorot + 3D-Vol	28.68dB
Meta-slice + 3D-Vol	**31.84dB**
Meta-random + 3D-Vol	31.14dB

Table 1: PSNR comparison between standard (Glorot), slice-based and random projection-based meta-learning initializers on the 3D-US fitting task.

3. Conclusion

We learn an implicit representation for 3D ultrasound using the SIREN neural field architecture, and leverage meta-learning to develop a domain-specific initialization that improves its quality. Future work includes representation of 3D spatial-temporal or "4D"-US, reconstruction methods specific to certain anatomy (e.g. echocardiography), and whether more targeted forms of projections for 2D-US would be helpful during meta-learning.

References

Denisa Muraru, Alice Niero, Hugo Rodriguez-Zanella, Diana Cherata, and Luigi Badano. Three-dimensional speckle-tracking echocardiography: benefits and limitations of integrating myocardial mechanics with three-dimensional imaging. *Cardiovascular diagnosis and therapy*, 8(1):101, 2018.

Vincent Sitzmann, Julien Martel, Alexander Bergman, David Lindell, and Gordon Wetzstein. Implicit neural representations with periodic activation functions. *Advances in Neural Information Processing Systems*, 33, 2020.

Matthew Tancik, Ben Mildenhall, Terrance Wang, Divi Schmidt, Pratul P Srinivasan, Jonathan T Barron, and Ren Ng. Learned initializations for optimizing coordinate-based neural representations. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 2846–2855, 2021.

Catalina Tobon-Gomez, Mathieu De Craene, Kristin Mcleod, Lennart Tautz, Wenzhe Shi, Anja Hennemuth, Adityo Prakosa, Hengui Wang, Gerry Carr-White, Stam Kapetanakis, et al. Benchmarking framework for myocardial tracking and deformation algorithms: An open access database. *Medical image analysis*, 17(6):632–648, 2013.

Yiheng Xie, Towaki Takikawa, Shunsuke Saito, Or Litany, Shiqin Yan, Numair Khan, Federico Tombari, James Tompkin, Vincent Sitzmann, and Srinath Sridhar. Neural fields in visual computing and beyond. *CoRR*, abs/2111.11426, 2021. URL https://arxiv.org/abs/2111.11426.