What is Neural Radiance Field:

1. https://www.techtarget.com/searchenterpriseai/definition/neural-radiance-fields-NeRF
2. https://dellaert.github.io/NeRF/

NeRF application in medical:

1. https://collab.dvb.bayern/display/TUMdlma/NeRF+Applications+in+Medical+Imaging
2.

Creating 3D X-ray image from 2D X-ray images using NeRF:

1. https://developer.nvidia.com/blog/creating-3d-visualizations-from-x-ray-data-with-deep-learning/
2. https://www.zib.de/projects/3d-reconstruction-anatomical-structures-2d-x-ray-images
3. https://healthcare-in-europe.com/en/news/deep-learning-helps-visualize-x-ray-data-in-3d.html
4. https://www.iitb.ac.in/en/research-highlight/x-ray-to-3d

NeRF Applications in Medical Imaging

In this blog post, we will discuss the applications of NeRF (Neural Radiance Fields) in Medical Imaging. First, we will introduce the motivation for the NeRF implementations in Medical. We will later talk about current works that tackle stated problems. We will go through their methodology and important implementation details and discuss the results obtained by the authors. Lastly, we will state conclusions made by the author of the blog post about summarize the applications of NeRF in medical domain.

Author: Daria Matiunina

Tutor: Yeganeh, Y. M.

Motivation and Problem

NeRF's objective is to create a 3D reconstruction and novel views generation. Why is there a need for 3D reconstructions in Medical Imaging? First of all, it can be beneficial or even crucial for **medical procedures** such as cardiac surgeries planning [1], jaw reconstructive surgeries [2], hip and knee arthroplasties surgery planning [3], lung nodules removal surgery [1] etc. Secondly, it is now widely used for **implant planning** with examples being teeth implants, hip implants [3], jaw modeling [2] and other implant modeling. It has also been claimed [1] to be useful for education purposes: for teaching medical trainees, residents, cardiac surgeons; with patient education - for understanding of a planned approach in surgeries [1], for instance.

However, there are problems in Medical Image Analysis that can be tackled with the advances of NeRF or a close approach - NF (Neural Fields) [14]. To name a few discussed in the literature:

1. Difficulty of preserving geometric and topological structures in biomedical image segmentation which can be improved with NF shape modeling.
2. Difficulty of obtaining high-quality surfaces with the current approaches where NeRF could potentially assist with high level detail surface reconstruction.
3. High memory requirements for organ segmentation on high resolution scans where NF continuous implicit function representation can be used.

Furthermore, there are problems created by the **limitations of medical data** which make the application of some current approaches difficult.

- Capturing the data as discrete grids.
- Labeling noise.
- Incomplete borders or occlusions.
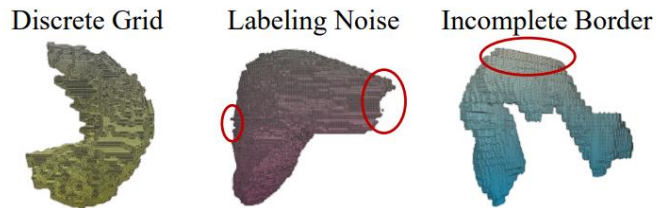- Lack of large datasets which is limiting the deep learning experiments [5].

Figure 1. Artifacts in data [4]

- Special cases: CT of spine and lower limb, pregnancy [6], cancer [5], ferromagnetic implants, non-contrastive scans [6]. Sometimes getting CTs is just not possible or dangerous for the patient. Non-contrastive scans are also claimed to give a bit worse segmentations [6].
- High costs: CT scans are more expensive than X-Ray or Ultrasound [6] and precise annotation is expensive [4].

Previous work

Before explaining the methodology of reviewed papers it is important to mention an important aspect of NeRF which as we will show later makes it *difficult to apply it to many tasks in medical domain*. NeRF has been trained on a set of synthetic and real world images that center around a non-opaque object. NeRF models the standard camera capturing process using the ray marching and radiance which is different from other capturing processes, e.g. computer tomography. We have found 5 papers that test NeRF for several very diverse tasks in medical domain (reconstruction of CT images from X-Rays, shape reconstruction from ultrasound images and endoscopy captures), and noticed much more research effort in applying the Neural Fields - models learning the mapping from a point location to a value with trained weights - and implicit representations ([4], [5], [8], [22], [23]).

We will mention the existing approaches in the most common tasks NeRF and Neural Fields have been applied to - **3D shape reconstruction and shape modeling**. This research direction and gathered works share similarities which we will describe in this post. We will also mention the goals and challenges of other papers that applied NeRF to their domain. Lastly, we will give an overview of research in NeRF enhancement that will be discussed in the context of future work for mentioned works.

3D Reconstruction from Ultrasound Images

NeRF has been researched for 3D reconstruction from ultrasound images as well as Endoscopic and X-Ray images [15], [17], [18], [4].

In this blogpost we will talk about ultrasound 3D shape reconstruction because it has been successful with NeRF. We will also mention other cases in the review of works.

The classic approaches to 3D reconstruction from ultrasound include voxel based, pixel based and function based methods. While the latter can be used for certain organs 3D reconstruction as was described in [9], the first two approaches are memory extensive or imprecise (pixel based images). Thus, work has been done in applying CNNs to sets of images as well as creating implicit representations of the shapes [6].
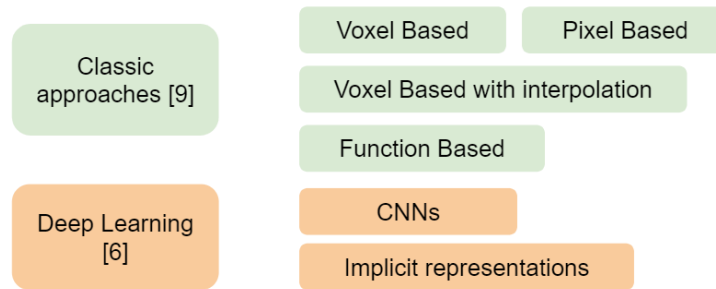
Figure 3. Categorization of approaches to 3D reconstruction from Ultrasound images using the study [6] and [9].

Shape Modeling in Medical Domain

Shape modeling in medical domain has been recently performed with different representations of the shapes - both explicit and implicit [4] with popular representations being signed distance fields, occupancy grids and implicit representations like Neural Fields. Approaches chosen in research have to deal with lack of data or noisy records and annotations in medical domain as we saw in Fig. 1. DeepSDF mentioned in the Fig. 4 has also served as a base for one of Neural Fields approaches described in Methodology.
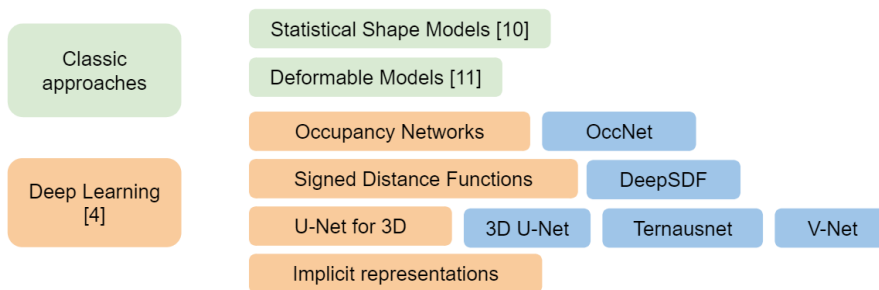

Figure 4. Categorization of approaches and examples for shape modeling using study [4] and [6].

NeRF works and their applications

The application of NeRF in all of the works reviewed relied heavily on certain assumptions or suffered from the design and limitations of NeRF. However, there has been a lot of work on improving NeRF in different directions which can be summarized in the Fig. 5. Thus, it would be of interest to adapt NeRF solutions and evaluate performance with new approaches (e.g. endoscopy 3D reconstruction or CT reconstruction [18]).
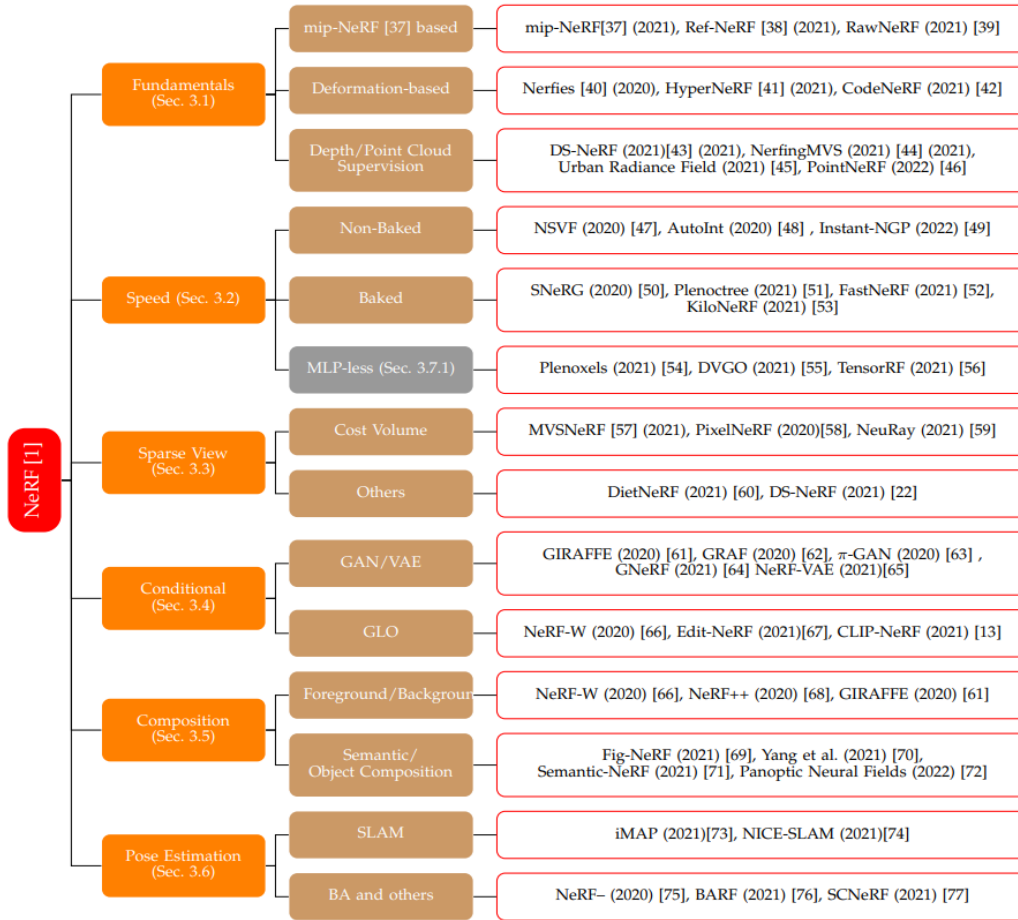
Figure 5. Source [19]. The taxonomy of NeRF-related papers. Different works tackling different limitations or experimenting with fundamental changes.

It has been challenging to collect the application fields for NeRF in Medical domain. The report by Gao et. al.[19] displays these popular fields of NeRF applications:
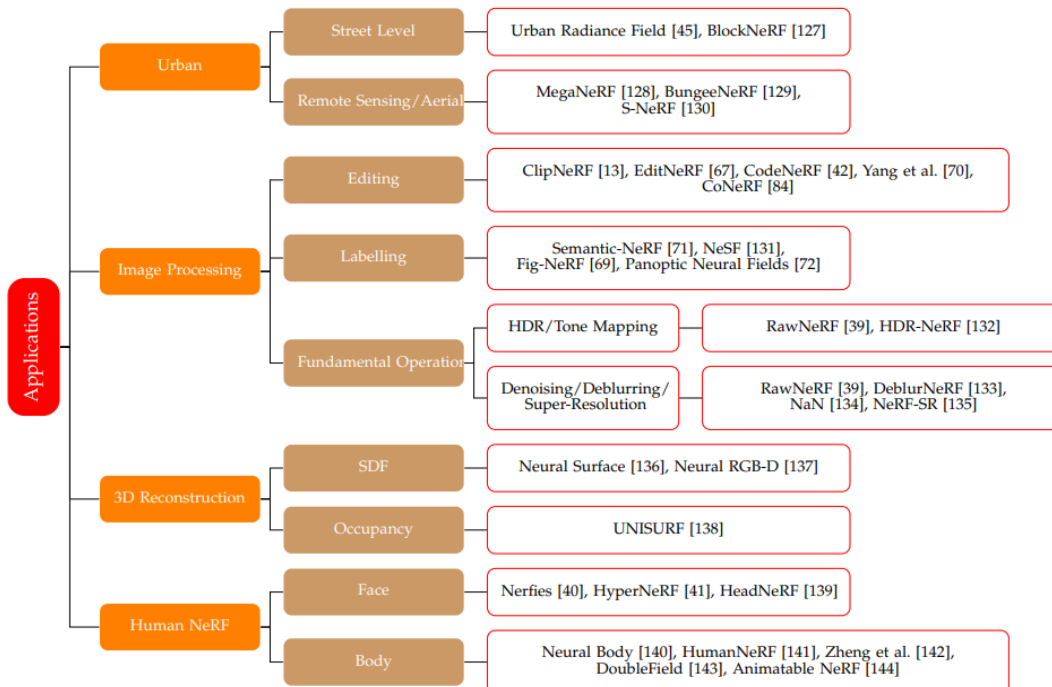

Figure 6. Source [19]. Applications of NeRF.

Methodology of NeRF and NeRF in Medical

NeRF. Neural Radiance Fields.

NeRF [12] has been introduced in 2019 as an approach to novel views rendering and scene 3D representation that is "contained" by a fully-connected, non-convolutional network. Basically, Neural Field stands for an MLP (Multi-Layer Perceptron) that has a 3D point (x, y, z) and camera viewing direction angles (α, β) as an input and outputs the volume density (e.g. zero if there is no object at this point) and view dependent radiance (rgb color) at the point from this viewing direction. NeRF uses the volume density and emitted radiance to characterize the scene structure and a traditional differentiable volume rendering to learn from the colors of the pixels in the input images.

Input   Output

R_1, t_1

R_2, t_2

...

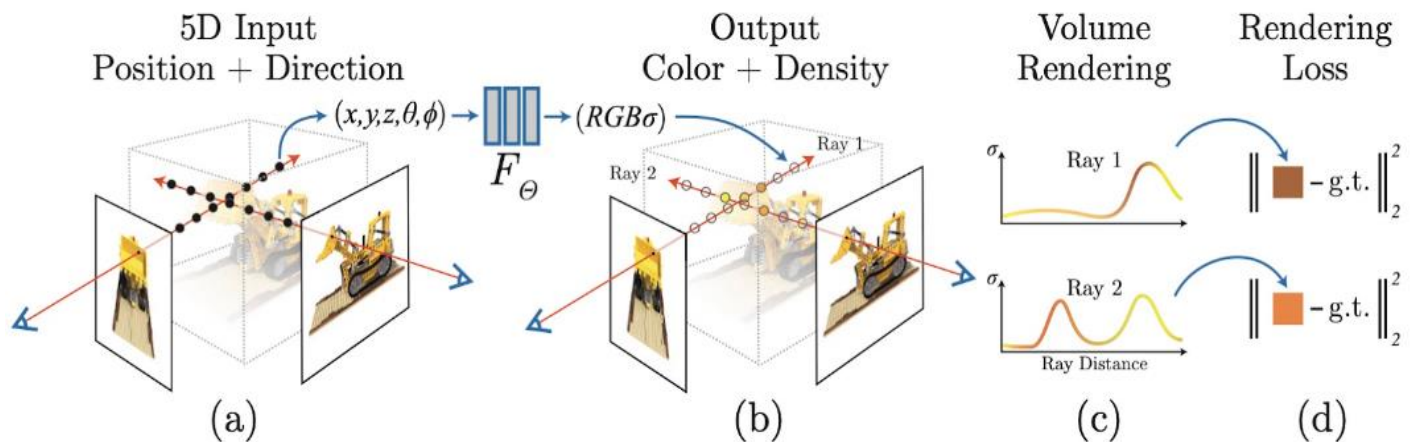Figure 2. NeRF's example input and output.

Figure 3. Source [12]. An overview of NeRF scene representation and rendering procedure: synthesize images by sampling 5D coordinates (location and viewing direction) along camera rays (a), feeding those locations into an MLP to produce a color and volume density (b), and using volume rendering techniques to composite these values into an image (c).

The **key features** of NeRF that are important in context of medical image analysis and 3D reconstruction are:

1. It's training for a single scene
2. Accurate camera estimations are needed
3. Captured as a dense set of high resolution camera views
4. Concentration of samples near opaque surfaces - type of scene (Fig. 7)
5. We can get depth estimation

6. Rendering is differentiable
7. NeRF might fail with non-Lambertian surfaces (Fig. 8) [13]

Figure 7. Dense grid of camera captures.

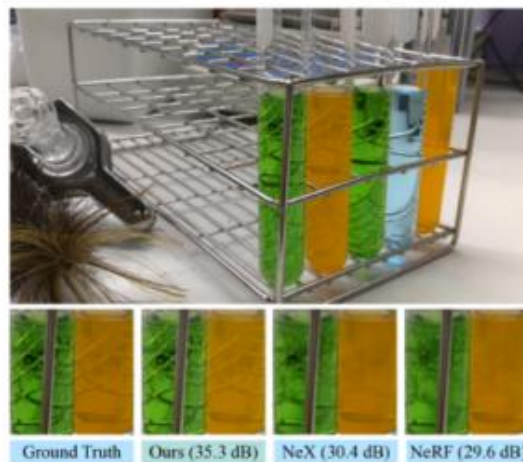Ground Truth   Ours (35.3 dB)   NeX (30.4 dB)   NeRF (29.6 dB)

Figure 8. Original NeRF implementation has difficulties with non-lambertian effects. Source [13].

NeRF and NF in Medical Domain

We have found a number of works following the NeRF approach for Medical Imaging, including this rendering process. It should be mentioned, that there is way more research in implicit representations for different applications in medical domain than the scientific work on NeRF for Medical. The implicit representations with a direct connection to a position in the scene are named in the literature as Neural Fields [19] and we will also list some of them and describe shortly as they are very similar to NeRF in terms of implicit representation of a scene or object.

The examples of works using NeRF in Medical are:

- 3D Ultrasound Spine Imaging with Application of Neural Radiance Field Method [15]
- MedNeRF: Medical Neural Radiance Fields for Reconstructing 3D-aware CT-Projections from a Single X-ray [16]
- NeAT: Neural Adaptive Tomography [17]
- 3D Reconstruction of Endoscopy Images with NeRF [18]

Many other papers discuss the Neural Fields and implicit representations in medical domain:

- Implicit Neural Representations for Medical Imaging Segmentation [5]
- Curvature-Enhanced Implicit Function Network for High-quality Tooth Model Generation from CBCT Images [8]
- ImplicitAtlas: Learning Deformable Shape Templates in Medical Imaging [4]
- Deep Implicit Statistical Shape Models for 3D Medical Image Delineation [22]
- CoIL: Coordinate-Based Internal Learning for Tomographic Imaging Results [23]

3D Ultrasound Spine Imaging with Application of Neural Radiance Field Method

The *motivation* for this paper is to create an alternative for radiograph measurement for scoliosis.

In this paper NeRF was used to get a better quality spine solid matter reconstruction. The bone structure on the image reconstructed by NeRF was clearer and more complete than of method used before - VNN (Voxel Nearest neighbors). The soft tissue was suppressed compared to the traditional reconstruction algorithm: for soft tissue the ultrasound propagation and reflection in different angles could vary greatly and lead to different responses in the image, while the bone tissue had the same reflection in the dataset. So NeRF neglected (suppressed) those varying soft tissue reflections and focused on static bone tissue.

In this case this has helped the authors to achieve better bone tissue reconstruction. In principle, NeRF can be suited to modeling view-dependent effects (non-Lambertian reflectance of specular and translucent surfaces). However, without regularization, this formulation allows degenerate solutions due to the the ambiguity between the surface and radiance, thus an incorrect shape can be matched with a high-frequency radiance function to minimize the optimization objective. In practice, NeRF avoids such solutions through its architecture, because the viewing direction is introduced only in the last layers of the model. This limits the expressiveness of the radiance function. Thus, at the cost of not handling the non-Lambertian effects well we don't get degenerate solutions.
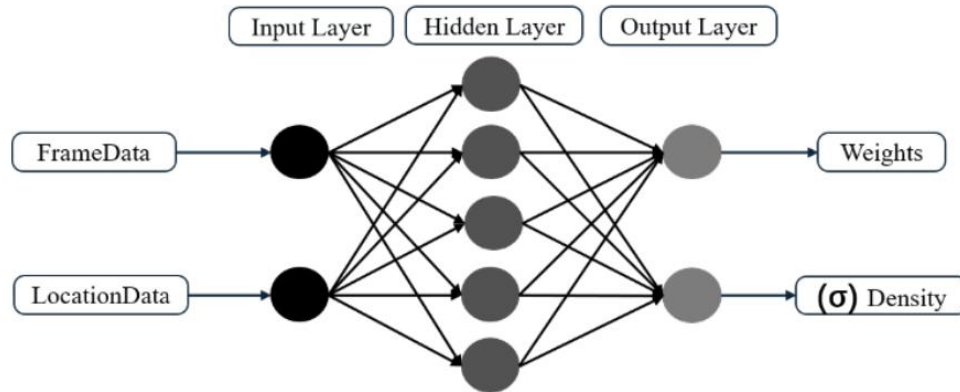

Figure 9. The MLP for ultrasound image reconstruction. Source [15]. Hidden layer is a fully-connected deep network.

In this papers authors used the NeRF architecture with a small change of a single channel output radiance as the data they worked on was not of regular red-green-blue pixel nature.

The MLP learns to predict the weights and the density. The weights can be seen as a probability density function along the ray that is being sampled (Fig. 10).

$$\widehat{w}_i = \frac{w_i}{\sum_{j=1}^{N_c} w_j} \quad (1)$$

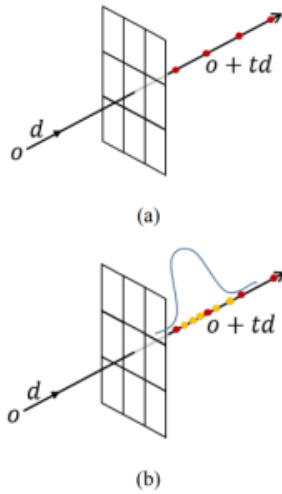$$w_i = T_i \cdot (1 - \exp(-\sigma_i \cdot \delta_i)) \quad (2)$$

Figure 10. Volume rendering. (a) Isometric volume sampling. (b) Hierarchical volume sampling. The weight here can be seen as probability density function along the ray.

Implicit Neural Representations for Medical Imaging Segmentation

The *motivation* for this paper is to create segmentation maps that are not discrete and are memory efficient.

How does this paper relate to NeRF? The authors created an implicit, trained segmentation function for high resolution medical scans. It is not a NeRF but a Neural Field according to [19] because it is a mapping from a point (x, y, z) to a value. The network maps a point (x, y, z)at a CT scan to an occupancy value (if the point is inside the organ or not).

Authors point out that using an implicit function for segmentation here brings these benefits:

- Allows to sample at any resolution during inference because learnt representations are continuous
- Converges faster than discrete voxel-based methods when there are size imbalances (organs have different sizes) (e.g. UNet - since it learns voxel distributions, it favors the larger organs and struggles to generate accurate predictions for the smaller ones). It was cited from another paper that INRs based decoders learn shape boundaries whereas CNN decoders learn voxel distributions, shape properties (e.g., surface curvature).
- Memory usage is independent of the resolution. Thus it can be very memory efficient to encode large 3D medical scans with continuous Implicit Neural Representation.

What's peculiar here: similar to the first paper [4], the multi-scale features are collected for every point and the mapping from the features for a point to the actual occupancy is trained. This way we get a scale-invariant mapping to a segmentation map, and we can inference for every desired resolution (we are only training on features for point).
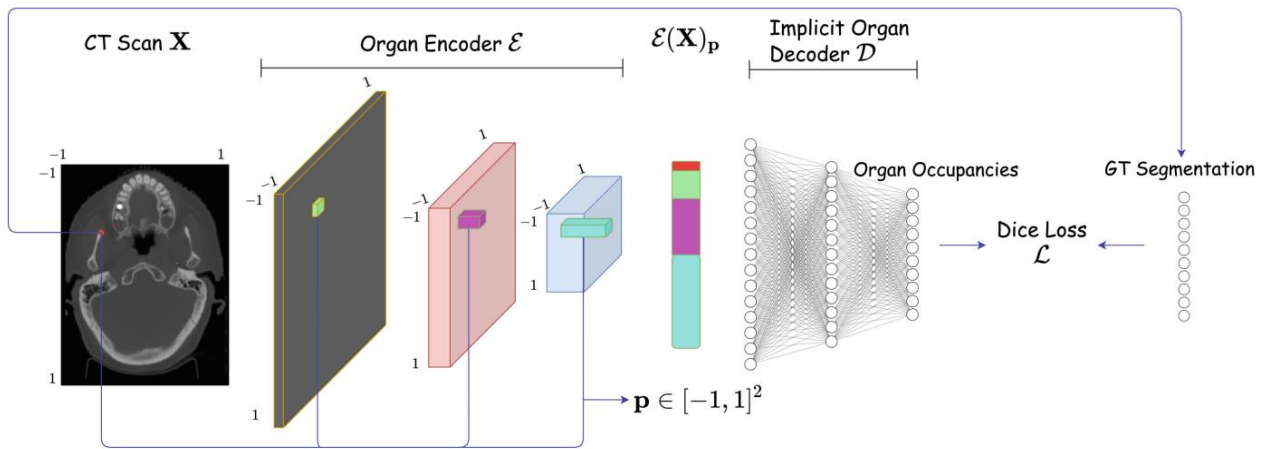
Figure 12. IOSNet architecture and training scheme, p stands for a point in a 2D input [5]

In the IOSNet architecture in Fig. 12 we can see an encoding of a point that is a concatenation of features extracted from different layers of encoder. Thus, it contains low level and high level features.

ImplicitAtlas: Learning Deformable Shape Templates in Medical Imaging

The motivation for this paper is perform precise and detailed shape completion, potential solution for dense correspondence and keypoint labeling.
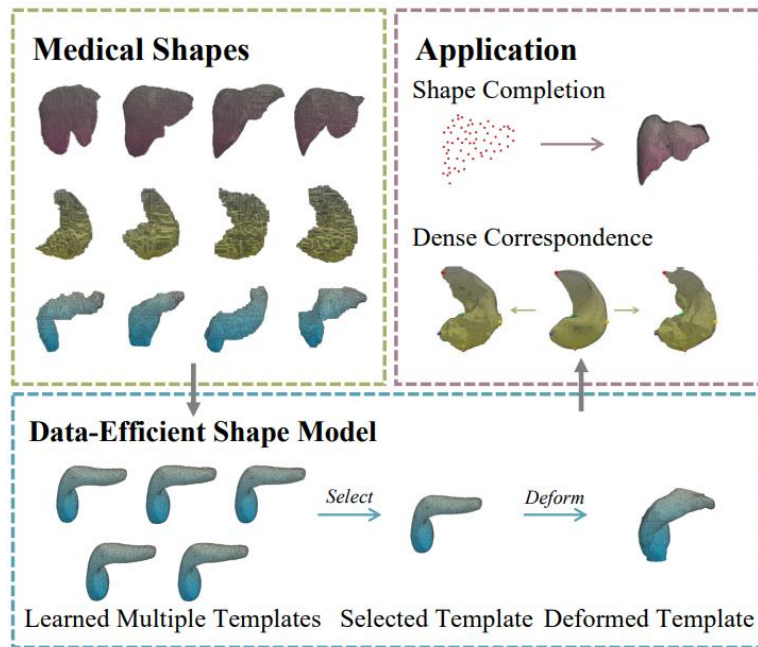

Figure 14. Use case. Reconstructing shapes by deforming one of learned templates.

How does this paper relate to NeRF: the shape representation in the approach is implicit and learned. It is not a NeRF but it is an approach that showcases shape modeling on limited and special medical data, plus shares the similarities with the previous mentioned paper. Authors show that building this implicit function using multi-scale features is important when working with medical data (because the data is limited, otherwise we need more to train similar approaches). Furthermore, having a template matching step helps the authors get better results with this implicit representation: a shape is less noisy because we optimize to match a template, and it is beneficial when the data is limited (not enough data to train shape representation without template matching). With template matching less computations are needed. At the same time, the deformation from the template is learned, so we train from the data and preserve the topology by using the template.

What was pointed out in the paper:

- Implicit representations in medical domain are often not perfect because large training datasets do not exist and in part because biomedical annotations are often noisy.
- When adapting this solution to multi-class shape reconstruction in medical imaging area, it will be a problem to deal with multiple objects in different poses and scales

Which methodology was suggested in the paper:

- Two convolutional nets (decoders of latent representations): one for template matching and one for deformation representation.
- Incorporate multiple templates to increase the representation capacity for known and unknown shapes (authors proved it does by testing on known and unknown shapes).
- Usage of templates when reconstructing is also a contribution.
- Use multi-scale features}to deal with the problem of limited data in medical domain (authors claimed that multi-scale features make the model less data-hungry than pure MLPs): so, extract feature maps at different resolutions instead of using only the final layers features.
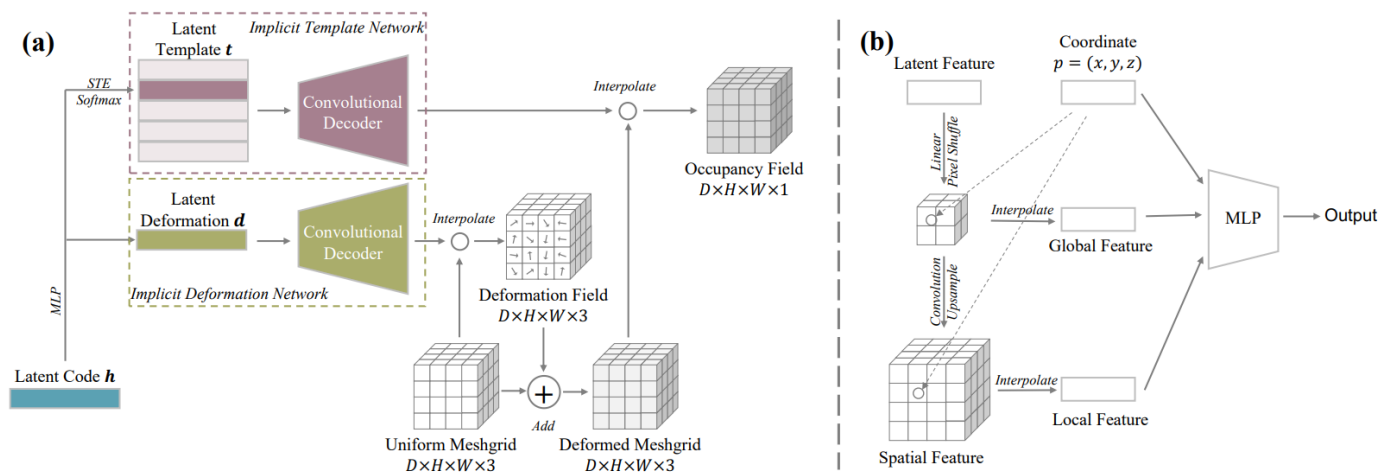

Figure 13. (a) Overview of ImplicitAtlas, (b) the architecture of the decoder [5].

Once again, similar to other papers mentioned, multi-scale features are used in the event of limited data. Here they are used in the Decoder part (Fig.13 b).

Evaluation and Results

The key results obtained by the authors of three reviewed papers are as follows:

3D Ultrasound Spine Imaging with NeRF:

- The bone structure with NeRF was clearer and more complete than of method used before - VNN (Voxel Nearest Neighbors). This is due to the special way NeRF handles non-Lambertian surfaces to avoid degenerate solutions. In this case, the suppression of soft tissues in the 3D shape has been beneficial for the goal of the project.
- The feasibility of NeRF for spine reconstruction depends highly on precise camera position estimations. In this study the authors did not have to estimate camera positions with COLMAP because they had a special attached sensor for positioning.
- Reconstruction with NeRF was claimed to be very computationally expensive.

Implicit Neural Representations for Medical Imaging Segmentation:

- Implicit representation for segmentation showed better segmentation of small organs than U-Net on different resolutions.
- The authors suggest an approach to obtaining high resolution segmentation maps with smaller memory requirements.
- The authors sho how we can leverage from multi-scale features when there is a lack of data.

ImplicitAtlas:

- Once again, multi-scale features were introduced to tackle the problem lack of data in organ annotations.
- The approach showed an improved shape reconstruction on limited noisy data.
- A possibility to model several shapes (templates) has been suggested and was experimentally shown to give better Dice Similarity Coefficient (DSC) and Normalized Surface Dice (NSD) metric results.

	NeRF US [15]	IOSNet [5]	ImplicitAtlas [4]
Data	Custom, SonixTABLET 640x480, 3D coordinates and orientation	AnatomyNet 2019 head and neck CTs and OARs segmentations, 271 CT scans and 9 OARs: (scan coordinate, label) pairs	Medical Segmentation Decathlon (MSD): liver, hippocampus, pancreas. Known and Unknown shapes
Evaluation baseline	VNN, radiograph measurements	U-Net	MLP of DeepSDF, MLP Decoder + Template, Conv Decoder
Evaluation criteria	Cobb angle	Dice score	Dice Similarity Coefficient, Normalized Surface Dice
Training setup	NeRF loss	UNet encoder for feature extraction, Dice loss	ImplicitAtlas Loss

Table 1. Details about data, evaluation baseline and loss function choice in reviewed papers.

In the first reviewed paper, 3D Ultrasound Imaging with NeRF, the goal of the project was to get a better calculation of Cobb angle than the one currently obtained with VNN. Authors claim, that with NeRF there is a slight improvement in this task and it is evaluated using the correlation with measurements made by radiograph (considered the most precise). They also evaluated the number of large discrepancy curves after curve measurement and noted that NeRF had less of those which is also better than VNN. An important finding of the paper was the suppression of reflections during NeRF training - for soft tissue in the region of interest it is mentioned that the ultrasound propagation and reflection in different
angles could varied to a big extend and cause different responses in the image, thus these image features could be called dynamic results. NeRF focused on reconstructing the static results instead - the bone tissue which was desirable.

	MAD	Correlation	Large Discrepancy Cures (>5°)
NeRF vs X-Ray	3.80	0.85	8
VNN vs X-Ray	4.44	0.82	11

Figure 14. Correlation, mean average distance and number of large discrepancy curves in comparison with radiograph [15].

In the paper of Implicit Neural Representations, authors show how their model, IOSNet was able to quicker learn small organs and tackle the problem of small organs detection/segmentation in general. They evaluated the model using Dice score and U-Net model as a baseline.
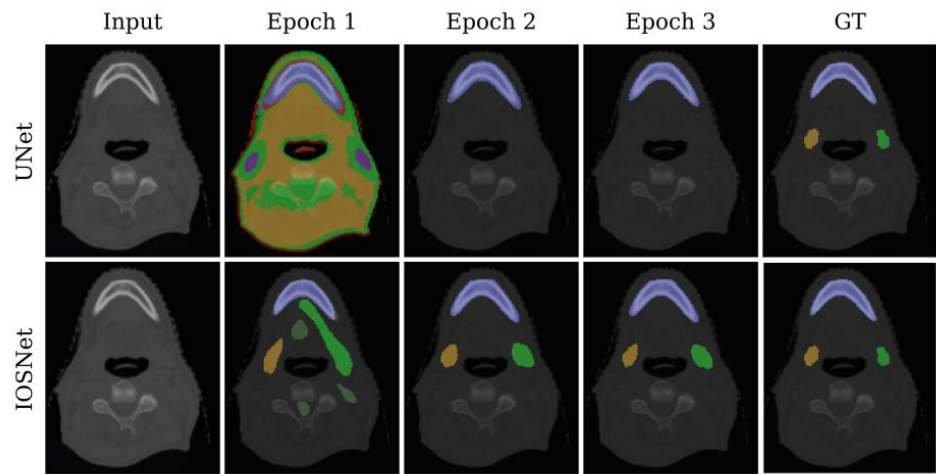

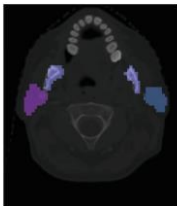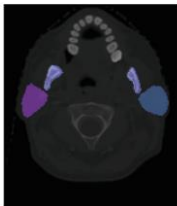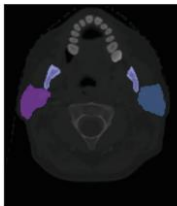Figure 16. IOSNet quicker learns small organs.
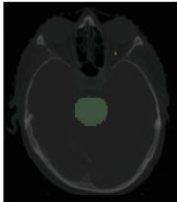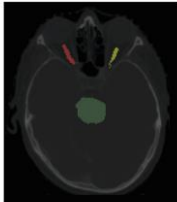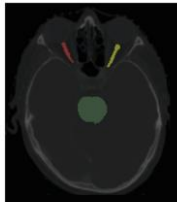
Input	Method	Dice
64^3	UNet (high-res)	7.3
	UNet (upsample)	54.3
	IOSNet	**64.0**
128^3	UNet (high-res)	48.2
	UNet (upsample)	62.4
	IOSNet	**73.2**

Table 2. Dice scores with different input sizes.

It has been shown that IOSNet also outperforms the UNet even under the setting of training on higher resolution.

Finally, for ImplicitAtlas it has been shown that using convolutional decoder significantly increased the capacity of shape representation. According to the authors, this indicates the importance of the spatial reductive bias for medical shapes.

The results have been evaluated using DSC (Dice Similarity Coefficient) and NSD (Normalized Surface Dice). In concept, DSC measures the volumes overlap while NSD measures the surfaces overlap.

ImplicitAtlas, which relies on multiple deformable templates improves the performance further than models not incorporating template matching.

Method	Liver (K5)		Hippocampus (K5)		Pancreas (K5)		Liver (U)		Hippocampus (U)		Pancreas (U)	
	DSC	NSD	DSC	NSD	DSC	NSD	DSC	NSD	DSC	NSD	DSC	NSD
MLP Decoder [38,44]	96.67	96.51	93.52	91.31	95.20	95.83	88.33	43.90	83.62	45.07	68.97	27.70
+ Template [12,72]	97.65	97.88	94.13	92.29	95.99	96.53	89.98	45.32	84.43	49.33	70.11	31.06
Conv Decoder [9,45]	98.41	98.38	95.46	87.73	96.67	97.00	91.26	55.13	87.10	50.92	71.79	29.15
ImplicitAtlas	**98.89**	**99.53**	**97.02**	**96.37**	**97.23**	**98.14**	90.64	48.27	88.37	54.40	74.71	34.87
ImplicitAtlas + reg.	98.71	99.05	96.48	93.93	96.90	97.31	**92.06**	**57.69**	**89.97**	**59.39**	**81.34**	**46.78**

Figure 17. ImplicitAtlas outperforms DeepSDF with Convolutional and MLP + template matching decoder.

D	MT	$\mathcal{L}_{LS}$	$\mathcal{L}_{DP}$	Liver (K)		Hippocampus (K)		Pancreas (K)		Liver (U)		Hippocampus (U)		Pancreas (U)	
				DSC	NSD	DSC	NSD	DSC	NSD	DSC	NSD	DSC	NSD	DSC	NSD
✓				98.22	98.01	95.62	91.17	96.15	95.62	96.52	85.43	93.18	74.54	90.61	74.43
	✓			82.65	26.18	80.13	30.98	55.39	15.66	79.24	22.42	76.82	27.21	51.64	14.16
✓	✓			**98.58**	**98.69**	**96.42**	**94.72**	96.85	97.30	96.59	85.95	93.54	76.99	93.38	81.11
✓	✓	✓		98.53	98.46	96.28	93.33	96.81	97.11	96.45	85.58	93.62	77.01	93.01	80.35
✓	✓		✓	98.52	98.42	96.32	93.10	**96.91**	**97.77**	96.69	86.21	93.79	77.13	**94.11**	**82.04**
✓	✓	✓	✓	98.50	98.33	96.09	92.85	96.76	97.03	**96.72**	**86.90**	**93.99**	**77.47**	93.31	80.92

Figure 18. For the unseen shapes the metrics with the regularization enforced were better than for learned shapes which shows that model was able to generalize better.

Lastly, it has been shown how the Laplacian smoothing and Deformation Penalty regularization improved the DSC (Dice Similarity Coefficient) and NSD (Normalized Surface Dice) scores on unseen shapes but decreased it a bit on model training data which means a better generalization of model.

My review
Strengths and weaknesses of presented papers
3D Ultrasound Spine Imaging with NeRF.

+ Data capturing process was described well. As camera position and capture grid is crucial for original NeRF implementation, the authors have described it in detail.

+ Experimental result of handling non-Lambertian effects in Medical domain.

- Limited information on training experiments, inference speed and model details.

Implicit Neural Representations for Medical Imaging Segmentation:

+ Extensive training experiments: dataset sizes, data modality, inference efficiency.

- Compared only to U-Net and not other trained implicit representations.

ImplicitAtlas:

+ Motivation stated well, good examples of challenging cases with data.

+ Experiments dealing with lack of data, including use of templates.

- How will this work with multiclass shape reconstruction? The authors have stated that it is a future research direction but multiple shapes from different viewing directions might introduce a lot of ambiguity during reconstruction, as different class shapes might be matched together from certain viewpoints.

Using NeRF and Neural Fields in Medical
Pros

1. suppression and highlighting: the mentioned previously **non-Lambertian effects are a challenge** for NeRF. While we saw how in ultrasound reconstruction this has helped the authors in the task of bone tissue reconstruction, in some other domains like tunnel-like capturing during endoscopy inability to reconstruct reflecting surfaces might affect the reconstructions negatively.
2. **high level of detail** in reconstructions: when combined with templates, multiple works showed NeRF and neural fields were able to create detailed high-fidelity reconstructions even with a lack of data.
3. implicit representations are **memory efficient**: we are no longer limited by discrete segmentation maps or voxel grids.
4. template matching or **enforcing a shape** is possible [19]: from my knowledge, recent works like Latent-NeRF [19] allow to enforce a shape template in other ways than suggested in reviewed papers. This might create new possibilities for 3D shape reconstruction.
5. **new implementations** tackle existing limitations: GRAF [26] to go away from a need for camera pose estimations, Latent-NeRF to improve the speed of training.

Cons

1. the inability to give good results with non-lambertian effects could be a problem in different application (e.g. endoscopy)
2. **lots of data needed + regular grid capture**. Furthermore, the patient has to be static during the capturing process. This is particularly difficult in medical domain. First, a dense grid of images is needed. This is researched further in versions of NeRF that allow **sparse capture grids**. The possibility to use for a **dynamicly changing scene** is also researched in new approaches like Nerfies, but in it's original form NeRF is unsuitable for dynamically changing scenes. Lastly, it seems from the application of NeRF to endoscopy, that **tunneled views are unsuitable** to NeRF. The scene has to be centered near a non-opaque object instead of inside it.
3. **long training** but can be solved with newer implementations as described in taxonomy from a recent report paper.
4. **camera positions needed** (also can be tackled by approaches like GRAF [26])
5. no generalization. I have not seen works where generalization to multiple shapes has been a goal other than the ImplicitAtlas [4]. In the latter, the multiple template network has suited the purpose.
6. **no uncertainty estimation**. The application of current NeRF-based methods in real scenarios is still limited since they are unable to quantify the uncertainty associated with the rendered views or estimated geometry [24].

   One work that aimed to work on this problem recently is Conditional-Flow NeRF: Accurate 3D Modelling with Reliable Uncertainty Quantification [24] and Stochastic neural radiance fields: Quantifying uncertainty in implicit 3d representations [25].

   Shortly, he idea of Stochastic Neural Radiance Fields was to model a distribution over all the possible radiance fields explaining the scene instead of learning a single radiance field as in original NeRF. During inference, this distribution is used to sample multiple color or depth predictions and compute a confidence score based on their associated variance.

Takeaways

1. It is often easy to implement **multi-scale features** in a model and get better results when the data is limited. In multiple papers authors leveraged from that with medical data.
2. **Data capturing** process is important if we want to use NeRF: we can match new NeRF implementations with the limitations of our case (e.g. sparse views or no camera pose estimation (GRAF [26])). However, special cases like endoscopy captures are drastically different from real-world scenes NeRF variations are developed for.
3. It can be beneficial to use **templates** or previous shape conditioning to ease the task for the model or deal with noisy data in medical domain. Sometimes we just don't have enough to train NeRF-like approaches on and captures are just noisy. Using priors is a good idea.
4. It might be a good idea to train for implicit representation of segmentation map in segmentation tasks **when the classes are imbalanced**. Small object detection is also known to be hard in research so implicit representations of segmentation maps might be an interesting research direction.

5. The **uncertainty estimation** for NeRF in Medical Domain should be paid special attention. With the recent works in uncertainty estimation for NeRF which are claimed to have an easy adaptation to the NeRF implementations, the implicit learned representations can be used for evaluation of approaches and to assist the professionals in the medical procedures.

References

[1] Mitzman B, "How-To: Creating a 3D Reconstruction of Your Patient's CT Scan", DOI:10.25373/ctsnet.13256381, 2020

[2] Narita, Masato, Takaki, Takashi, Shibahara, Takahiko, Iwamoto, Masashi, Yakushiji, Takashi, Kamio, Takashi, "Utilization of desktop 3D printer-fabricated "Cost-Effective" 3D models in orthognathic surgery", Maxillofacial Plastic and Reconstructive Surgery, 42. 10.1186/s40902-020-00269-0, 2020

[3] Hart, Zhe Su, Henckel, Di Laura, Schlüter-Brust, "3D-CT: A better map for hip surgery", Royal National Orthopaedic Hospital (RNOH), & Chair of Academic Clinical Orthopaedics, University College London (UCL), St. Franziskus Hospital Köln, https://www.opnews.com/2017/10/3d-ct-better-map-hip-surgery/14077

[4] J. Yang, U. Wickramasinghe, B. Ni, and P. Fua, "ImplicitAtlas: Learning deformable shape templates in medical imaging," IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 2022, pp. 15 840–15 850.

[5] M. O. Khan and Y. Fang, "Implicit neural representations for medical imaging segmentation," Medical Image Computing and Computer Assisted Intervention – MICCAI 2022, L. Wang, Q. Dou, P. T. Fletcher, S. Speidel, and S. Li, Eds. Cham: Springer Nature Switzerland, 2022, pp. 433–443.

[6] S. Hosseinian, H. Arefi, "3D Reconstruction from multi-view medical X-ray images – review and evaluation of existing methods", Photogrammetry, Remote Sensing and Spatial Information Sciences, Volume XL-1/W5, 2015 International Conference on Sensors & Models in Remote Sensing & Photogrammetry, 2015, DOI:10.5194/isprsarchives-XL-1-W5-319-2015

[7] Valenzuela, Waldo & Vermathen, Peter & Boesch, Chris & Nolte, Lut-Peter & Reyes, Mauricio. "iSix - Image Segmentation in Osirix", 30th Annual Scientific Meeting of the European Society for Magnetic Resonance in Medicine and Biology, 2013

[8] Y. Fang, Z. Cui, L. Ma, L. Mei, B. Zhang, Y. Zhao, Z. Jiang, Y. Zhan, Y. Pan, M. Zhu, and D. Shen, "Curvature-enhanced implicit function network for high-quality tooth model generation from cbct images," in Medical Image Computing and Computer Assisted Intervention – MICCAI 2022, L. Wang, Q. Dou, P. T. Fletcher, S. Speidel, and S. Li, Eds. Cham: Springer Nature Switzerland, 2022, pp. 225–234

[9] Solberg, Lindseth, Torp, E. Blake, Nagelnus Hernes, "Freehand 3D Ultrasound Reconstruction Algorithms - A review", 2007, Ultrasound in Med. & Biol., Vol. 33, No. 7, pp. 991–1009, World Federation for Ultrasound in Medicine & Biology, DOI:10.1016/j.ultrasmedbio.2007.02.015

[10] Heimann, Meinzer, "Statistical shape models for 3D medical image segmentation: A review", Medical Image Analysis, Vol. 13, 2009, pp. 543-563, DOI:10.1016/j.media.2009.05.004.

[11] Tim McInerney, Demetri Terzopoulos, "Deformable models in medical image analysis: a survey", Medical Image Analysis, 1996, Vol.1 , pp. 91-108, DOI:10.1016/S1361-8415(96)80007-7.

[12] Ben Mildenhall and Pratul P. Srinivasan and Matthew Tancik and Jonathan T. Barron and Ravi Ramamoorthi and Ren Ng, "NeRF: Representing Scenes as Neural Radiance Fields for View Synthesis", ECCV 2020, 2020

[13] Ben Mildenhall and Pratul P. Srinivasan and Rodrigo Ortiz-Cayon and Nima Khademi Kalantari and Ravi Ramamoorthi and Ren Ng and Abhishek Kar, "Local Light Field Fusion: Practical View Synthesis with Prescriptive Sampling Guidelines", ACM Transactions on Graphics (TOG) 2019

[14] Xie, Takikawa, Saito, Litany, Yan, Khan, Tombari, Tompkin, Sitzmann, Sridhar, "Neural Fields in Visual Computing and Beyond", EUROGRAPHICS 2022, Volume 41, arXiv:2111.11426v4, 2022

[15] H. Li, H. Chen, W. Jing, Y. Li, and R. Zheng, "3D Ultrasound Spine Imaging with application of Neural Radiance Field Method," in 2021 IEEE International Ultrasonics Symposium (IUS), 2021

[16] Corona-Figueroa et. al., "MedNeRF: Medical Neural Radiance Fields for Reconstructing 3D-aware CT-Projections from a Single X-ray", arXiv:2202.01020v3, 2022

[17] Rückert et. al., "NeAT: Neural Adaptive Tomography", arXiv:2202.02171v1, 2022

[18] Qin Ying Chen, "3D Reconstruction of Endoscopy Images with NeRF", Doctoral Dissertation, NYU Tandon School of Engineering, 2023

[19] Gao et. al. "NeRF: Neural Radiance Field in 3D Vision, A Comprehensive Review", IEEE Transactions on Pattern Analysis and Machine Intelligence, arXiv:2210.00379v2, 2022

[20] J. Shen, A. Agudo, F. Moreno-Noguer, and A. Ruiz, "Conditional-flow nerf: Accurate 3d modelling with reliable uncertainty quantification," arXiv:2203.10192, 2022

[21] J. Shen, A. Ruiz, A. Agudo, and F. Moreno-Noguer, "Stochastic neural radiance fields: Quantifying uncertainty in implicit 3d representations," arXiv:2109.02123, CoRR, vol. abs/2109.02123, 2021

[22] Raju, Miao, Jin, Lu, Huang, Harrison, "Deep Implicit Statistical Shape Models for 3D Medical Image Delineation", arxiv:2104.02847, 2021

[23] Sun, Liu, Xie, Wohlberg, Kamilov, "CoIL: Coordinate-based Internal Learning for Imaging Inverse Problems", arXiv:2102.05181v1, 2022

[24] J. Shen, A. Agudo, F. Moreno-Noguer, and A. Ruiz, "Conditional-flow NeRF: Accurate 3D modelling with reliable uncertainty quantification," arXiv:2203.10192, 2022

[25] Shen, A. Ruiz, A. Agudo, and F. Moreno-Noguer, "Stochastic neural radiance fields: Quantifying uncertainty in implicit 3D representations," CoRR, vol. abs/2109.02123, arXiv:2109.02123, 2021

[26] Schwarz, Katja and Liao, Yiyi and Niemeyer, Michael and Geiger, Andreas, GRAF: "Generative Radiance Fields for 3D-Aware Image Synthesis", Computer Vision and Pattern Recognition (cs.CV), arXiv:2007.02442, 2020