Continuum Robotics and NeRF for 3D Reconstruction of Subglottic Stenosis

Jinjie Sun, Julia Wiercigroch
Department of Computer Science, University of Toronto

Motivation

- Subglottic stenosis (SGS) is a rare disease that gradually narrows the airway between trachea and vocal cords
- Clinicians only get a top-down view from endoscopic exams to determine the length and width to justify surgical intervention

Figure 1: Top-Down View of SGS from Real (Top) and Simulated (Bottom) Endoscopic Video

- Computer-assisted diagnosis with continuum robotics and endoscopic 3D reconstruction can provide clinicians with a measurement tool for SGS

Figure 2: Proposed System for Measuring SGS

- **Goal:** Compare neural radiance field methods to reconstruct a virtual SGS model with sparse camera viewpoints, limited texture, and scene specularities

Related Work

Structure from Motion Reconstruction

- Textureless surfaces, reflections and poor illumination makes it difficult to detect and match features [1]

Neural Radiance Fields (NerF)

- Model considers RGB reconstruction loss
- Training time is long and requires significant number of images of the same object [2]

Neural Radiance Fields with Monocular Depth and Normal Cues (MonoSDF)

- Model considers RGB reconstruction, depth, normal and Eikonal losses [3]
- Additional losses help reconstructions from sparse viewpoints and complex scenes [3]

References

[1] Velez, A., Marcinczak, J., Grigat, R., Structure from motion approaches to 3D reconstruction in minimal invasive laparoscopy, ICIAR, 2012

[2] Mildenhall, B., Srinivasan, P., Tancik, M., *et. al.*, NeRF: Representing Scenes as Neural Radiance Fields for View Synthesis, ECCV, 2020

[3] Yu, Z., Peng, S., Niemeyer, M., *et. al.,* MonoSDF: Exploring Monocular Geometric Cues for Neural Implicit Surface Reconstruction. NeurIPS, 2022

Methods For Virtual SGS Reconstruction

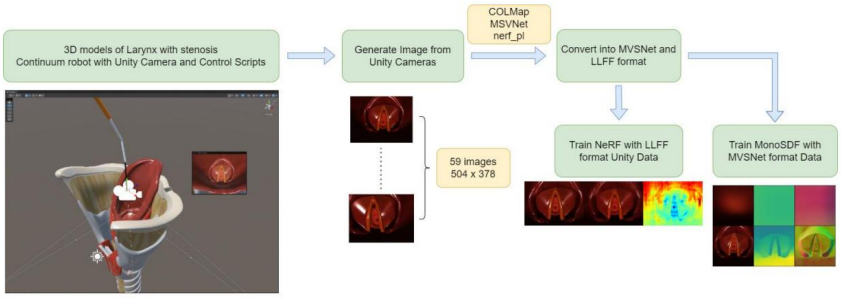

Figure 3: Virtual SGS NeRF Reconstruction Workflow

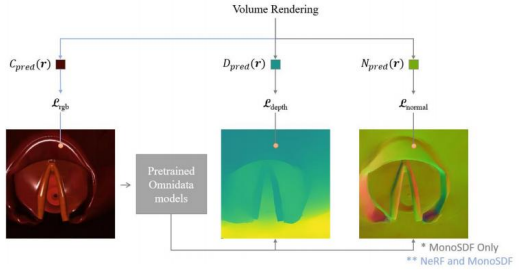

Figure 4: Overview of NeRF and MonoSDF Algorithms

Experimental Results

- **Trained 3 models with Unity data**
  - NeRF with the full dataset (59 images)
  - NeRF with the subset of the dataset (35 images)
  - MonoSDF with the full dataset (59 images)

Table 1: PSNR results from reconstructed models and ground truth

Method	Training Epoch	**Mean**	Standard Deviation	Image 0	Image 3	Image 8	Image 45	Image 50
NeRF	1	25.32	1.30	24.25	26.87	26.82	24.9	23.75
NeRF(35 Image)	1	17.28	1.20	16.69	15.80	16.50	18.56	18.84
MonoSDF	500	18.89	2.27	20.97	20.79	20.45	16.25	16.00

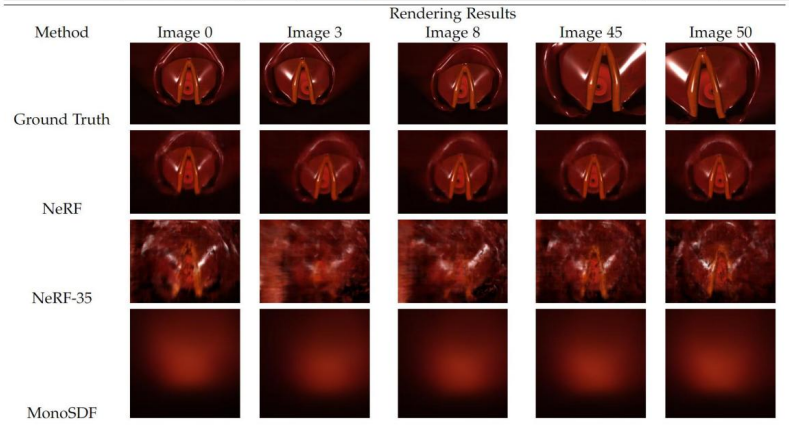

Figure 5: Rendering results of different reconstruction models on the test images

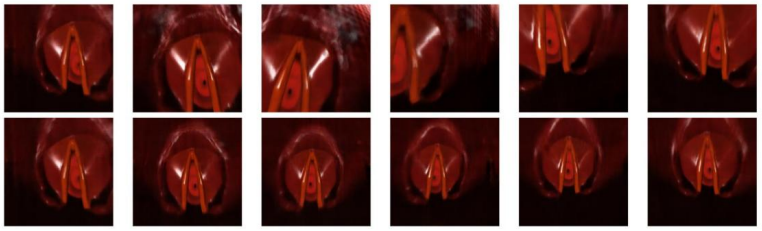

Figure 6: Rendering results of NeRF on different simulated camera positions