# Sierra Bonilla

Computational Scientist
University College London Hospitals

+44 7713 961437
sierrabonilla14@gmail.com
github.com/smbonilla
linkedin.com/in/sierra-bonilla
sierrabonilla.com

## Scene Reconstruction in Endoscopic Imaging Using Neural Rendering Frameworks

– Supervisor: Dr. Sophia Bano (sophia.bano@ucl.ac.uk)

## Research Proposal

Fetoscopy and cystoscopy are both minimally invasive endoscopic medical procedures used to view the fetus and the bladder, respectively. These procedures regularly involve the insertion of a thin, flexible tube containing a monocular camera and a light source and are performed for surgical procedural guidance or diagnosis [1]. Compared with open surgery, they have many advantages including shorter hospital stays, less trauma to the patient, and smaller skin scars. However, they are limited in their field-of-view and maneuverability [2]. These limitations in fields-of-view and low image quality of the camera make it difficult to accurately understand the anatomy. However in recent years, there has been a surge of research in the domains of computer graphics and computer vision that is rapidly changing the field of endoscopic imaging [3].

Classically computer graphics techniques have been created for and applied to video games and computer generated imagery (CGI) for movies. Content creation in computer graphics has relied on copious manual work of CGI artists and ongoing research on representing 3D spaces on a 2D screen. Simultaneously research in the field of computer vision, the process of deriving high-level information from visual inputs, has advanced significantly. While computer graphics take the approach of building a *physical* model of an object, computer vision relies on a creating a *statistical* model of the object. Neural rendering combines these approaches by first representing and then generating, thus addressing both reconstruction and rendering needs. By incorporating the knowledge learned in computer graphics on representation of 3D spaces on a 2D screen, machine learning can take scene reconstruction a step further by creating photo-realistic outputs of varying arbitrary views [1].

Neural rendering frameworks (NRFs), like generative query networks (GQN) and neural radiance fields (NeRFs), are powerful scene reconstruction tools for reconstructing 3D digital versions of the real world from 2D images and videos, even from sparse input views. In 2018, Eslami et al. proposed the GQN model: a framework for machines to learn about their surroundings, composed of two parts – a representation network and a generative network [4]. The representation network works by taking the observations as inputs, like snapshot images from the endoscope camera, and produces a representation of the scene as a vector. The generative network then generates a prediction of the scene given the representation vector from a different perspective. The two models are trained jointly to increase chance of creating ground-truth scene.

Neural radiance fields (NeRFs) are another type of neural rendering framework that can generate new views of 3D scenes based on an incomplete 2D image set of the data. NeRFs use a non-convolutional deep network with a singular 5D coordinate input (spatial location and viewing direction) and the output is volume density and view-dependent radiance at a specific location. NeRFs have been shown to be remarkably efficient and accurate for synthesizing novel views of complex scenes, although the exact location and viewing direction of the object is necessary [5].

Using NRFs for creating 3D photo-realistic models of the surgical scenes has the potential to mitigate the aforementioned limitations of fetoscopic and cystoscopic procedures and therefore improve surgical outcomes. This technique can be used to provide a more comprehensive 3D view of the scene even from sparse or limited fields of view for better understanding of the anatomy leading to more accurate diagnoses. This technique can also be used for developing further context aware computer-assisted intervention systems. With a better representation of the surgical scene, further models will have more information for accurate real time monitoring of surgical phases, assessing surgical skills, and alerting surgeons about possible complications [2]. In this research proposal, I propose to investigate the use of NRFs, like GQNs and NeRFs, for scene reconstruction in endoscopic imaging, with a particular focus on fetoscopy and cystoscopy.

## Research Questions

1. What existing NRF techniques are available for inferring 3D scenes from low-resolution and monocular camera feeds?
2. Are these techniques applicable for fetoscopic and cystoscopic medical procedures and what are the possible applications?
3. What improvements can be made to the accuracy and robustness of NRF-based scene reconstruction?
4. How can these proposed methods be evaluated in a real-world setting?
5. What are the challenges and ethical concerns associated with implementing the proposed methods in a clinical setting?

## Research Methodology

1. Perform a literature review on NRF techniques and analyze their effectiveness for inferring 3D scenes from low-resolution and monocular camera feeds.
2. Speak with surgeons to determine challenges associated with implementing AI systems such as these in clinical settings and what ethical and social concerns should be considered.
3. Test the possible NRF techniques on the available fetoscopic and cystoscopic datasets and evaluate the results.
4. Identify areas of improvement in the accuracy and robustness of NRF-based scene reconstruction and develop techniques to address the identified areas of improvement.
5. Discuss potential applications of the proposed methods in fetoscopic and cystopscopic medical procedures.
6. Develop research papers to document the findings of the research.

## Outcome

I expect to achieve a comprehensive study of existing NRF techniques and analyze their effectiveness for inferring 3D scenes from low-resolution and monocular camera feeds. I will also explore, develop, and validate techniques for improving the accuracy and robustness of NRF-based scene reconstruction. Finally, I will discuss the potential applications of the proposed methods in fetoscopic and cystopscopic medical procedures.

## References


[1] M. Baumhauer, M. Feuerstein, H.-P. Meinzer, and J. Rassweiler, “Navigation in endoscopic soft tissue surgery: perspectives and limitations,” *Journal of endourology*, vol. 22, no. 4, pp. 751–766, 2008.

[2] “Artificial intelligence in minimally invasive surgery | sophia bano,” Dec 2021. [Online]. Available: https://www.youtube.com/watch?v=7h-PYvFe38Y

[3] S. Bano, F. Vasconcelos, M. Tella-Amo, G. Dwyer, C. Gruijthuijsen, E. Vander Poorten, T. Vercauteren, S. Ourselin, J. Deprest, and D. Stoyanov, “Deep learning-based fetoscopic mosaicking for field-of-view expansion,” *International journal of computer assisted radiology and surgery*, vol. 15, no. 11, pp. 1807–1816, 2020.

[4] S. A. Eslami, D. Jimenez Rezende, F. Besse, F. Viola, A. S. Morcos, M. Garnelo, A. Ruderman, A. A. Rusu, I. Danihelka, K. Gregor *et al.*, “Neural scene representation and rendering,” *Science*, vol. 360, no. 6394, pp. 1204–1210, 2018.

[5] B. Mildenhall, P. P. Srinivasan, M. Tancik, J. T. Barron, R. Ramamoorthi, and R. Ng, “Nerf: Representing scenes as neural radiance fields for view synthesis,” in *ECCV*, 2020.