Designing a Synthetic Ultrasound-Based Machine Learning Framework to Differentiate Renal Calculi From Ultrasound Artifacts and Evaluating Domain Transfer: Year 2

Kulkarni, Archit (School: Melbourne High School)

Kidney stones are solid mineral deposits that cause severe pain or harm to the urinary tract if not diagnosed and treated early. However, medical machine learning models require large labeled clinical datasets that are difficult to obtain. The goal of this project was to develop a machine learning framework that distinguished kidney stones from ultrasound artifacts using synthetic ultrasound images. This study investigated whether physics-based ultrasound simulation could generate data sufficient to train a convolutional neural network (CNN) with validation accuracy greater than 75%. Using Field II simulation software, 400 B-mode images were synthesized using controlled parameters. A ResNet-18 CNN was trained on images of stone-only and stone with reverberation artifacts. The model achieved 100% accuracy on the synthetic validation dataset, exceeding the target. To test real-world transfer, the model was evaluated on 1884 real ultrasound images (1001 stone, 883 no-stone). The model assigned higher probabilities to real stone-present images than to no-stone images (mean scores: 0.257 vs. 0.168). Statistical testing confirmed this difference was highly significant (KS $p < 0.001$, Mann-Whitney $p < 0 .001$). UMAP dimensionality reduction was used to map similarity between the two classes. A clear separation between clusters suggested that the learned features captured meaningful differences between real stone and no-stone images. These findings demonstrate the potential of synthetic images as an alternative training source for medical machine learning models when clinical data is limited. Future work will improve synthetic data realism using Generative Adversarial Networks (GANs) and additional simulated artifacts.

Awards Won: