Endometriosis Modeling 101

This paper explains some details about The Endo Project for the general public. These instructions are located in my Github [here](https://github.com/awicklund/EndoObjDetect). The Jupyter Notebook used for the project is located [here](https://github.com/awicklund/EndoObjDetect/blob/main/Google%20Colab/Copy_of_Roboflow_TensorFlow2_Object_Detection.ipynb). (A Jupyter Notebook is like a word document that runs computer code for you. For more information about Jupyter Notebook, click [here](https://jupyter.org/).)

What is Endometriosis?

Endometriosis is defined as: a condition resulting from the appearance of endometrial tissue outside the uterus and causing pelvic pain. Endometrial tissue is the tissue created in the uterus every month to hold and protect a fertilized egg. Sometimes this tissue grows on the outside of the uterus, on the ovaries, or even other organs such as the kidneys, bladder, or liver. It is most commonly found in persons with a uterus, but may develop in those who do not have a uterus or have had a hysterectomy.

What am I looking at in Github?

The EndoObjDetect repository contains The Endo Project. The Endo and my\_tfod folders are folders that can be looked at if you are interested in building your own model. The Google Colab folder contains two Jupyter Notebooks. One is a copied notebook about how to make an object detection model, and the other is how I made changes to it to suit the model’s needs. The folders Finished examples and Graphs both contain images of what the model produces as well as how the model runs.

What am I looking at in Finished examples?

Finished examples has 10 images without annotations added, and 10 with the annotations added. Those that start with “alpha” have annotations. The green rectangle indicates where my model guessed the endometriosis to be, and the percentage is how confident the model is of its answer. The light spots on the “alpha” images are where doctors say the endometriosis is located. To find out more about why this model is missing a few spots, feel free to read the Conclusions and Recommendations section of the technical paper in my Github.

What do the graphs mean?

Loss is the inability to classify the image properly. When the model misses its target, loss increases. We want loss to be low. Steps per second is the speed the model processes, which we want to be high. The lightly shaded parts of the graph are the original data and the jagged darker orange line is the smoothed result. It’s easier to read after applying some “smoothing” math to the graphs. The RPN graphs are more interesting since they determine how well the rectangle is placed, rather than “Is this spot background? Or endometriosis?”

How can I make a model of my own?

I would suggest giving the technical paper a read before starting. It explains the issues I had to iron out to get the model working. It’s mostly prep work with the images, and can be done by altering the code slightly in the copied notebook. If you need help finding a database, try [Google Dataset Search](https://datasetsearch.research.google.com/). Also, if you are interested in looking at multiple, different objects in an image, try the [COCO](https://cocodataset.org/#home) dataset. Happy Modeling!

Got a question? Or a neat project to share? E-mail me at [theendoproject@yahoo.com](mailto:theendoproject@yahoo.com)