Auto-PCOS Classification Challenge

Team name: KLIV-IITKGP

Team member names and affiliation:

- Raj Krishan Ghosh (Centre of Excellence in Artificial Intelligence, Indian Institute of Technology Kharagpur, India)
- Abhishek Kumar (Centre of Excellence in Artificial Intelligence, Indian Institute of Technology Kharagpur, India)
- Dr. Debdoot Sheet (Department of Electrical Engineering, Indian Institute of Technology Kharagpur, India)

Contact number: 9038361203

Email address: rajkrishanghosh@gmail.com

Brief write up about the pipeline:

Our approach uses EfficientNet-B7 based network for Auto-PCOS Classification. To facilitate faster training, we used pretrained EfficientNet-B7 model and finetune it by training for 10 epochs. The input ultrasound images (which are of different sizes) are scaled to a size of 256×256. While splitting the training data into training and validation datasets, we ensured equal fraction of healthy and unhealthy sample distribution to limit overfitting to any one category. We experimented with other input image sizes, but 256×256 gave best validation results, corresponding to epoch 2, after which the model started to overfit to the training dataset. We also tried using different information to the three channels, including edge detection and emboss, but it made the training worse.
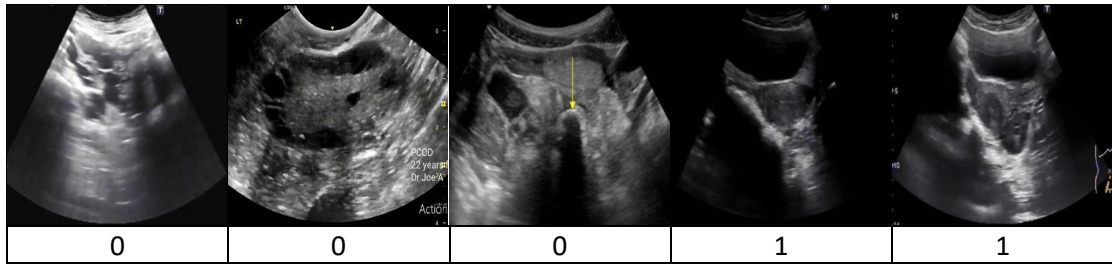
Figure of the developed pipeline:
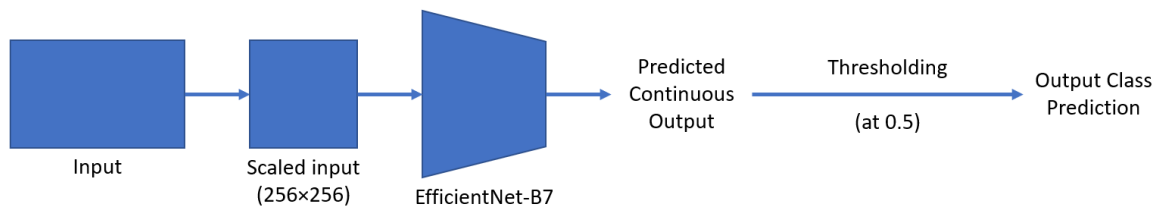
Achieved results on validation dataset:

- Table of achieved evaluation metrics:

Metric	Value
Accuracy	0.7878315132605305
Precision	0.8461538461538461
Recall	0.30386740331491713
F1-Score	0.4471544715447155
AUC-ROC	0.6410641364400673

- Best 5 frames with classification on testing data:

0	0	0	1	1

- We could not finish interpretability part, we are therefore submitting without the interpretability images.

Achieved results on testing dataset including:

- Best 5 frames with classification on testing data:
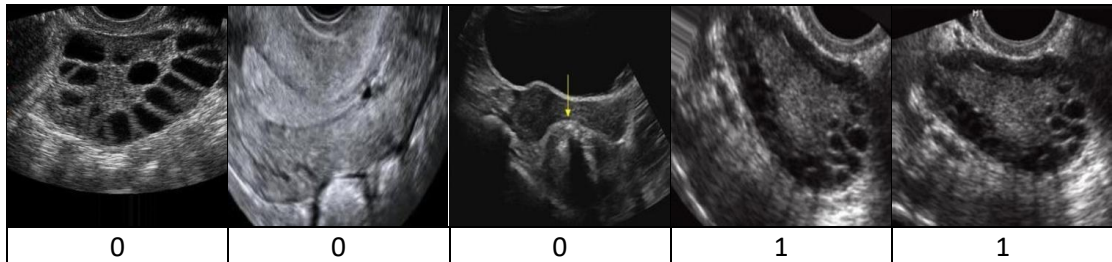
0	0	0	1	1

- We could not finish interpretability part, we are therefore submitting without the interpretability images.

NOTE: Running the codes (sequentially) after downloading from GITHUB:

- Put *PCOSGen-test* and *PCOSGen-train* folders in *Dataset/* folder (empty right now to upload on GITHUB).
- Run *data_preprocessing.py* to split the data into training and testing datasets.
- Run train.py to train the model (the model .pth file is not uploaded to GITHUB as it has 732.17MB file size)
- Run *validation.py* to predict on validation dataset. The results are in *train_val/val_predicted/val_predicted.xlsx*
- Run *test.py* to predict on validation dataset. The results are in *test/Predicted/Predicted.xlsx*
- Run *validation_evaluation.py* to calculate evaluation metrics.

NOTE: Running Locations of README.pdf and Prediction.xlsx in GITHUB repository:

- *README.pdf* can be found in base directory.
- Test predictions can be found in *test/Prediction/Prediction.xlsx*