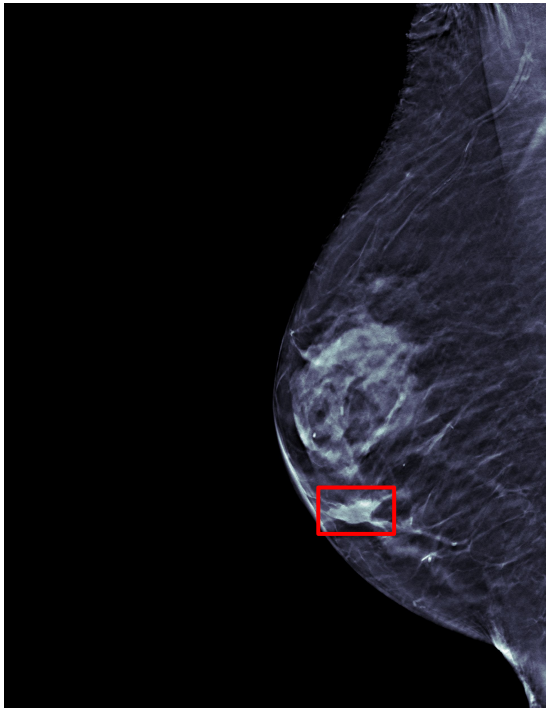
Object Detection on Digital Breast Tomosynthesis Images with YOLO

Joy Fu
05/11/2021

Background

- Digital Breast Tomosynthesis (DBT)
  - More accurate
  - More time-consuming
- Automated methods
  - Computer-aided detection
  - Computer-aided diagnosis
- Objective
  - Object detection on DBT
  - Lower specificity and precision (while not resulting in a significant increase in sensitivity)



Ioannis Sechopoulos, Seminars in Cancer Biology, https://doi.org/10.1016/j.semcancer.2020.06.002

Method: You Only Look Once (YOLO)

- Object detection
  - Finding all the objects in an image and drawing the so-called bounding boxes around them
  - Two general groups
    - Algorithms based on classification: RCNN, Fast-RCNN, RetinaNet, etc.
    - Algorithms based on regression: YOLO, Single Shot Multibox Detector (SSD), etc.
- Regression-based object detection
  - Faster, but less accurate in general
- YOLO
  - Reframe object detection as **a single regression problem (convolution network)**
  - **You Only Look Once** at an image to predict what objects are present and where they are
  - Uses features from the **entire image** to predict
  - Predicts **all bounding boxes** across **all classes** for an image **simultaneously**

J. Redmon, S. Divvala, R. Girshick, and A. Farhadi, “You Only Look Once: Unified, Real-Time Object Detection,” arXiv:1506.02640 [cs], May 2016, Accessed: May 10, 2021. [Online]. Available: http://arxiv.org/abs/1506.02640

Method: You Only Look Once (YOLO)

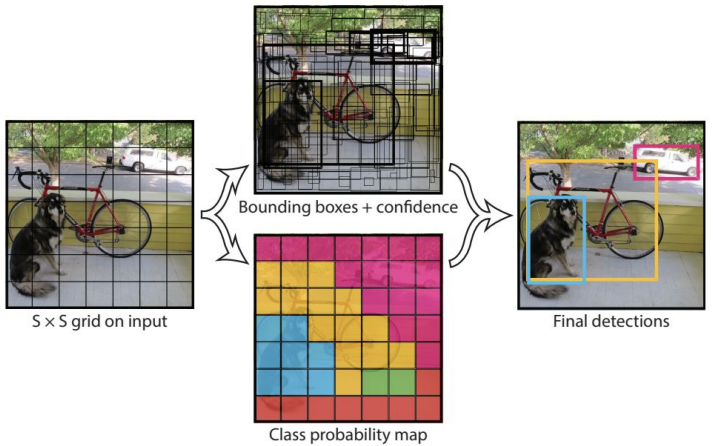
- Unified detection
  - Divides the input image into an S × S grid
  - Each grid cell predicts B bounding boxes and confidence scores for those boxes

$$\Pr(\text{Object}) * \text{IOU}_{\text{pred}}^{\text{truth}}$$

(IOU: intersection over union)

  - Each bounding box consists of 5 predictions
    - **(x, y)**: center of the box
    - **width, height**
    - **confidence**
  - Each grid cell also predicts C conditional class probabilities

$$\Pr(\text{Class}_i|\text{Object})$$



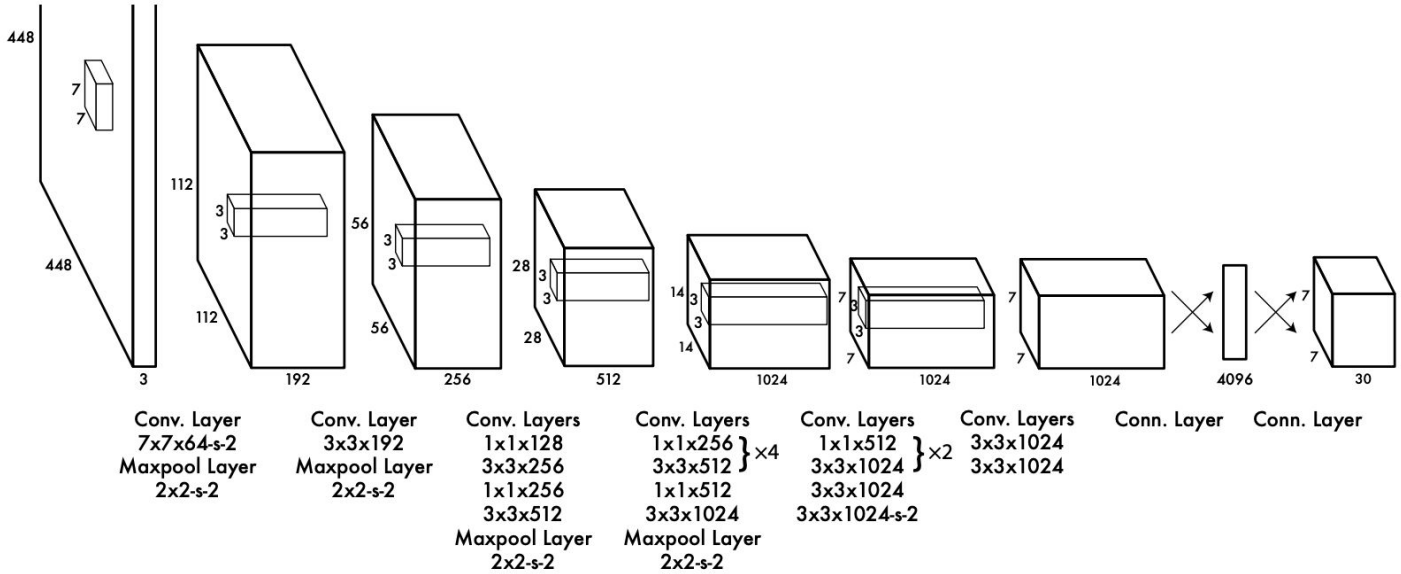
  - At test time: class-specific confidence scores for each box

$$\Pr(\text{Class}_i|\text{Object}) * \Pr(\text{Object}) * \text{IOU}_{\text{pred}}^{\text{truth}} = \Pr(\text{Class}_i) * \text{IOU}_{\text{pred}}^{\text{truth}}$$

J. Redmon, S. Divvala, R. Girshick, and A. Farhadi, “You Only Look Once: Unified, Real-Time Object Detection,” arXiv:1506.02640 [cs], May 2016, Accessed: May 10, 2021. [Online]. Available: http://arxiv.org/abs/1506.02640

Method: You Only Look Once (YOLO)

- Full network architecture
  - 24 convolutional layers followed by 2 fully connected layers
  - 1 × 1 reduction layers
  - Optimize for sum-squared error



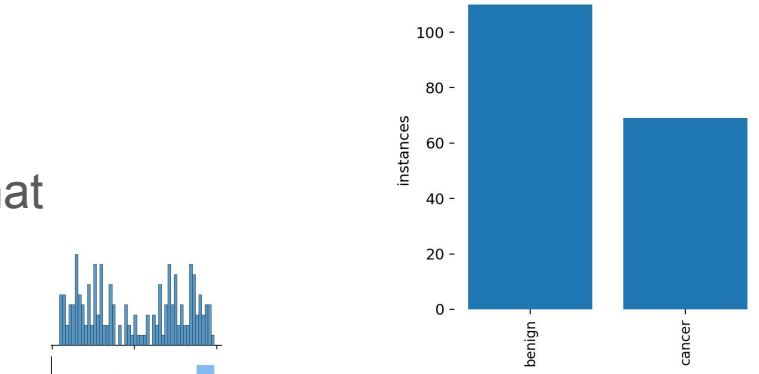
Method: You Only Look Once (YOLO)

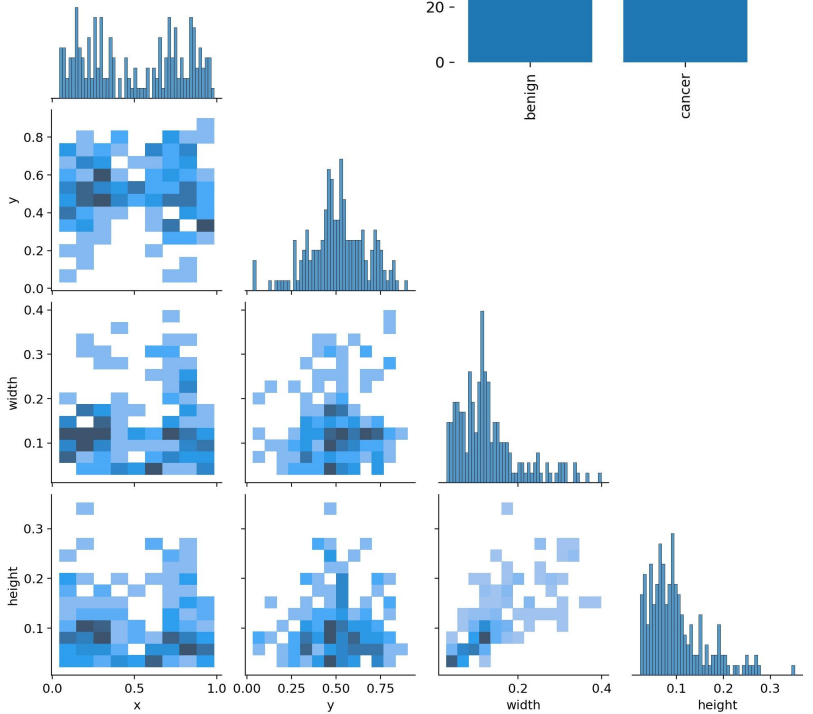
- Pros
  - Extremely fast
  - Reasons globally about the image when making predictions
  - Learns generalizable representations of objects
- Cons
  - Lags behind in accuracy
  - Struggles with small objects

Method: Workflow

```
images
  train
  val
labels
  train
  val
```

- Built the dataset and convert to YOLO format
- Two configuration files
  - Specify architecture to be used
  - Number of classes
- Train-test (0.9-0.1) split
  - N_{train} = 179, N_{test} = 21
- Model training
  - Labels: benign, cancer, (background)
- Model evaluation
  - Multiple metrics



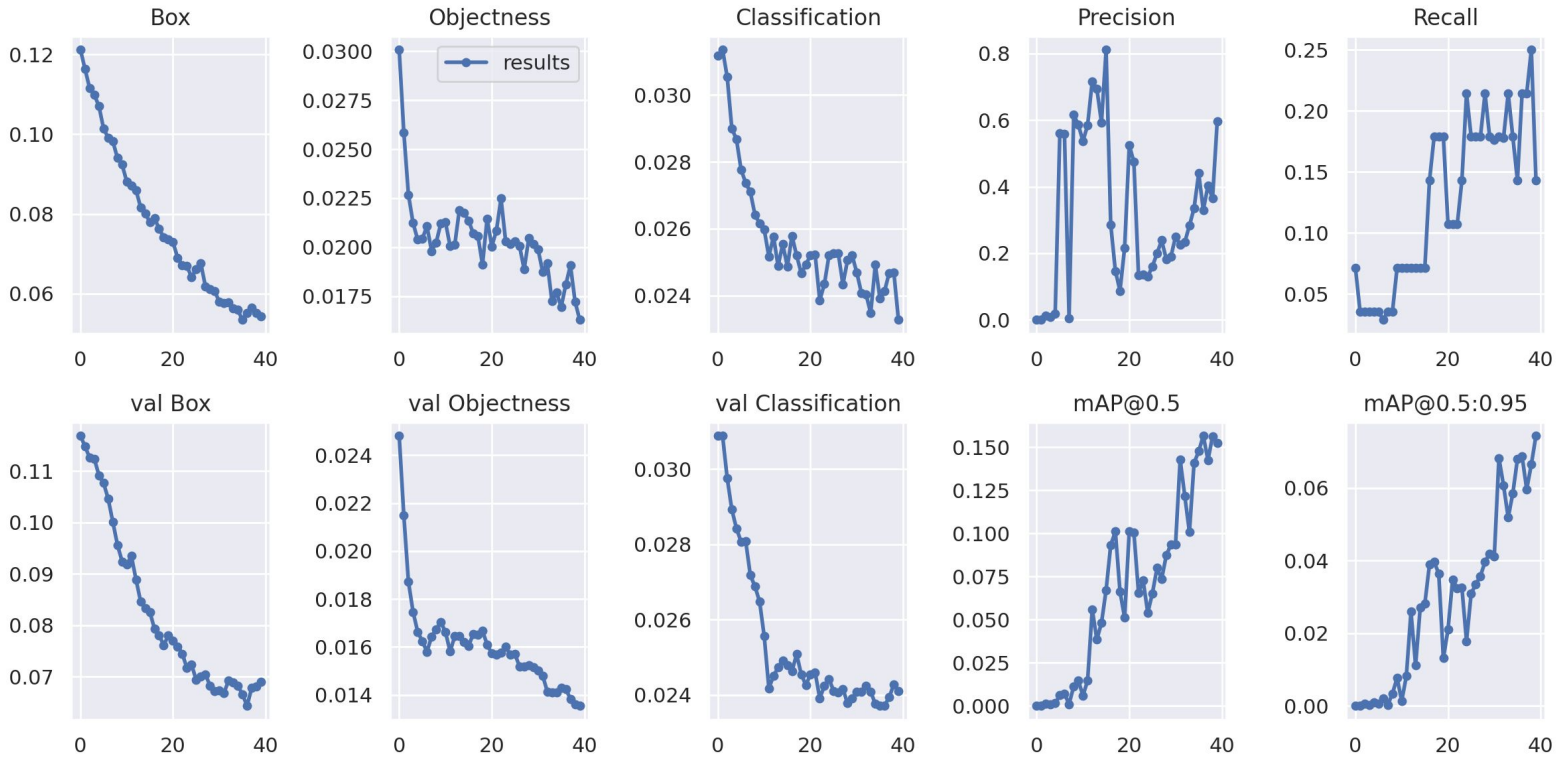


0 0.26064359295031164 0.12173736807287297 0.03206412825651302 0.026048026048026047

Preliminary results

Model: YOLOv5s (the smallest YOLOv5)

- Using their default architecture and parameters
- Image size 640
- Batch 16
- Epochs 40
- ~ 2hrs to train

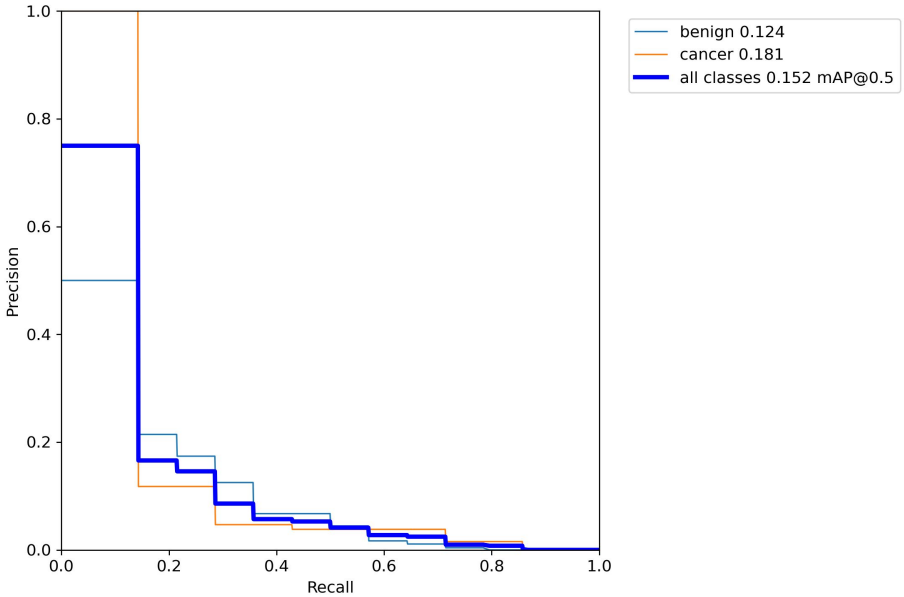


Preliminary results

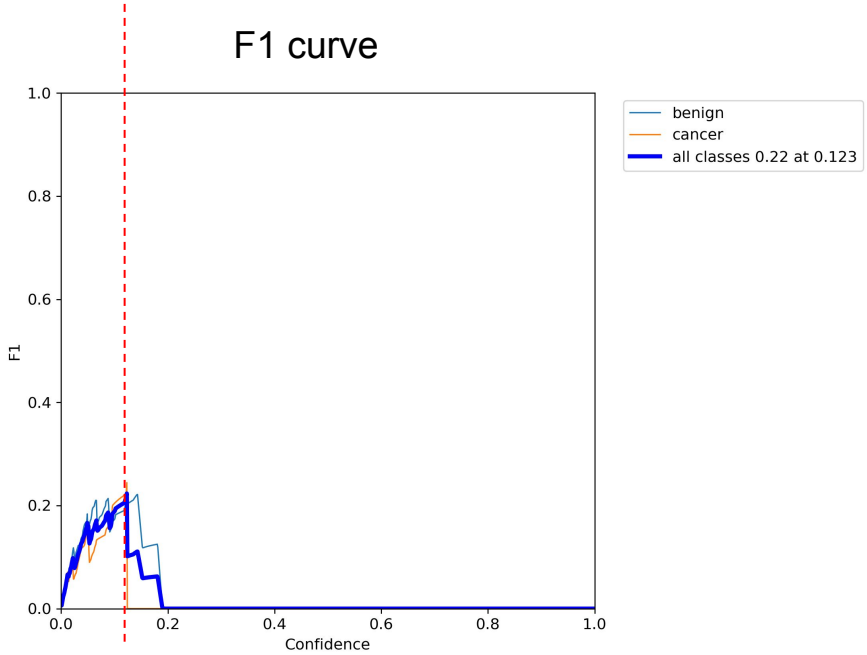
Model: YOLOv5s (the smallest YOLOv5)

- Using their default architecture and parameters
- Image size 640
- Batch 16
- Epochs 40
- ~ 2hrs to train

Precision-recall curve



F1 curve

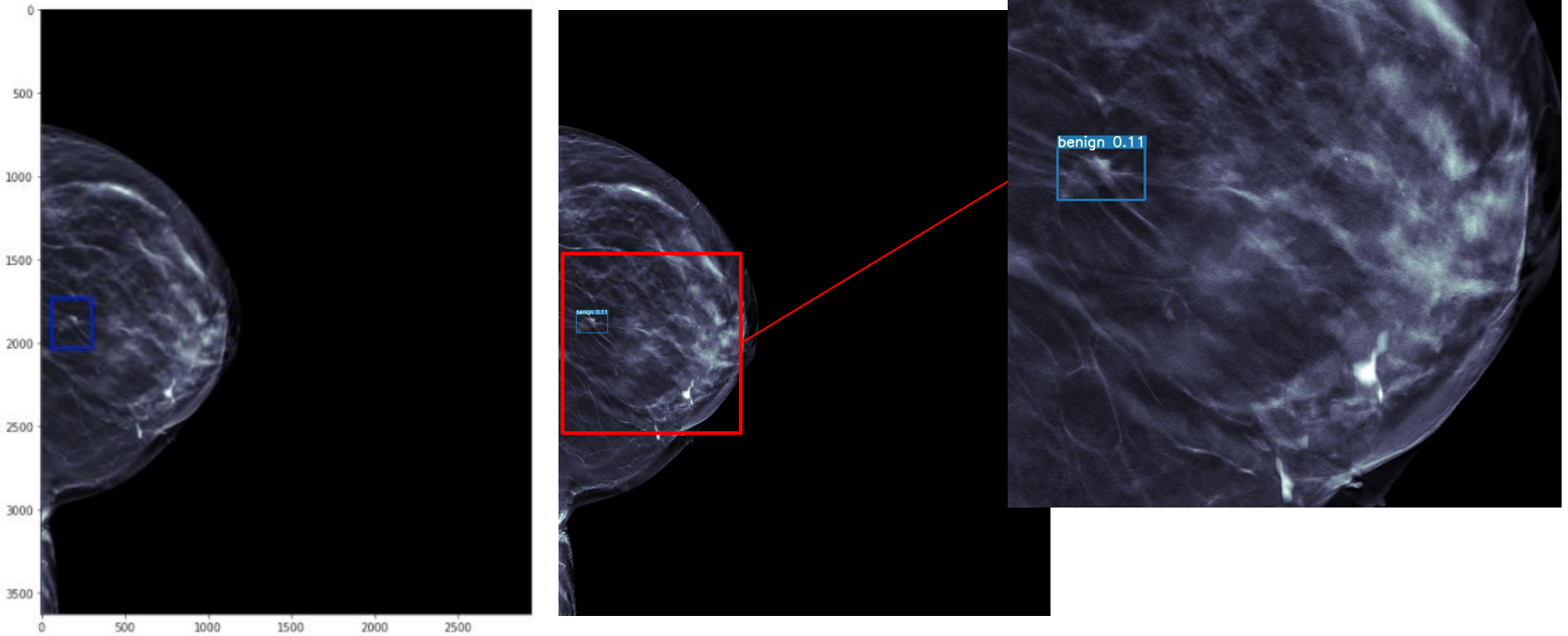


Preliminary results

Predict: YOLOv5s (the smallest YOLOv5)

- Using best weights trained
- Confidence 0.1

True label: benign

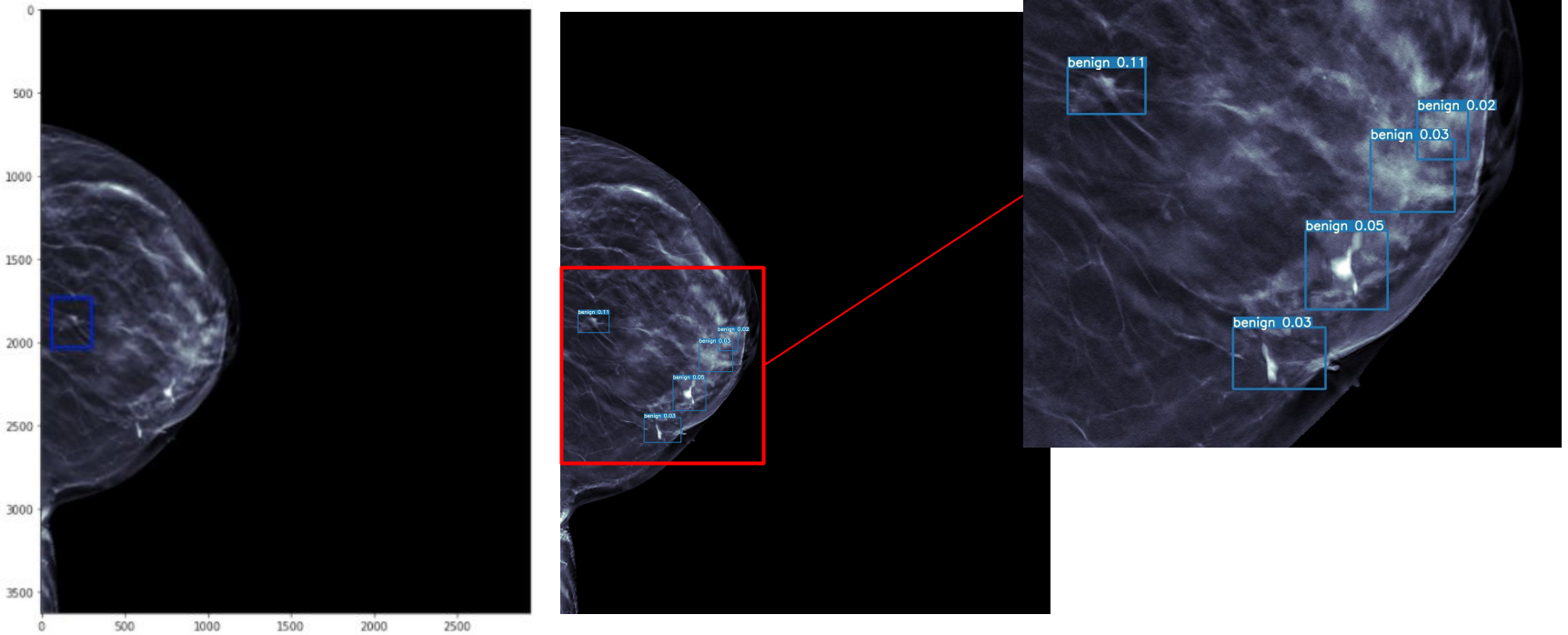


Preliminary results

Predict: YOLOv5s (the smallest YOLOv5)

- Using best weights trained
- Confidence 0.02

True label: benign

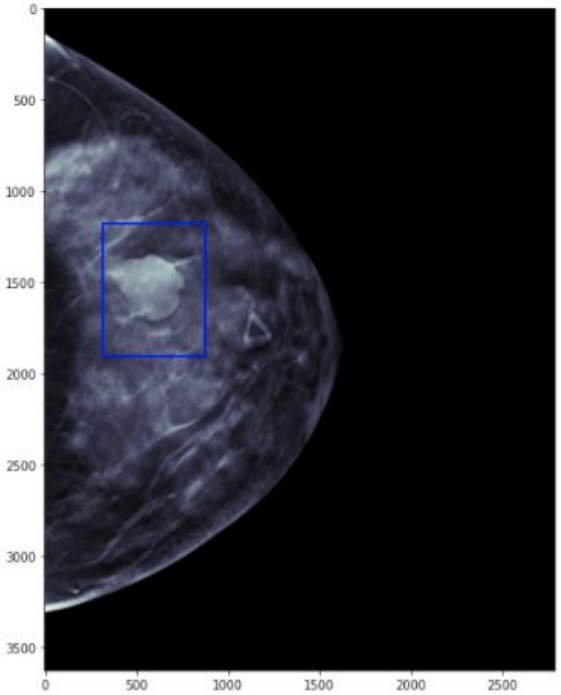


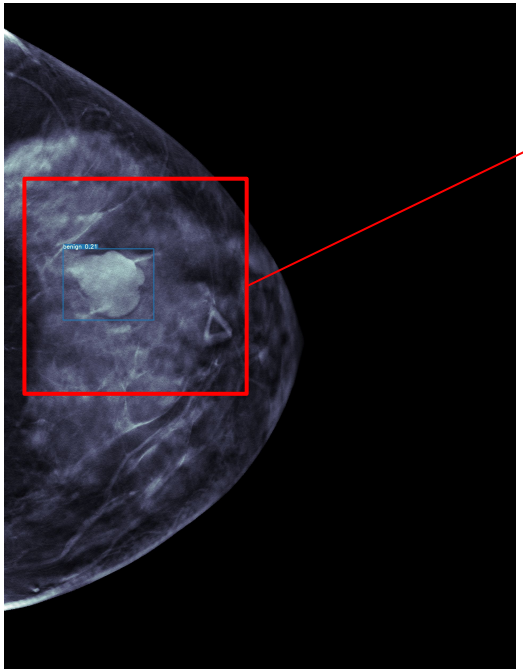
Preliminary results

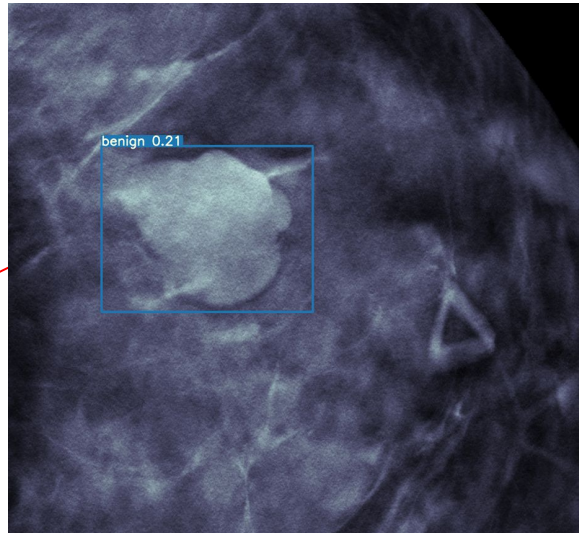
Predict: YOLOv5s (the smallest YOLOv5)

- Using best weights trained
- Confidence 0.02

True label: benign





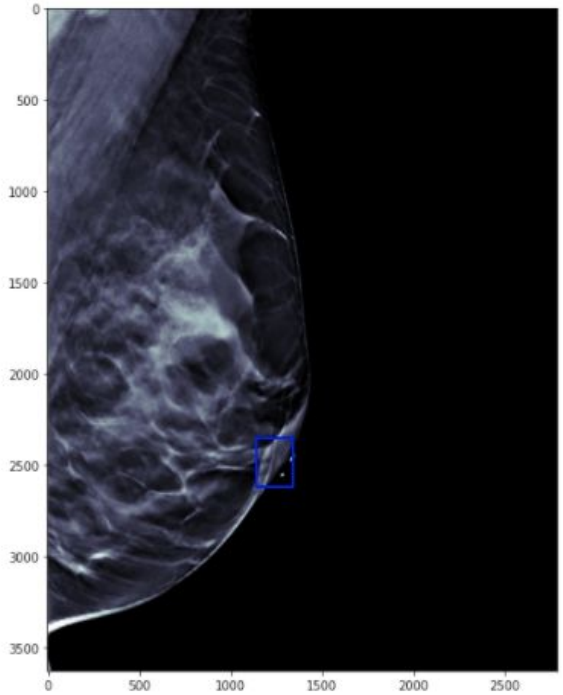


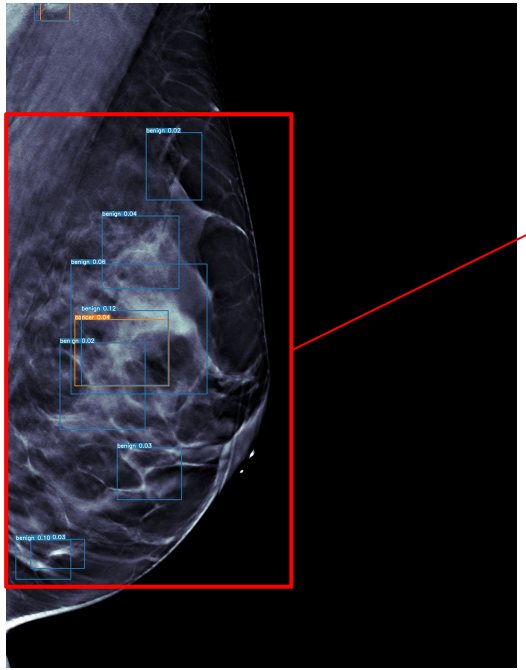
Predict: YOLOv5s (the smallest YOLOv5)

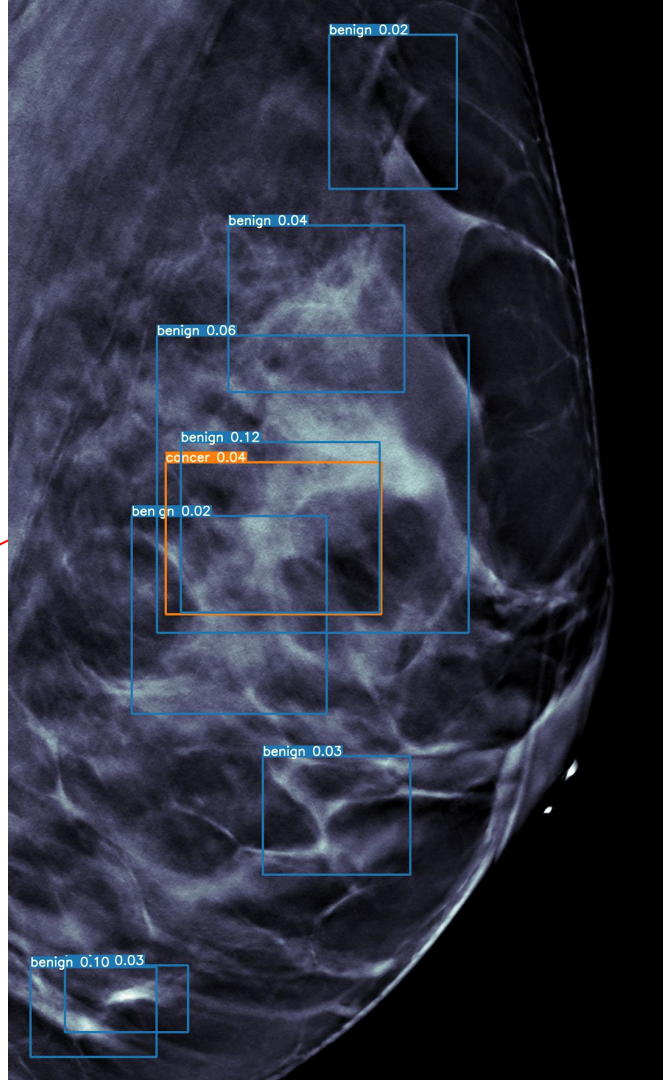
- Using best weights trained
- Confidence 0.02

Preliminary results

True label: benign







Reflections and future works

- The current model is not ideal for ROI detection and classification
- The current train-test split doesn’t include cancer in the test set
- Try larger YOLOv5 frameworks