Deep Learning in Computer Vision
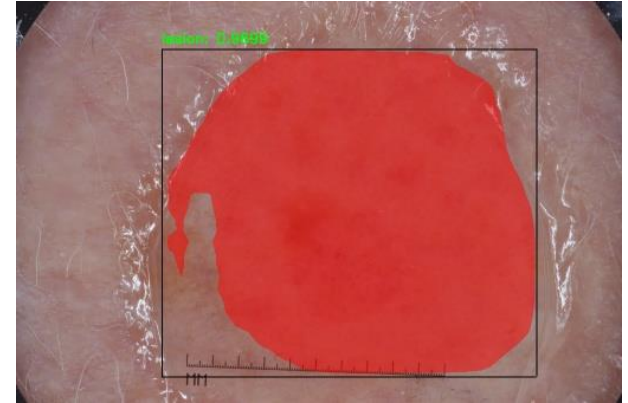
Mask R-CNN
for
Instance Segmentation

Instructor: Dr. Mehrandezh
Student: Marzieh Zamani

Mask R-CNN for Instance Segmentation

▸ Applications:

◦ Segmenting and masking skin lesions

◦ Segmenting and masking round pills

# From Object Detection to Segmentation

- Object detection:
  - Detecting objects in an image
  - Output: bounding box coordinates and class labels

- Segmentation:
  - Partitioning an image into parts
  - No understanding of the resulting segmented image parts
  - Output: pixel-wise mask for all object classes

Semantic Segmentation vs. Instance Segmentation

▸ **Semantic segmentation** treats multiple objects of the same class as a single entity;

▸ **Instance segmentation** treats multiple objects of the same class as distinct individual objects (or instances).
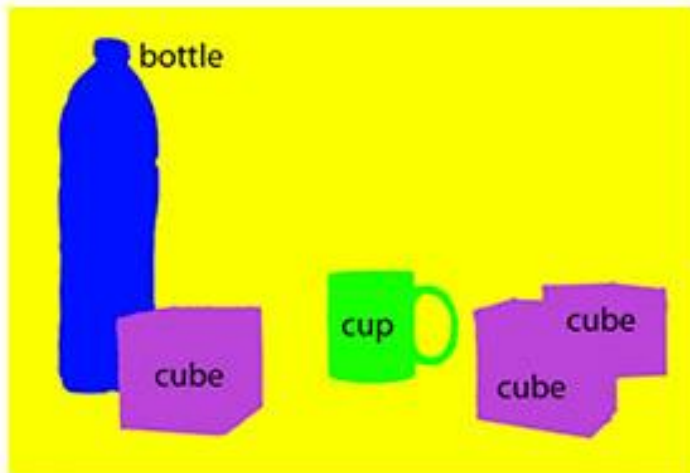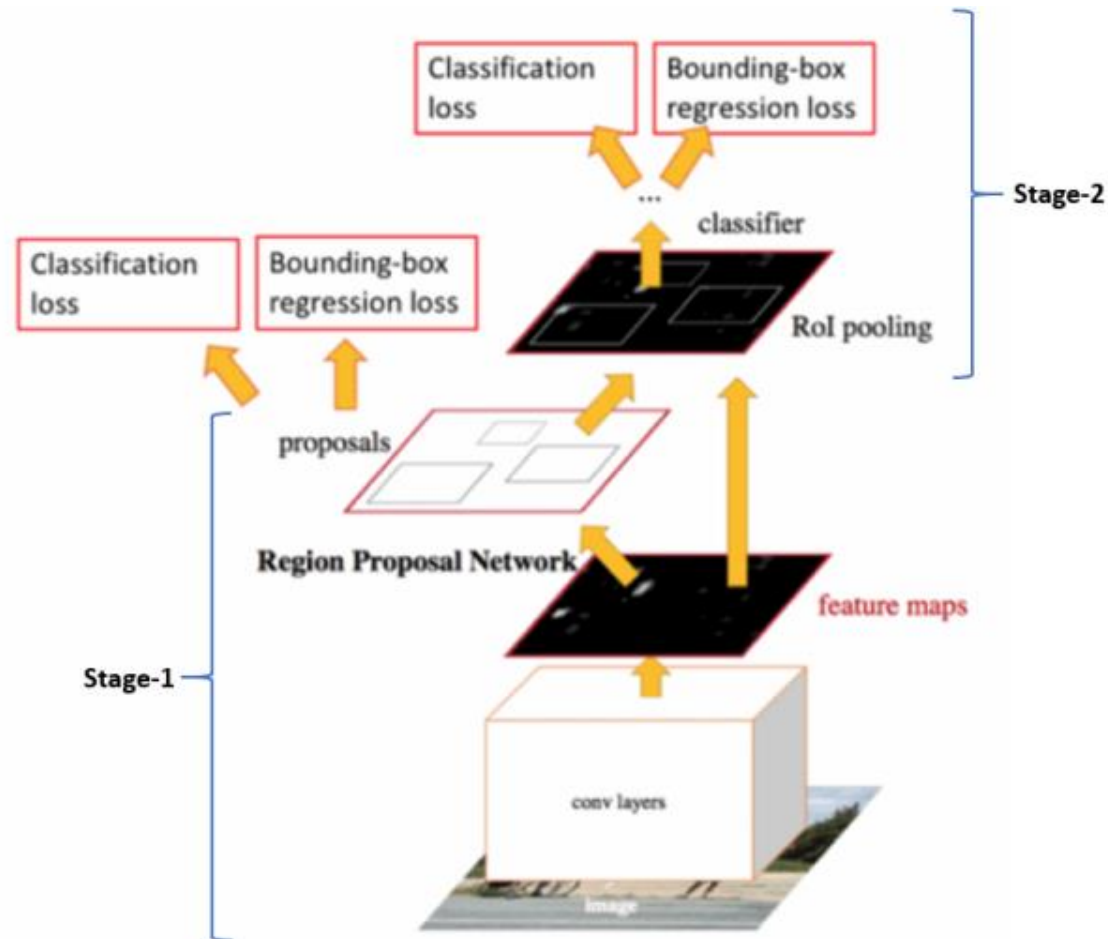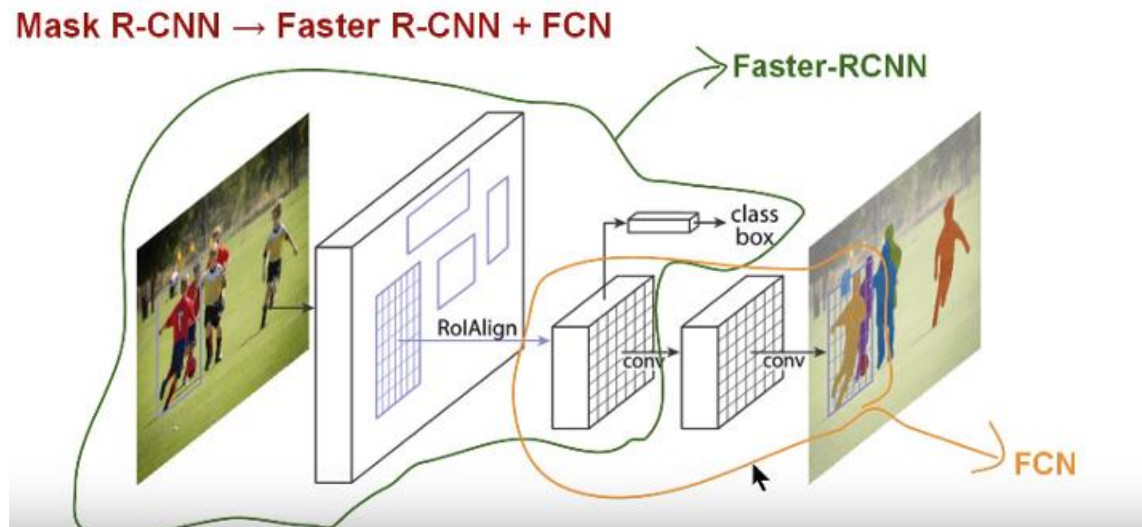

(c) Semantic segmentation
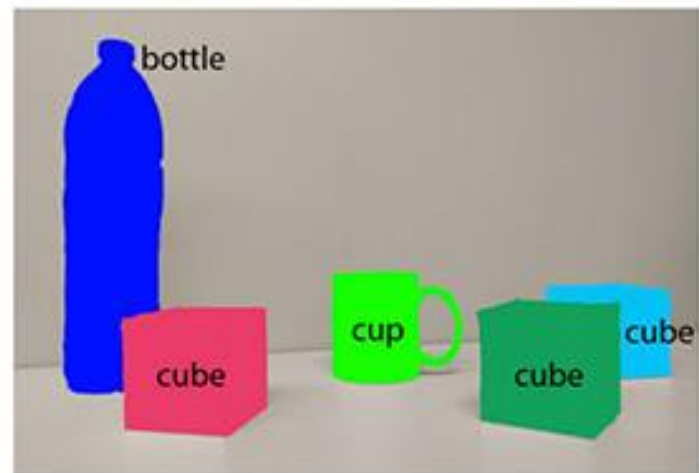
(d) Instance segmentation

Faster R-CNN

▸ Stage 1 (RPN) => Proposal ROI
▸ Stage 2 => Object Bounding-Box & Class
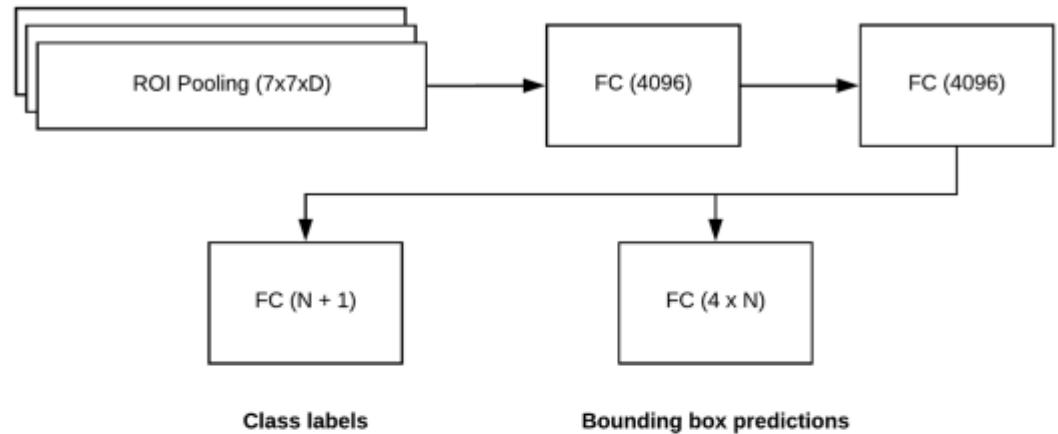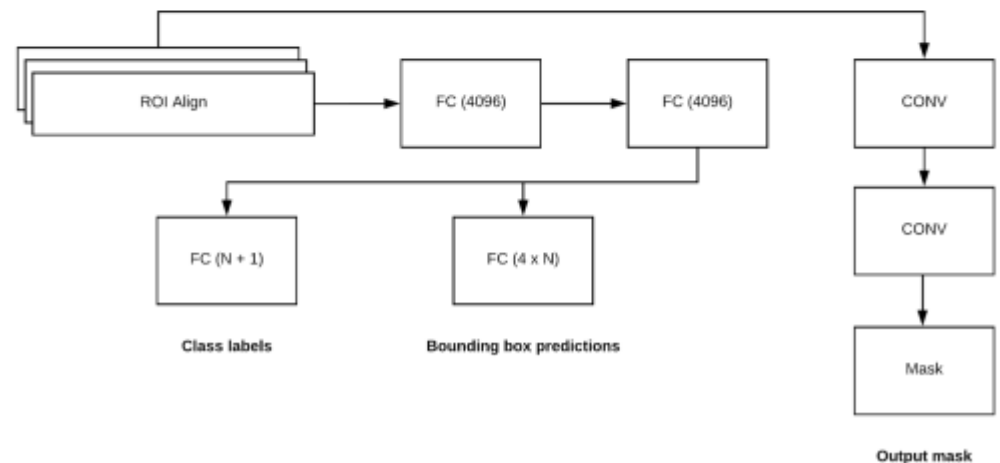
Mask R-CNN=Faster R-CNN + FCN

Mask R-CNN=Faster R-CNN + FCN

- Mask R-CNN=>
  - Class labels
  - Object bounding box
  - Pixel-wise masks

Faster R-CNN

ROI Pooling (7x7xD) → FC (4096) → FC (4096)

FC (N + 1) — Class labels

FC (4 x N) — Bounding box predictions

Mask R-CNN

ROI Align → FC (4096) → FC (4096)

FC (N + 1) — Class labels

FC (4 x N) — Bounding box predictions

CONV → CONV → Mask

Output mask

Mask RCNN for Segmenting Skin Lesion

▸ Melanoma is also the deadliest form of skin cancer

▸ Early detection of melanoma => survival rates of 95%.

Mask RCNN for Segmenting Skin Lesion

▸ Dataset: The International Skin Imaging Collaboration (ISIC) Skin Lesion (2018)

◦ Training_Input: 2594 images of skin lesions

◦ Training_GroundTruth: 2594 corresponding masks
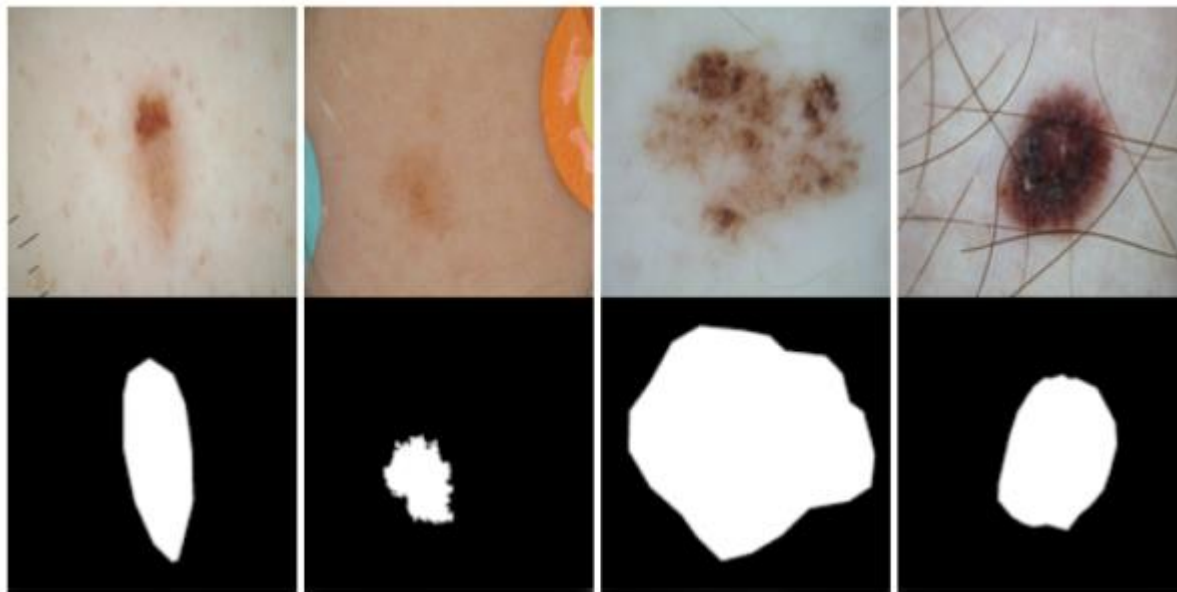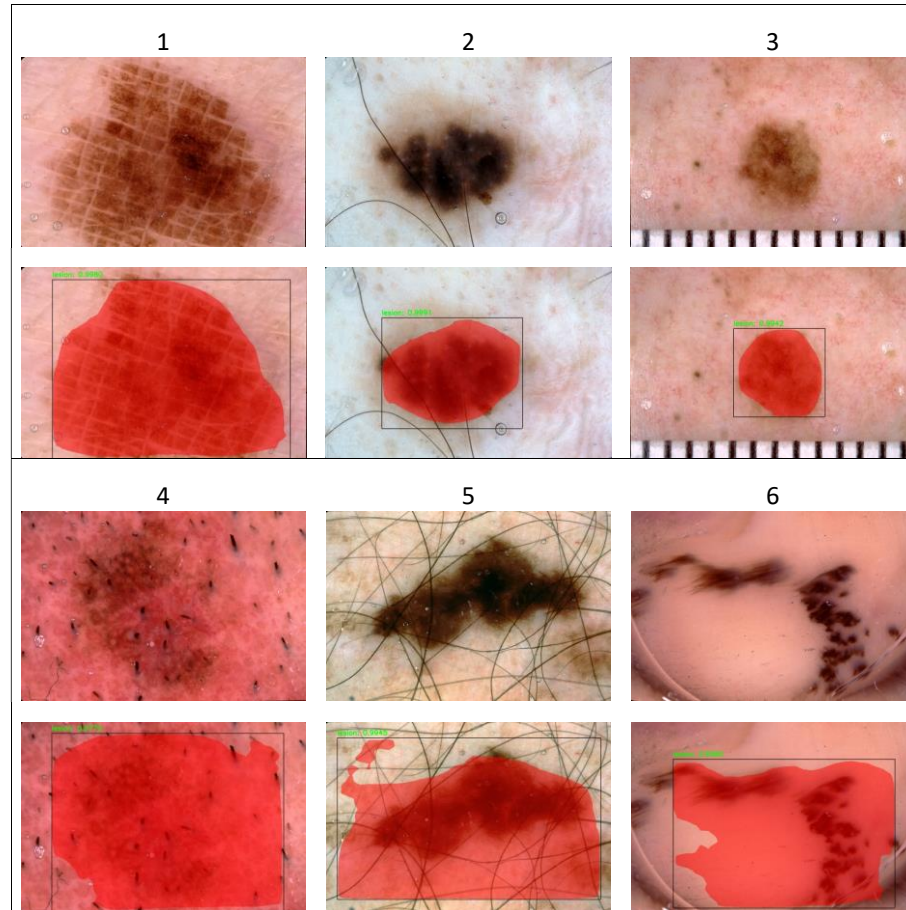
Mask RCNN for Segmenting Skin Lesion

▸ Dataset: The ISIC Skin Lesion (2018)

Usage	Number of Images	Source Dataset
Training	80% x 2000 = 1600	ISIC Skin Lesion Dataset (2018)
Evaluation	20% x 2000 = 400	ISIC Skin Lesion Dataset (2018)
Prediction	594	ISIC Skin Lesion Dataset (2018)

Training Mask RCNN for Segmenting Skin Lesion

- Initial weights from mask_rcnn_coco.h5

- Training Head (20 epochs)

- Training all layers (20 epochs)
  - Unsatisfactory due to memory shortage

Mask RCNN for Segmenting Skin Lesion => Results

# Mask RCNN for Segmenting Skin Lesion => Remarks

▶ **Remarks from results:**

◦ These 6 images of skin lesions and their corresponding masks have been selected form prediction dataset in a way to represent the best and worst performances of Mask RCNN

◦ Results 1 to 4 show the significant ability of Mask RCNN network for segmenting skin lesions;

◦ On the other hand, results 5, and 6 show that the mask precision would suffer for challenging skin lesions (e.g. when highly scattered, covered by hair or etc.)

Mask RCNN
for Segmentation of Round Pills

▸ Over 3.3 million injuries and deaths / year due to the incorrect pill being taken.

▸ => the National Library of Medicine's (NLM) pill identification challenge

Mask RCNN for Segmenting Round Pills

▸ Dataset: The National Library of Medicine's (NLM) 2016 pill identification challenge
  ◦ Consumer–quality Images
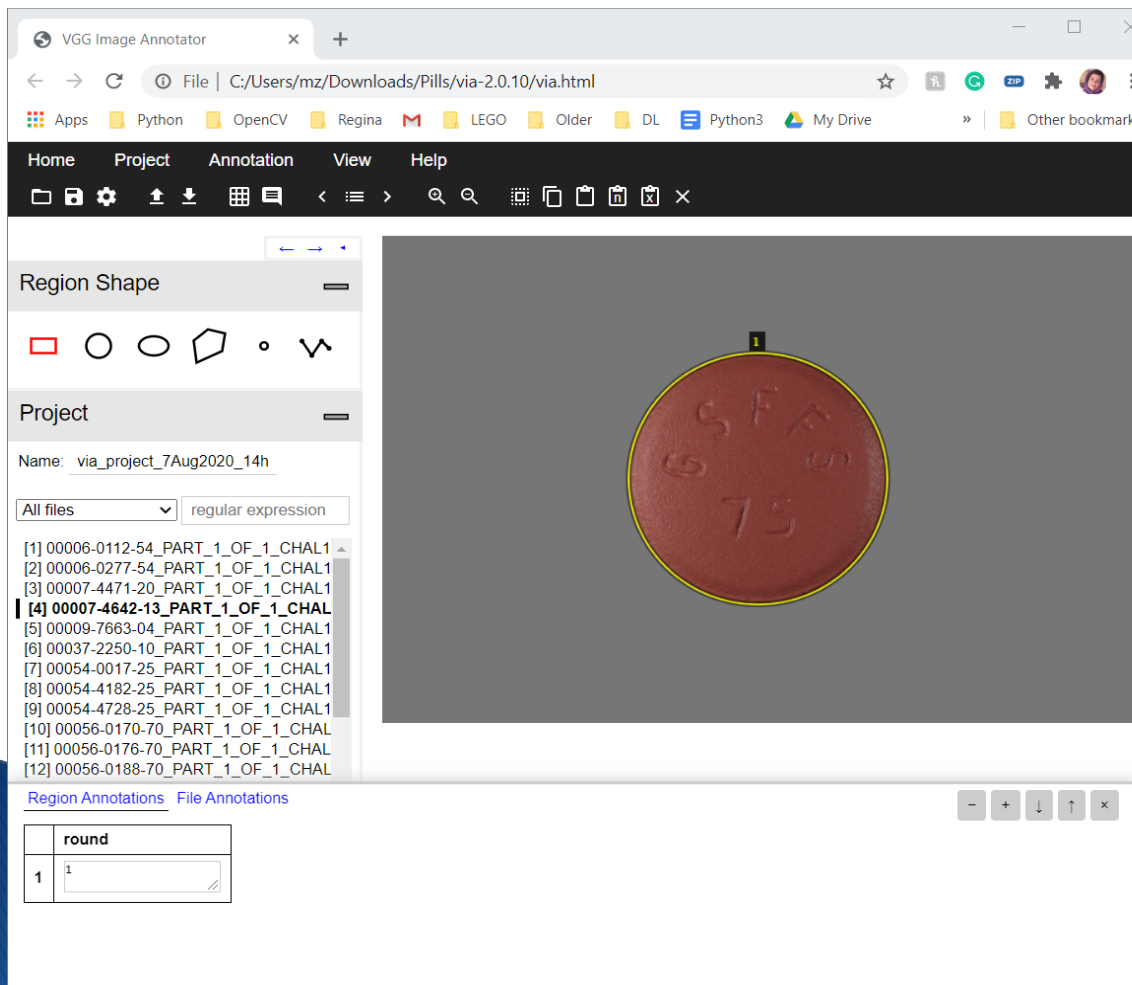 ◦ Reference Images
  ◦ Ground Truth Table
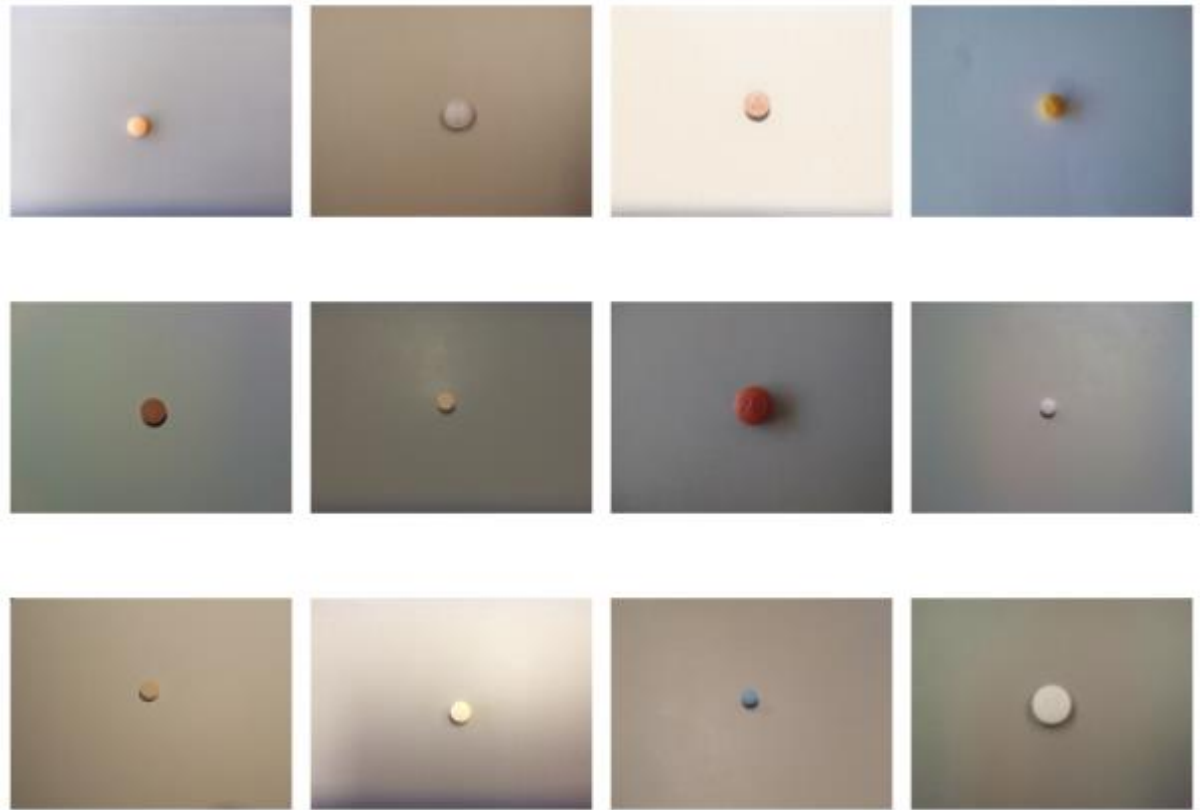
Mask RCNN for Segmenting Round Pills

▸ Used Datasets

Usage	Number of Images	Source Dataset	Sample Image
Training	75% x 60 = 45	NLM (2016) pill dataset- Reference Images	
Evaluation	25% x 60 = 15	NLM (2016) pill dataset- Reference Images	
Prediction	100	NLM (2016) pill dataset- Consumer-quality Images	
Prediction	20	Google Images Search Results	

Mask RCNN for Segmenting Round Pills

▸ Annotating images in VGG Image Annotator (VIA)

Training Mask RCNN for Segmenting Round Pills

- Initial weights from mask_rcnn_coco.h5

- Training Head (10 epochs)

- Training all layers (10 epochs)
  - Unsatisfactory due to memory shortage

Mask RCNN for Segmenting Round Pills => Predicting Consumer-quality Images
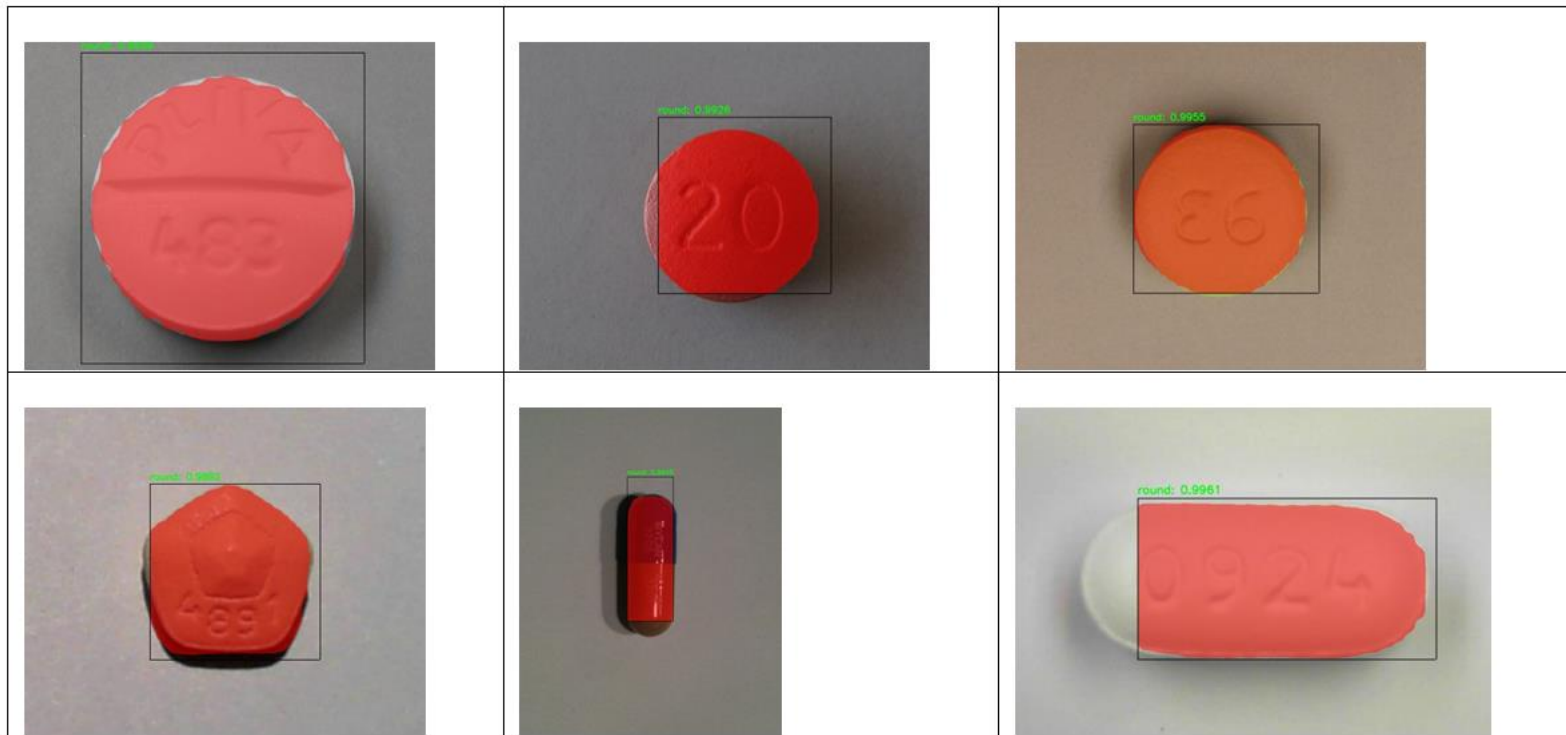
Mask RCNN for Segmenting Round Pills => Predicting Consumer-quality Images

▸ **Remarks from results:**

◦ Mask RCNN is successfully segmenting and masking pills from Consumer-quality Images;

◦ However, the problem is that it mistakenly detects non-round pills as round.

Mask RCNN for Segmenting Round Pills => Predicting Images from Internet

Mask RCNN for Segmenting Round Pills => Predicting Images from Internet

▸ **Remarks from results:**

◦ Mask RCNN precision is significantly lower when applied to images of pills in hands;

◦ For 30% of images, Mask RCNN is able to segment all round pills; However, for the remaining 70% (such as the result 6) segmentation is not satisfactory.

Required software and libraries for implementation of Mask RCNN 2.1

Hardware/Software/Library	Version	Useful Link
Operating System	Windows 10	
GPU	GEFORCE GTX 1660 Ti	
Visual Studio	2017 (RTW and all updates)	https://docs.nvidia.com/cuda/archive/10.0/cuda-installation-guide-microsoft-windows/index.html
NVIDIA CUDA	10.0	https://www.tensorflow.org/install/source_windows#gpu
NVIDIA cuDNN	7.4	https://www.tensorflow.org/install/source_windows#gpu
Python	3.6	
tensorflow_gpu	1.13.1	
Keras	2.2.2	
opencv-python		
opencv-contrib-python		
imgaug		
ipython		
imutils		
Mask RCNN	2.1	https://github.com/matterport/Mask_RCNN

References

▸ Deep Learning for Computer Vision with Python (Bonus Bundle) by Dr. Adrian Rosebrock

▸ https://developers.arcgis.com/python/guide/how-maskrcnn-works/

▸ https://towardsdatascience.com/semantic-segmentation-popular-architectures-dff0a75f39d0

▸ https://challenge.isic-archive.com/data#2018

▸ https://github.com/matterport/Mask_RCNN

▸ https://pir.nlm.nih.gov/challenge/submission.html

Thanks to

Behnam, and Amir
and specially Niloufar for sharing her satisfactory configuration with me.

Thanks for your attention

>> Questions are welcome