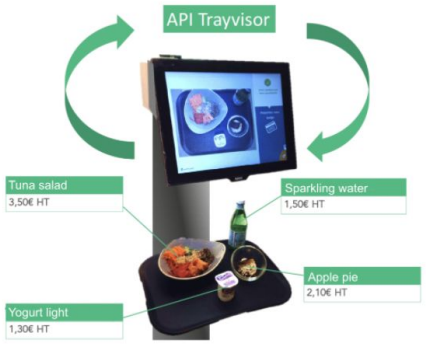
Deep learning based food instance Recognition

Classification - Segmentation Deep Learning Model

Use case: Self Service - Virtual Cashier - Calories Estimation

Thuc Dang - Tommy Nguyen - Minh Nhat
Moscow - Melbourne - Paris
08/05/2022



Content

- Introduction
- Methodology
- Result/ Discussion
- Conclusion

Introduction

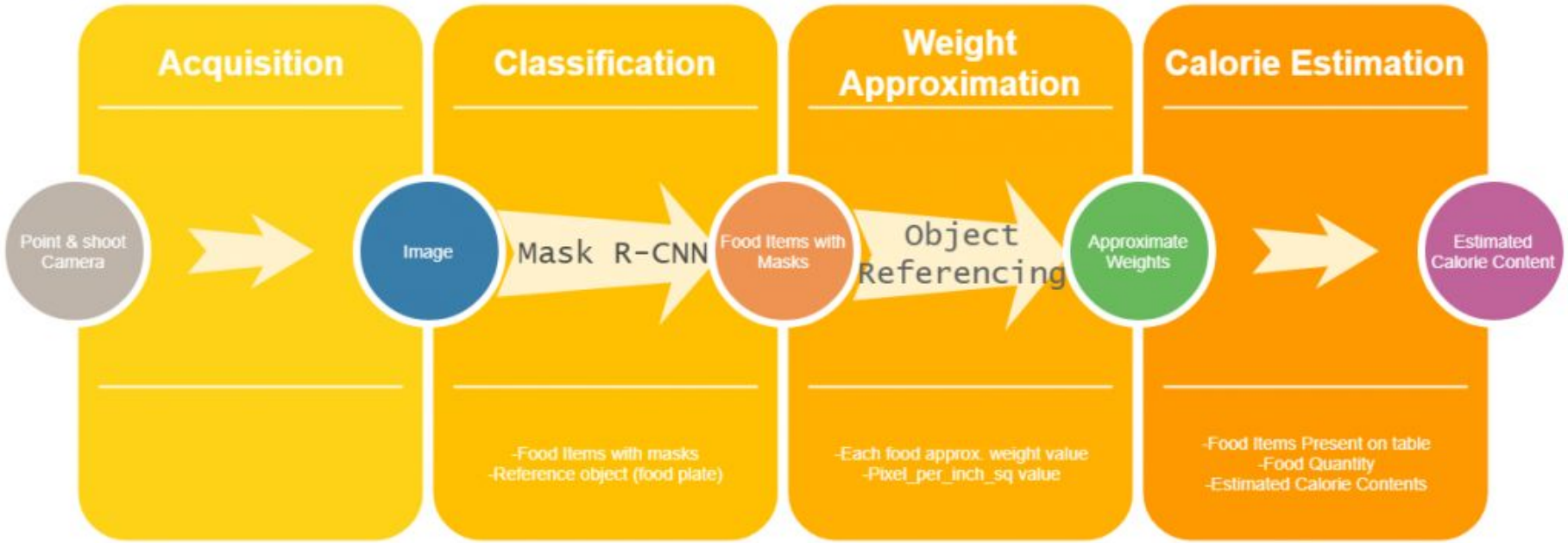
- **Goal** : Experience real-life AI project from (A-Z) scratch
- According to WHO 20% of deaths worldwide are attributable to an unhealthy diet nowadays.
- 39% of adults aged 18 years and over were overweight in 2016, and 13% were obese.

VS

SOTA Food Recognition

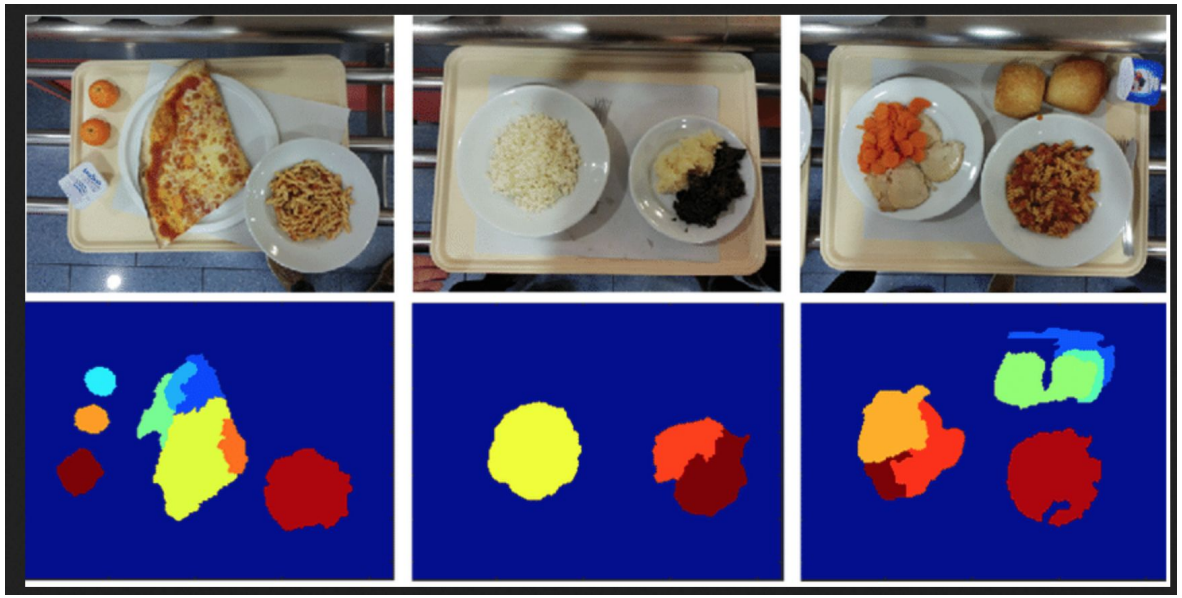
Image-Based Calorie Estimation using Deep Learning

Kitamura, Aizawa et al 2009
Aguilar et al., 2017.



Methodology - dataset

image crawling



Public dataset: Unimib2016

- Set of 1027 canteen food tray captured from top view
- 3616 food instances
- 73 food classes
- Polygonal boundaries
- 79% accuracy using CNN

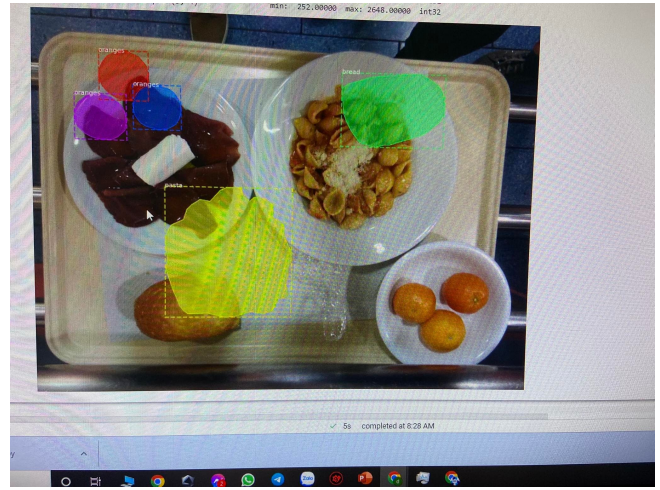
Data Preprocessing

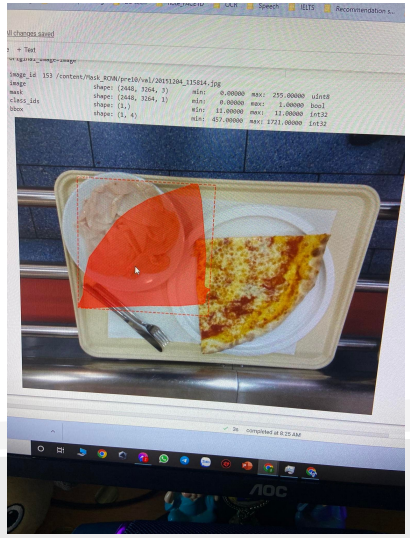
image labelling

Issues regarding to data treatment.

extension .mat -> .json

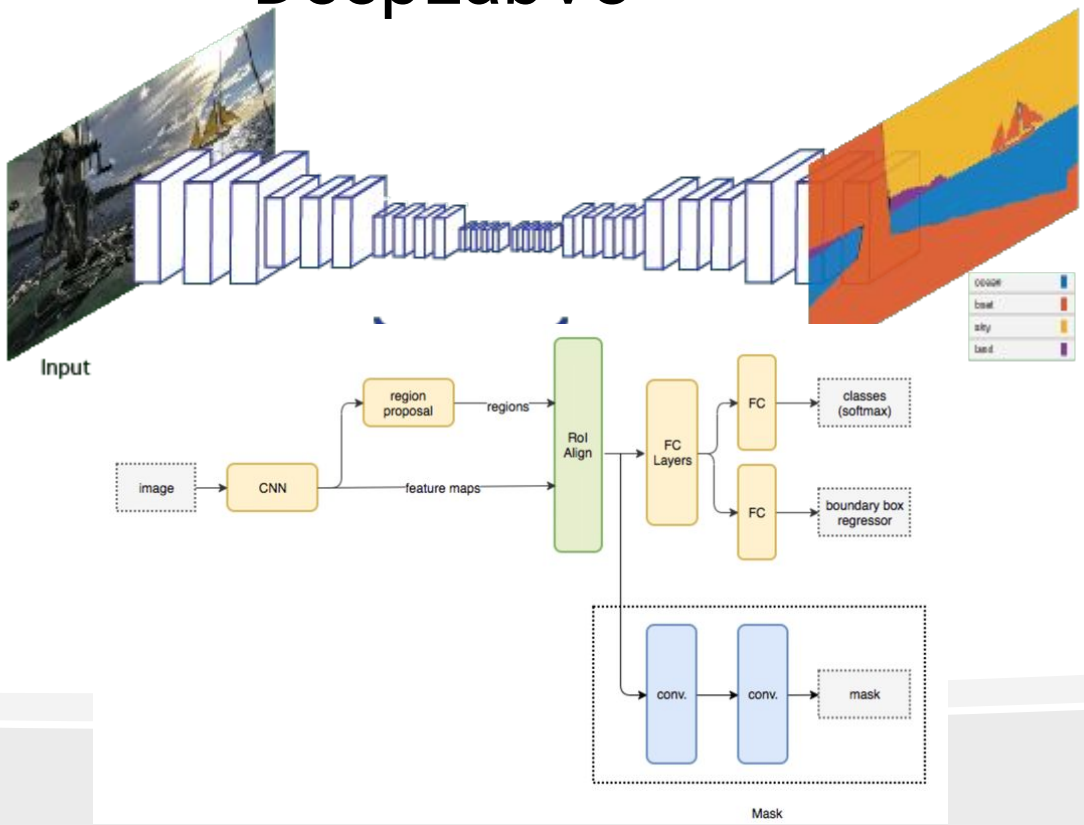
image rotation





Methodology - Model Training with Colab Pro

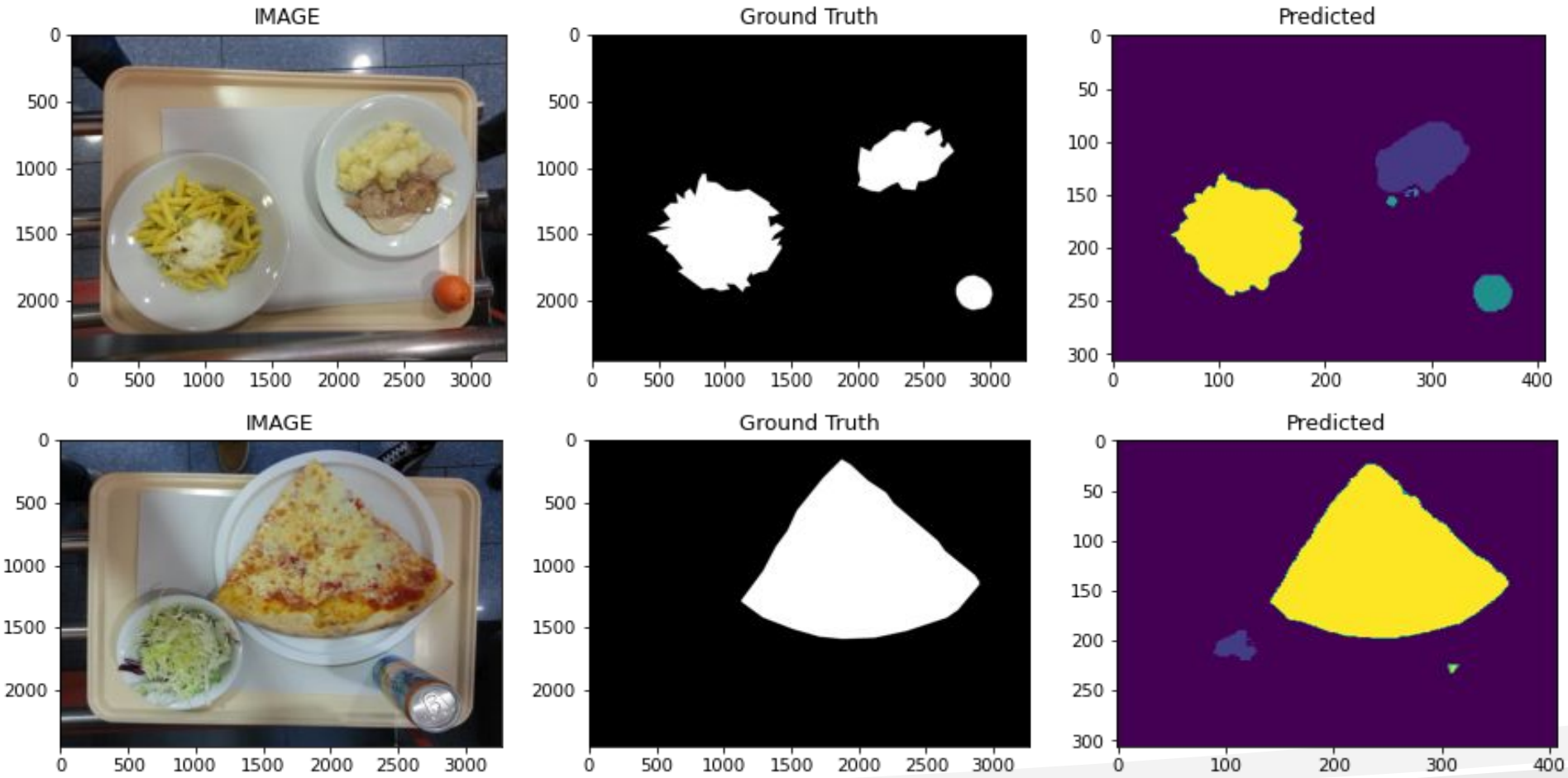
- Unet
- Mask RCNN
- DeepLabV3



Training detail for
Mask RCNN
SGD ,
optimizer learning rate
batch size
epoch
pretrained models

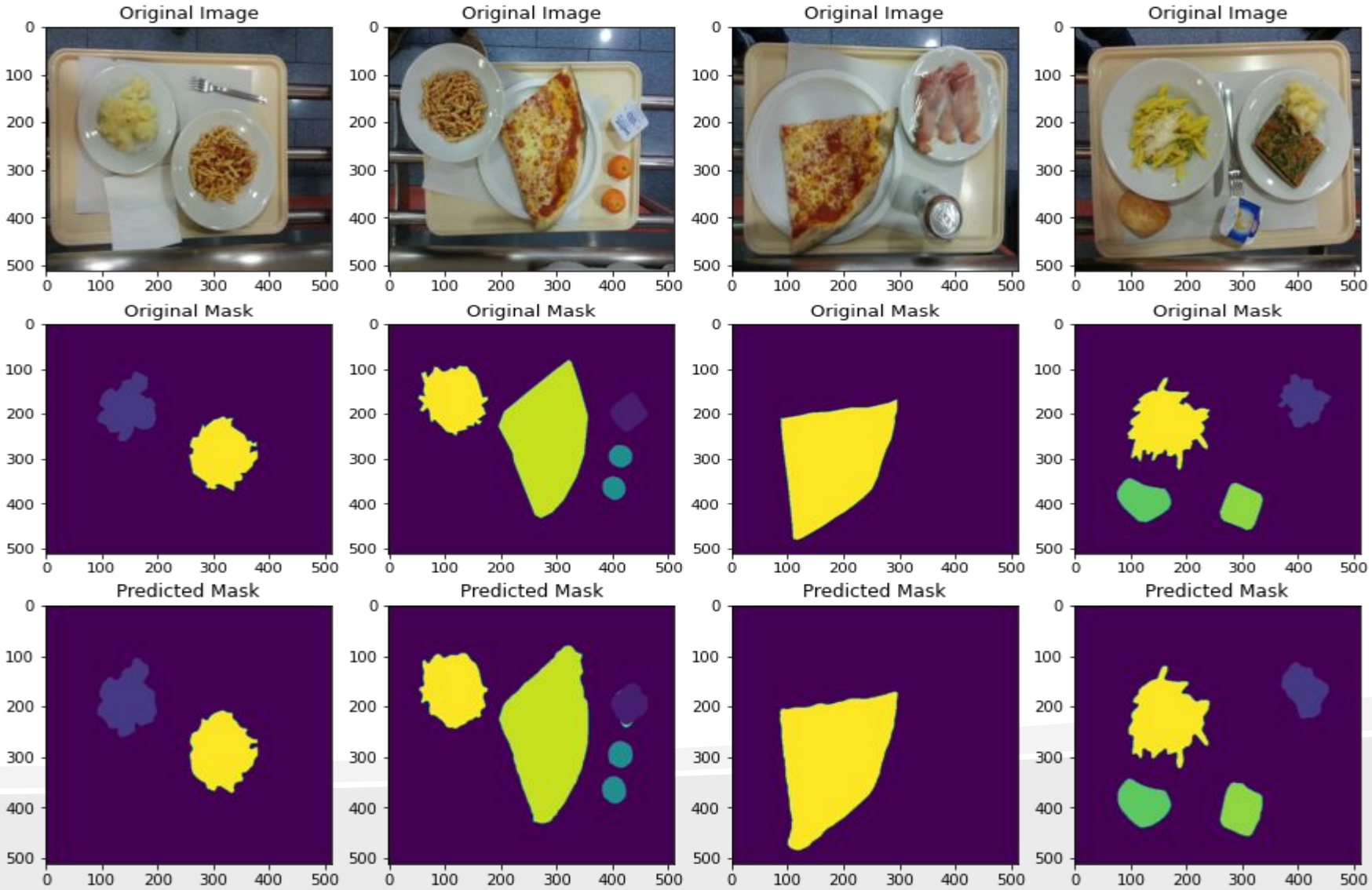
Evaluation metric
COCO
mAP

Results / Discussion (**Unet**from : Loss 0.05) 12hours

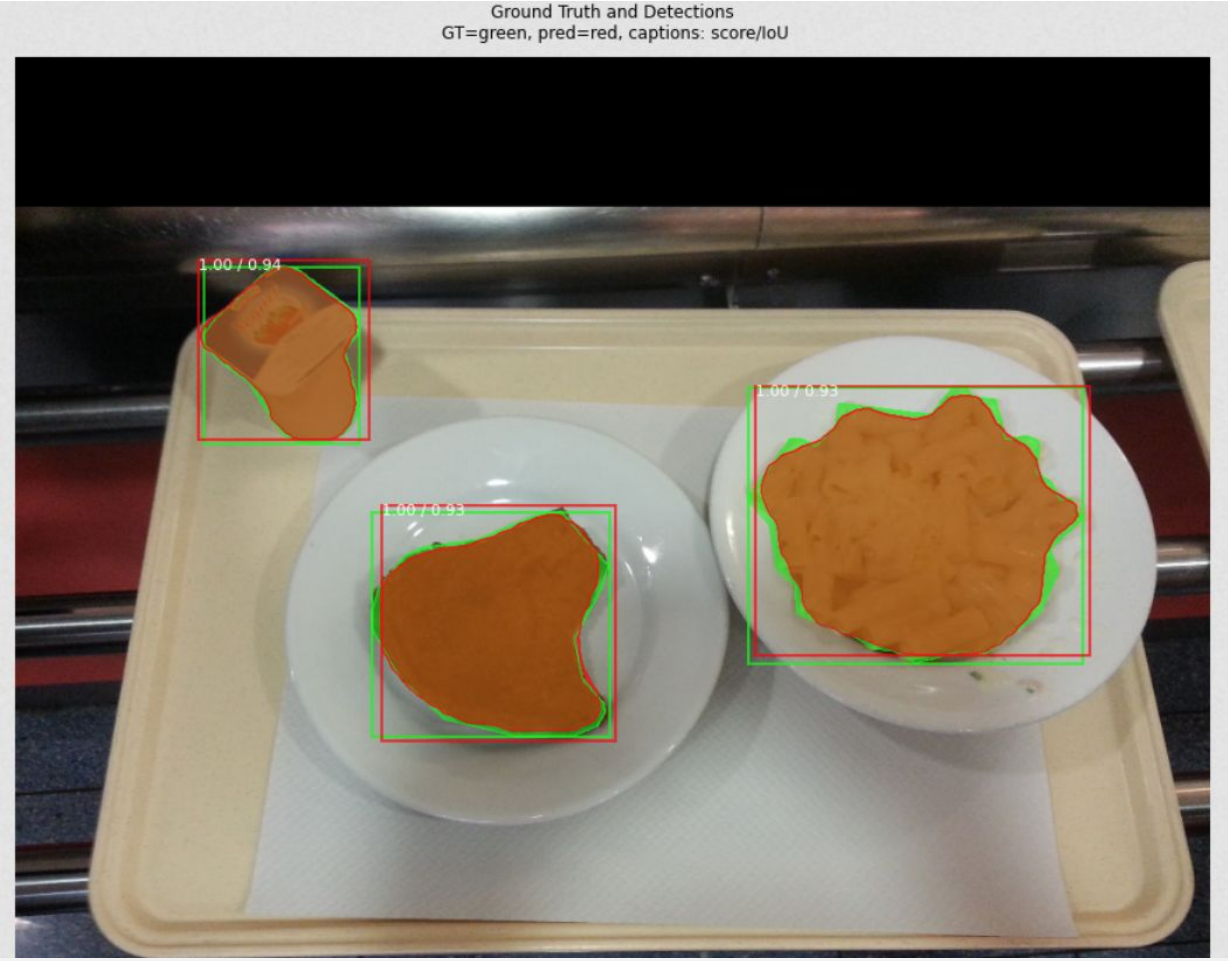


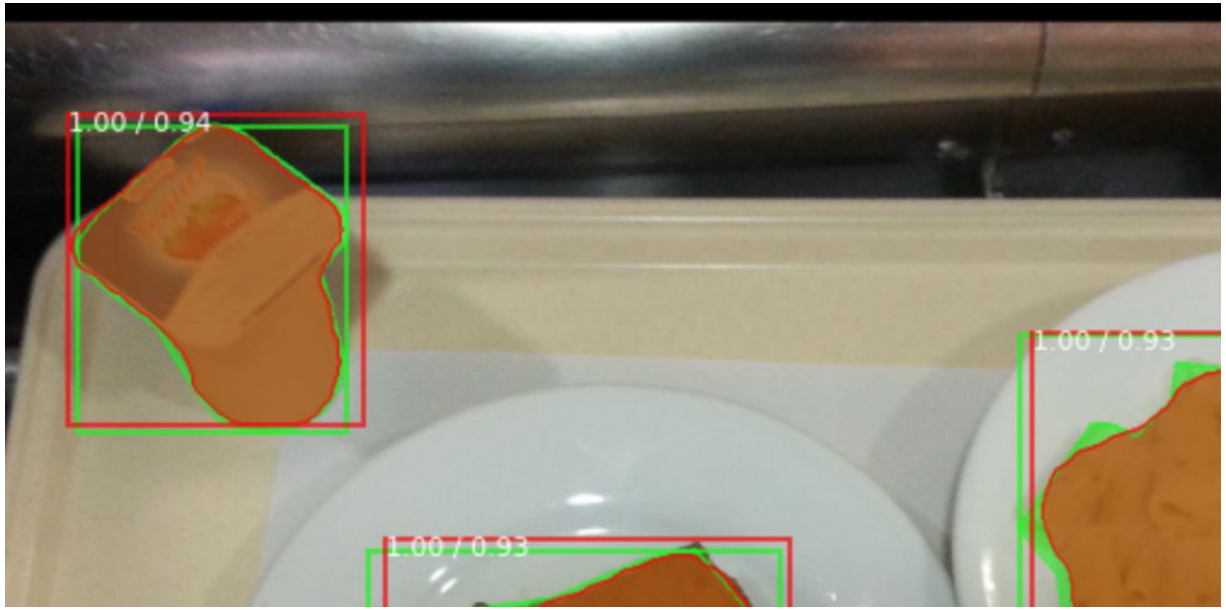
Results / Discussion (**DeepLab-V3+**)

Loss Train: 0.02
Loss Val: 0.09



Results / Discussion (MaskRCNN: mAP: 0.903)

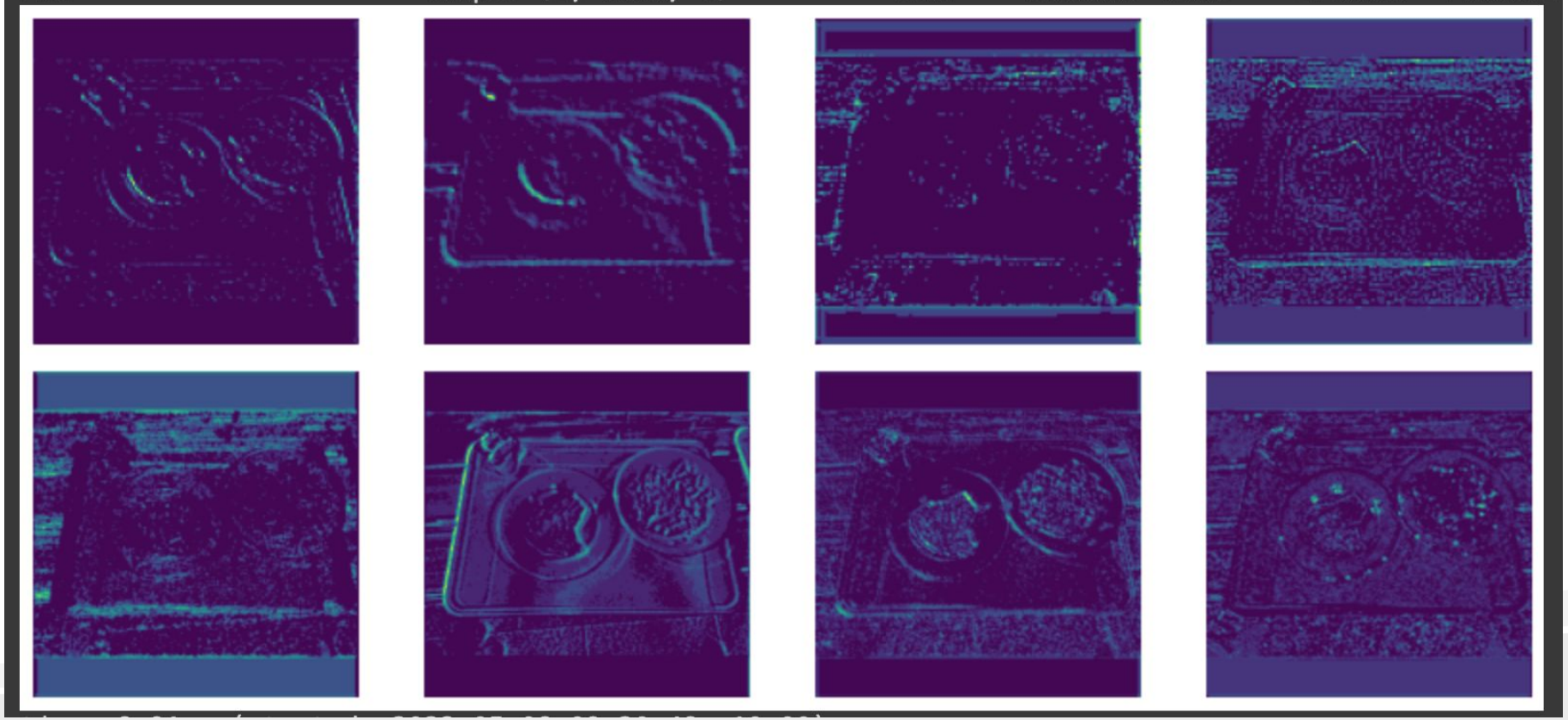




GT: Green | Pred:Orange | Caption =score/IOU

Mask RCNN Activation Layers

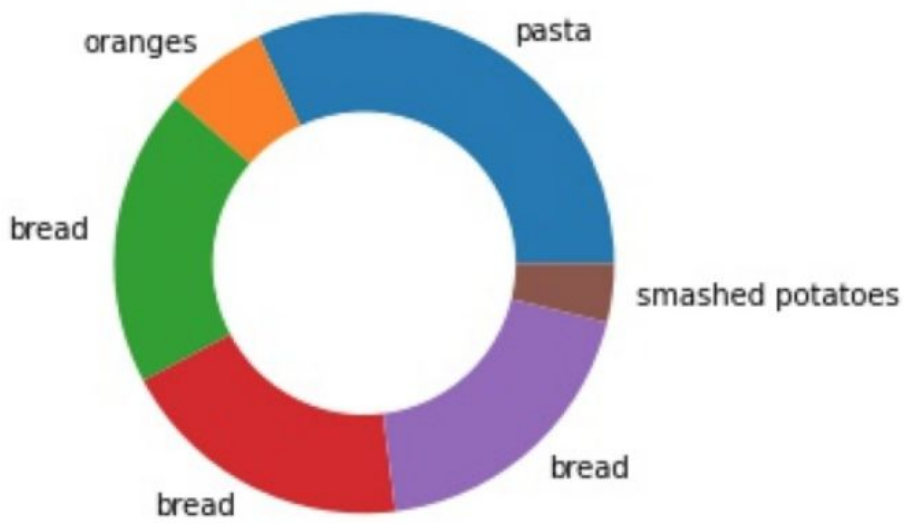
```
input_image              shape: (1, 1024, 1024, 3)    min: -123.70000  max:  151.10001  float32
res2c_out                shape: (1, 256, 256, 256)    min:    0.00000  max:   23.40121  float32
res3c_out                shape: (1, 128, 128, 512)    min:    0.00000  max:   27.75632  float32
rpn_bbox                 shape: (1, 261888, 4)        min:  -10.78870  max:   44.14271  float32
roi                      shape: (1, 1000, 4)          min:    0.00000  max:    1.00000  float32
```



Application

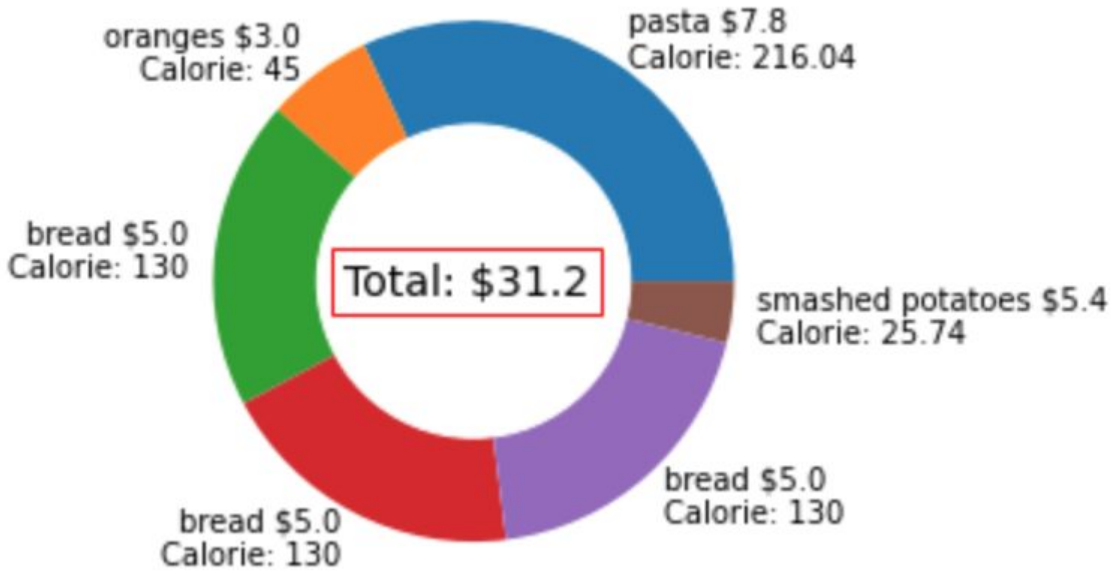
$$PixelsPerInchSQ = \frac{MaskedPlatePixel}{RealPlateArea}$$

$$RealFoodArea = \frac{MaskedFoodPixel}{PixelsPerInchSQ}$$



Application

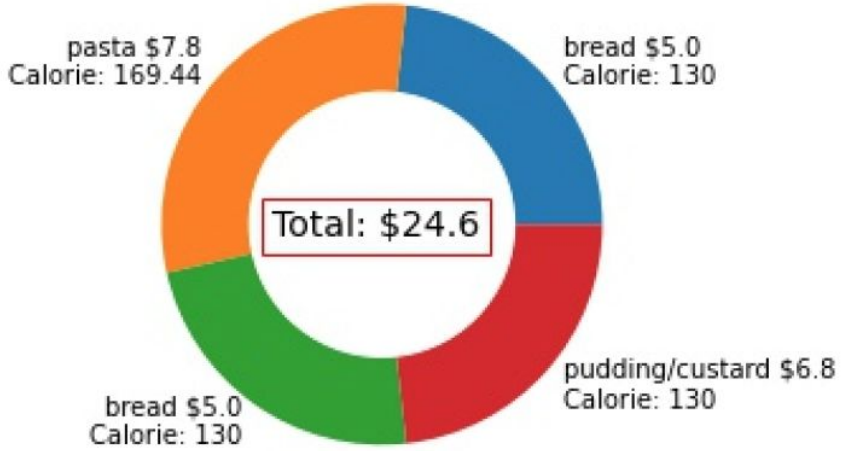
```
pasta with 216 calories
oranges with 45 calories
bread with 130 calories
bread with 130 calories
bread with 130 calories
smashed potatoes with 25 calories
```



Dict: 13 labels | Price + Calories

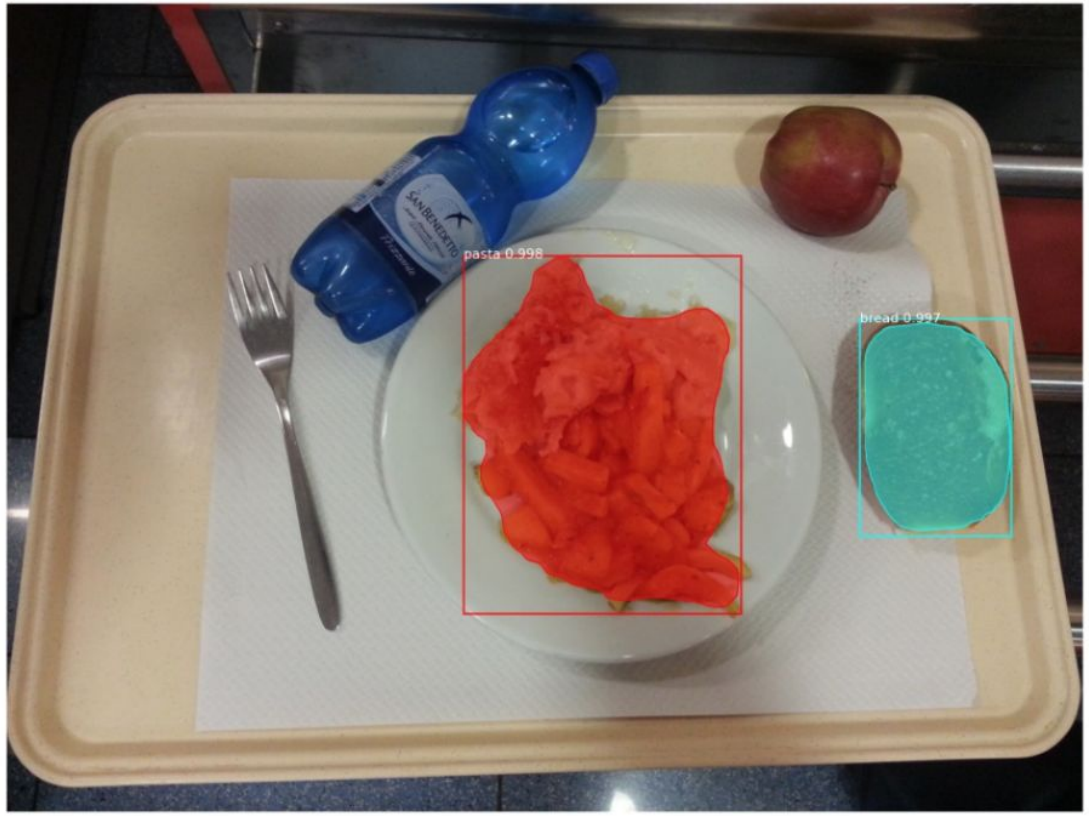
Application

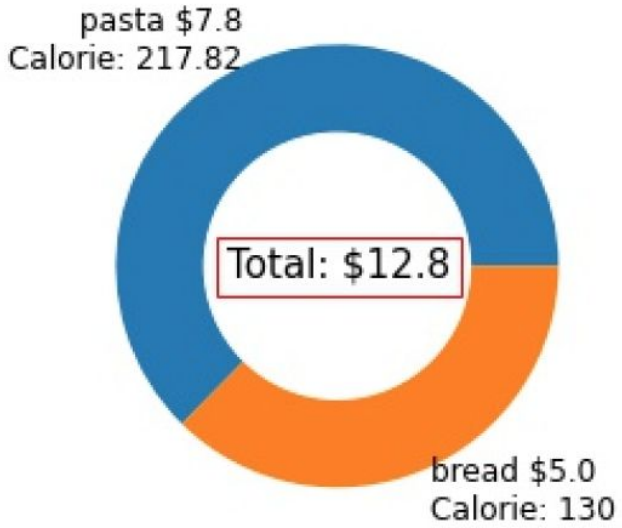
```
bread with 130 calories
pasta with 169 calories
bread with 130 calories
pudding/custard with 130 calories
```



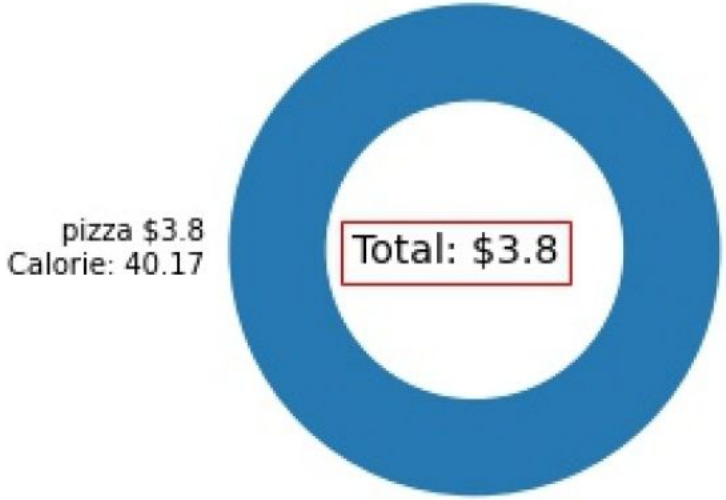
Application

```
pasta with 217 calories
bread with 130 calories
```

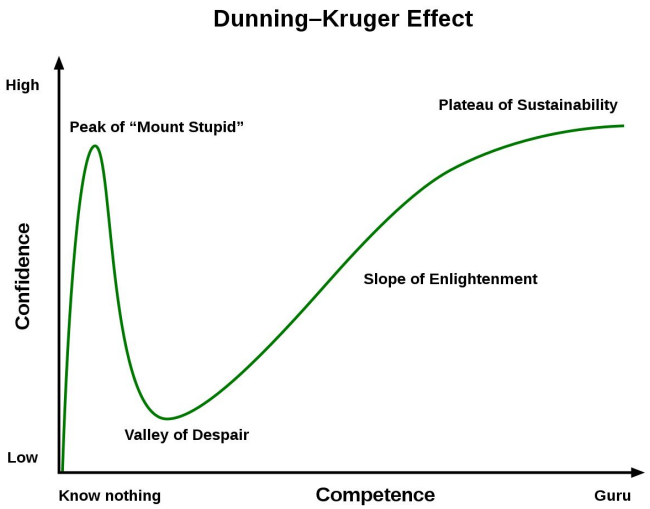




Application/ Not in dataset



Conclusion



- Experience a DL project from scratch
- Define problem
- Source of Data
- Project Management

Co

References

UNIMIB2016 Food Database :

https://mldta.com/dataset/unimib2016-food-database/

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

https://www.digitalocean.com/community/tutorials/how-to-make-a-web-application-using-flask-in-python-3

https://www.segmentationstudyguide.com/understanding-market-segmentation/market-segmentation-examples/market-segmentation-example-fast-food/