Supervised learning

Assignment 4

Name	ID
Marwan Mohamed abdelmonem	20190513
Youssef Hesham Mohamed	20190648

Report for custom dataset Assignment using Yolo deep learning

It's required to implement a machine learning algorithm for classification of the custom dataset using YOLO v5s

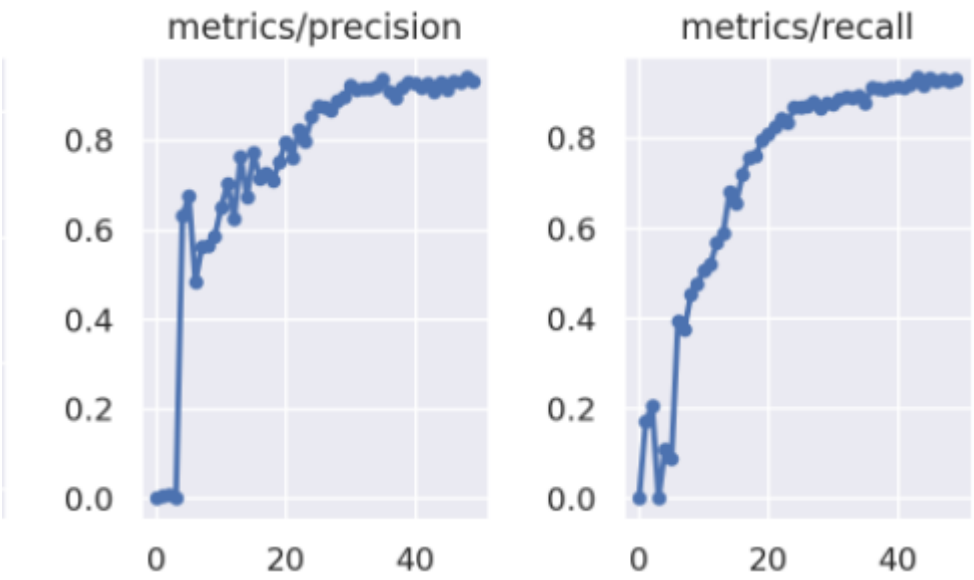
briefing phase:

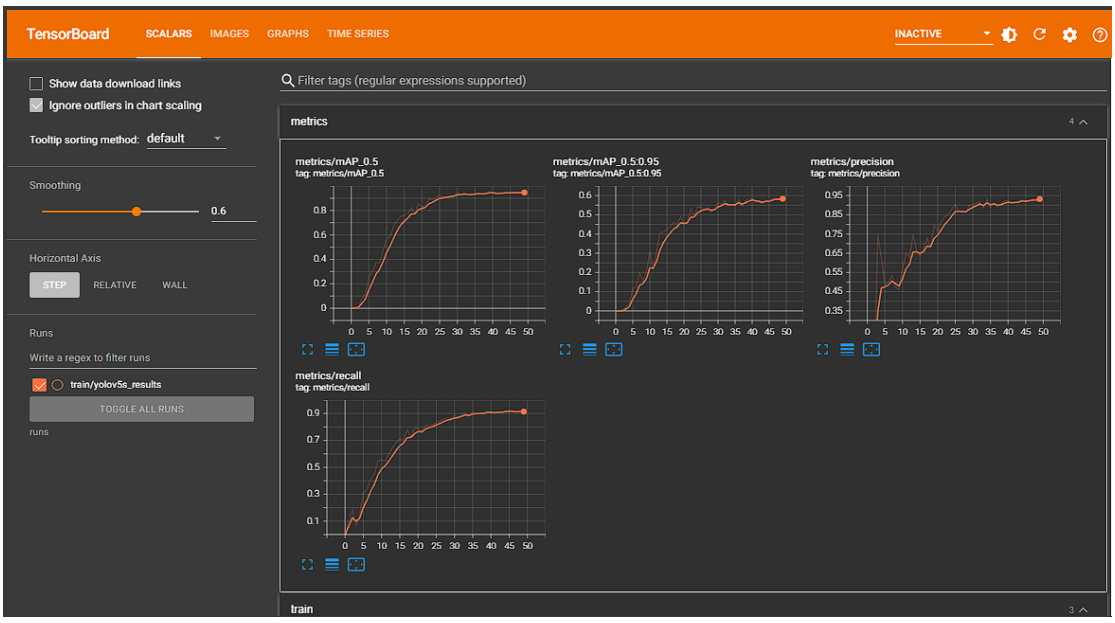
```
- apple
- banana
- beef
- blueberries
- bread
- butter
- carrot
- cheese
- chicken
- chicken_breast
- chocolate
- corn
- eggs
- flour
- goat_cheese
- green_beans
- ground_beef
- ham
- heavy_cream
- lime
- milk
- mushrooms
- onion
- potato
- shrimp
- spinach
- strawberries
- sugar
- sweet_potato
- tomato
```

We used custom dataset of 30 class representing different kinds of food in a fridge where 2896 samples are used for training and 103 for validation, 320 image size and 51 sample for training

Yolo architecture:

- Precision , recall low in early epochs and increase with a logarithmic growth





(Performance metrics using tensorboard)

- We used 24 layers in yolo architecture and 283 in training and 7333307 parameters

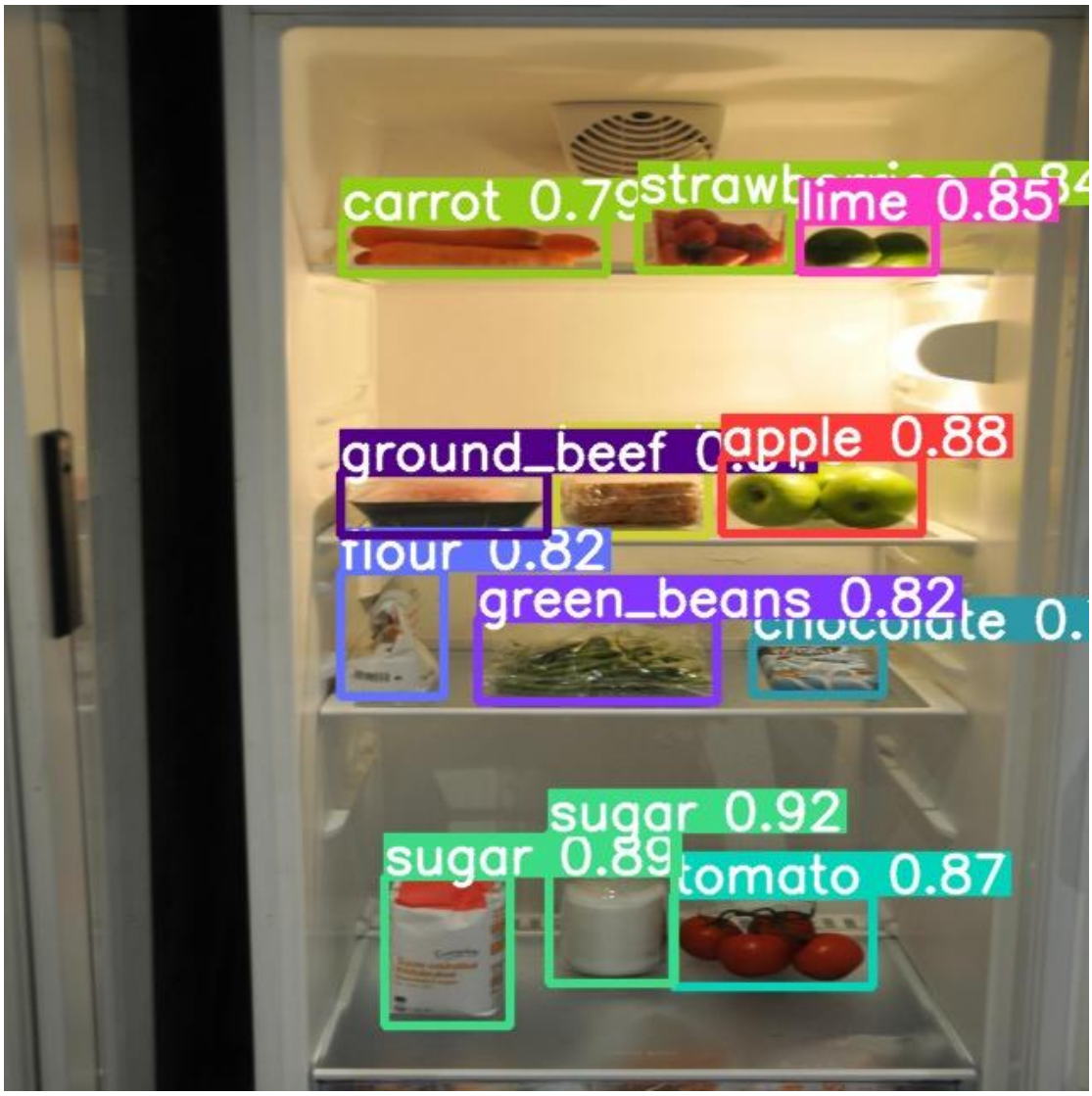
```
[[-1, 1, Focus, [64, 3]],  # 0-P1/2
 [-1, 1, Conv, [128, 3, 2]],  # 1-P2/4
 [-1, 3, BottleneckCSP, [128]],
 [-1, 1, Conv, [256, 3, 2]],  # 3-P3/8
 [-1, 9, BottleneckCSP, [256]],
 [-1, 1, Conv, [512, 3, 2]],  # 5-P4/16
 [-1, 9, BottleneckCSP, [512]],
 [-1, 1, Conv, [1024, 3, 2]],  # 7-P5/32
 [-1, 1, SPP, [1024, [5, 9, 13]]],
 [-1, 3, BottleneckCSP, [1024, False]],  # 9
]
```

```
[[-1, 1, Conv, [512, 1, 1]],
 [-1, 1, nn.Upsample, [None, 2, 'nearest']],
 [[-1, 6], 1, Concat, [1]],  # cat backbone P4
 [-1, 3, BottleneckCSP, [512, False]],  # 13

 [-1, 1, Conv, [256, 1, 1]],
 [-1, 1, nn.Upsample, [None, 2, 'nearest']],
 [[-1, 4], 1, Concat, [1]],  # cat backbone P3
 [-1, 3, BottleneckCSP, [256, False]],  # 17 (P3/8-small)

 [-1, 1, Conv, [256, 3, 2]],
 [[-1, 14], 1, Concat, [1]],  # cat head P4
 [-1, 3, BottleneckCSP, [512, False]],  # 20 (P4/16-medium)

 [-1, 1, Conv, [512, 3, 2]],
 [[-1, 10], 1, Concat, [1]],  # cat head P5
 [-1, 3, BottleneckCSP, [1024, False]],  # 23 (P5/32-large)

 [[17, 20, 23], 1, Detect, [nc, anchors]],  # Detect(P3, P4, P5)
```

- Activation function: Leaky ReLU activation function is used in hidden layers and the sigmoid activation function is used in the final detection layer
- Grid size 1 * 3
- Number of anchors 3
- optimization function for training is SGD
- loss function is IOU loss as it gave us the minimum loss

➢ Predicted images:

We use 10 images with number of class, x coordinate , y coordinate , width , height



4 0.578125 0.455469 0.140625 0.0640625

10 0.749219 0.609375 0.120312 0.05

6 0.432813 0.222656 0.24375 0.0484375

15 0.546094 0.602344 0.220313 0.0765625

13 0.357031 0.578125 0.0953125 0.115625

26 0.654688 0.213281 0.140625 0.0609375

16 0.404687 0.458594 0.190625 0.0578125

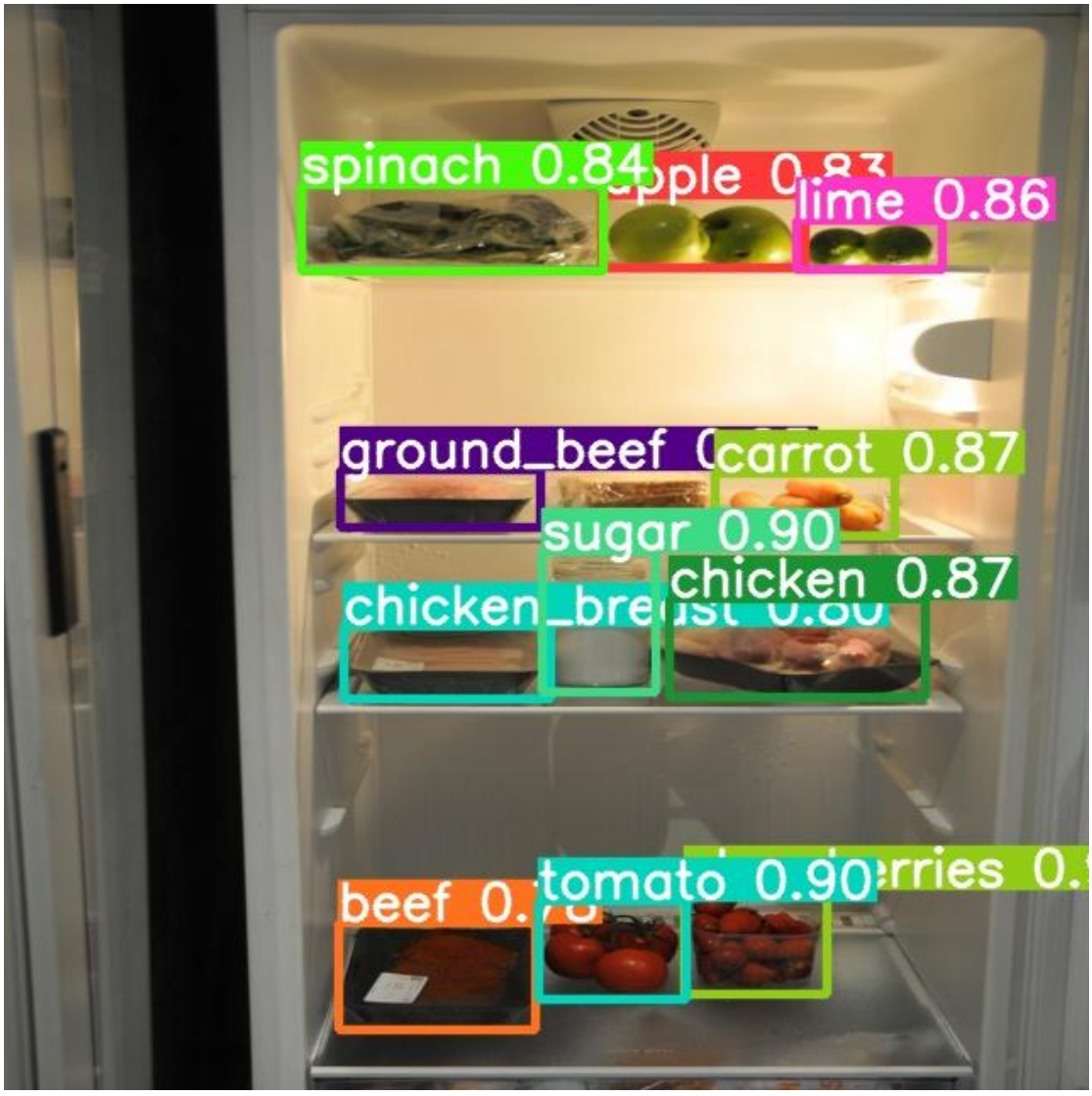
19 0.795313 0.221875 0.125 0.046875

29 0.707812 0.860937 0.184375 0.084375

0 0.753125 0.450781 0.184375 0.0703125

27 0.407813 0.870313 0.115625 0.1375

27 0.559375 0.830469 0.115625 0.139062



2 0.398438 0.894531 0.18125 0.0984375

9 0.407813 0.605469 0.190625 0.0671875

0 0.64375 0.208594 0.1875 0.0671875

16 0.401563 0.454688 0.184375 0.053125

25 0.4125 0.205469 0.278125 0.0796875

19 0.796875 0.221094 0.134375 0.0453125

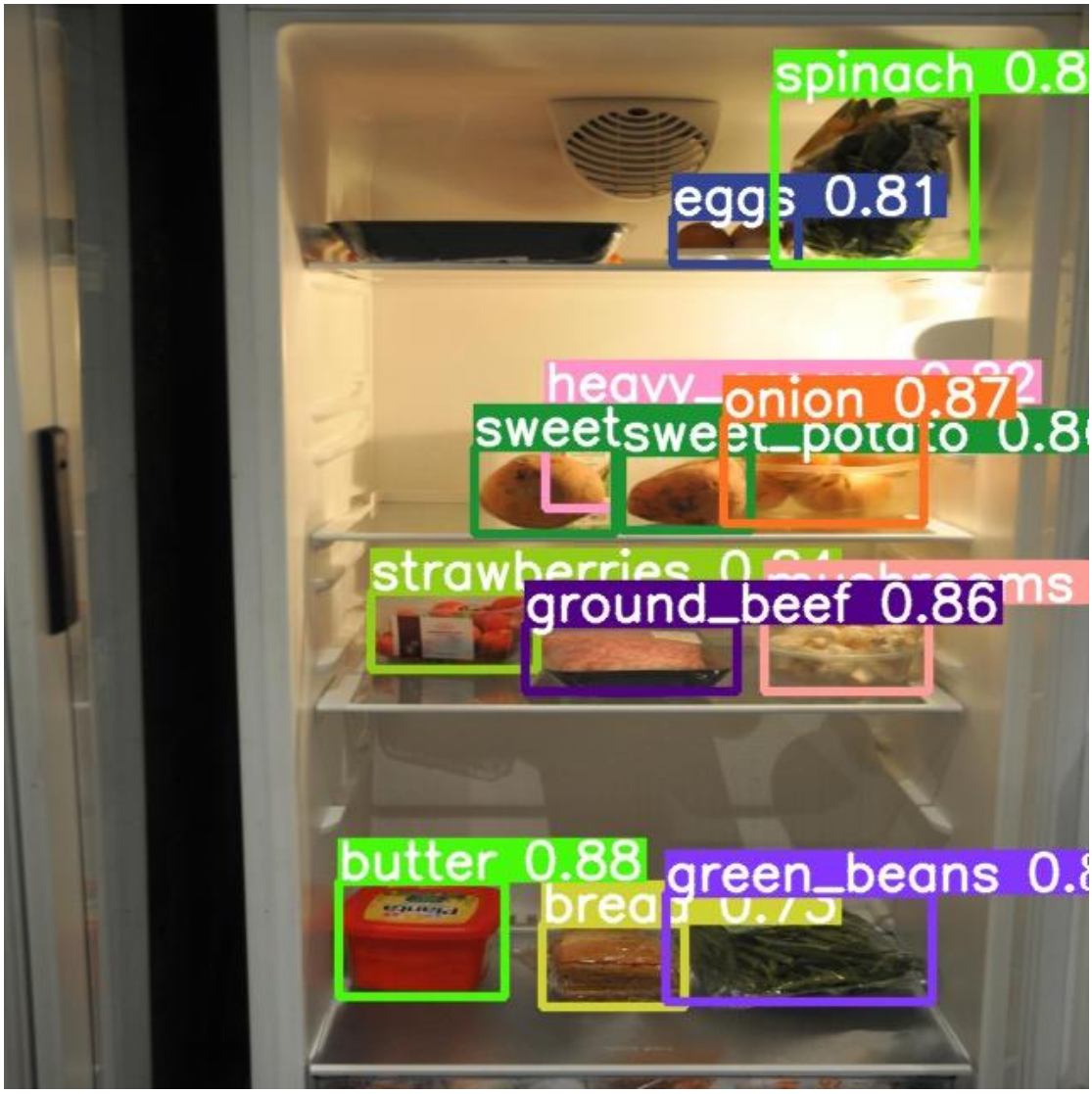
8 0.730469 0.592188 0.235938 0.090625

6 0.7375 0.460156 0.165625 0.0578125

26 0.692187 0.8625 0.13125 0.096875

29 0.560156 0.870313 0.135937 0.090625

27 0.547656 0.567969 0.104687 0.129688



4 0.560938 0.885156 0.13125 0.0765625

12 0.673437 0.217187 0.115625 0.04375

18 0.530469 0.415625 0.0640625 0.096875

28 0.497656 0.447656 0.129688 0.0796875

15 0.732031 0.86875 0.245312 0.1

26 0.414062 0.576563 0.153125 0.071875

21 0.775781 0.592188 0.151562 0.08125

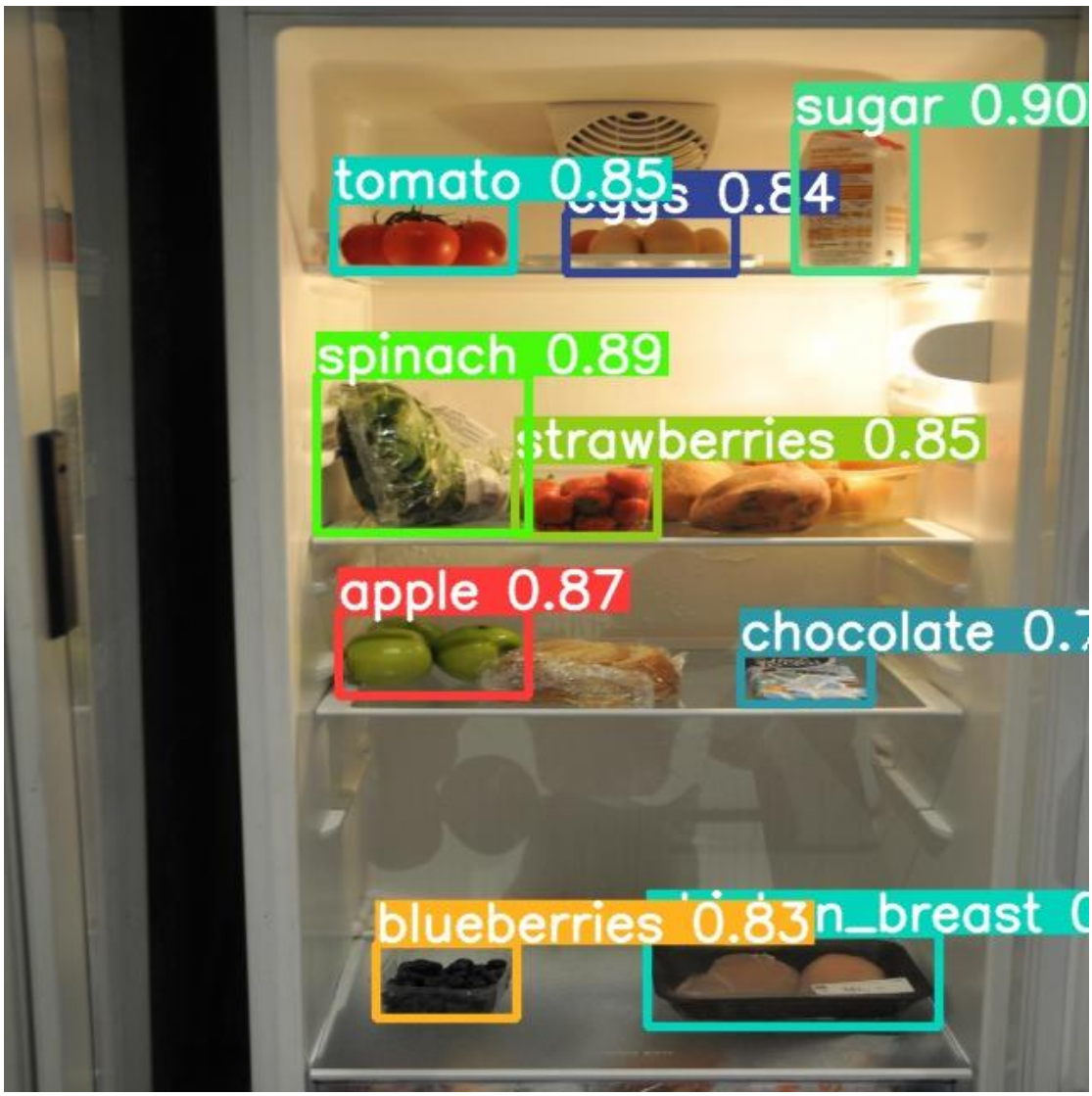
16 0.577344 0.601562 0.195312 0.0625

28 0.628125 0.448438 0.11875 0.071875

25 0.801562 0.160156 0.184375 0.157813

22 0.754687 0.429688 0.184375 0.096875

5 0.384375 0.860937 0.153125 0.10625



10 0.738281 0.617969 0.120312 0.0484375

9 0.726562 0.896094 0.26875 0.0859375

3 0.407031 0.896875 0.129688 0.06875

12 0.595312 0.217969 0.15625 0.0546875

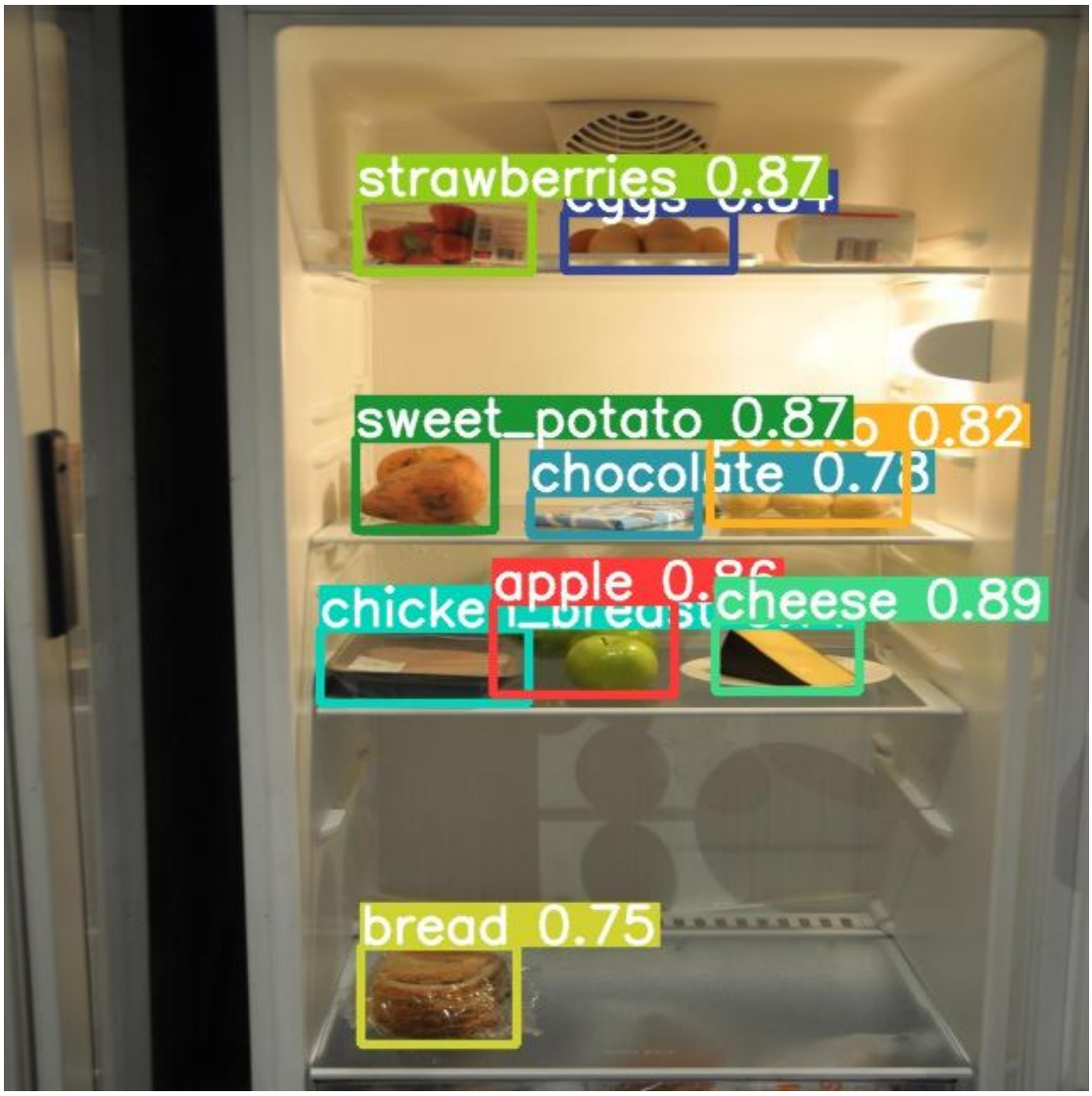
29 0.385938 0.210938 0.165625 0.065625

26 0.535937 0.452344 0.13125 0.0703125

0 0.395312 0.595312 0.175 0.078125

25 0.385156 0.411719 0.195312 0.145313

27 0.783594 0.176563 0.110937 0.134375



4 0.4 0.910937 0.14375 0.090625

9 0.386719 0.607813 0.192188 0.06875

10 0.561719 0.467969 0.154687 0.0390625

23 0.740625 0.441406 0.18125 0.0703125

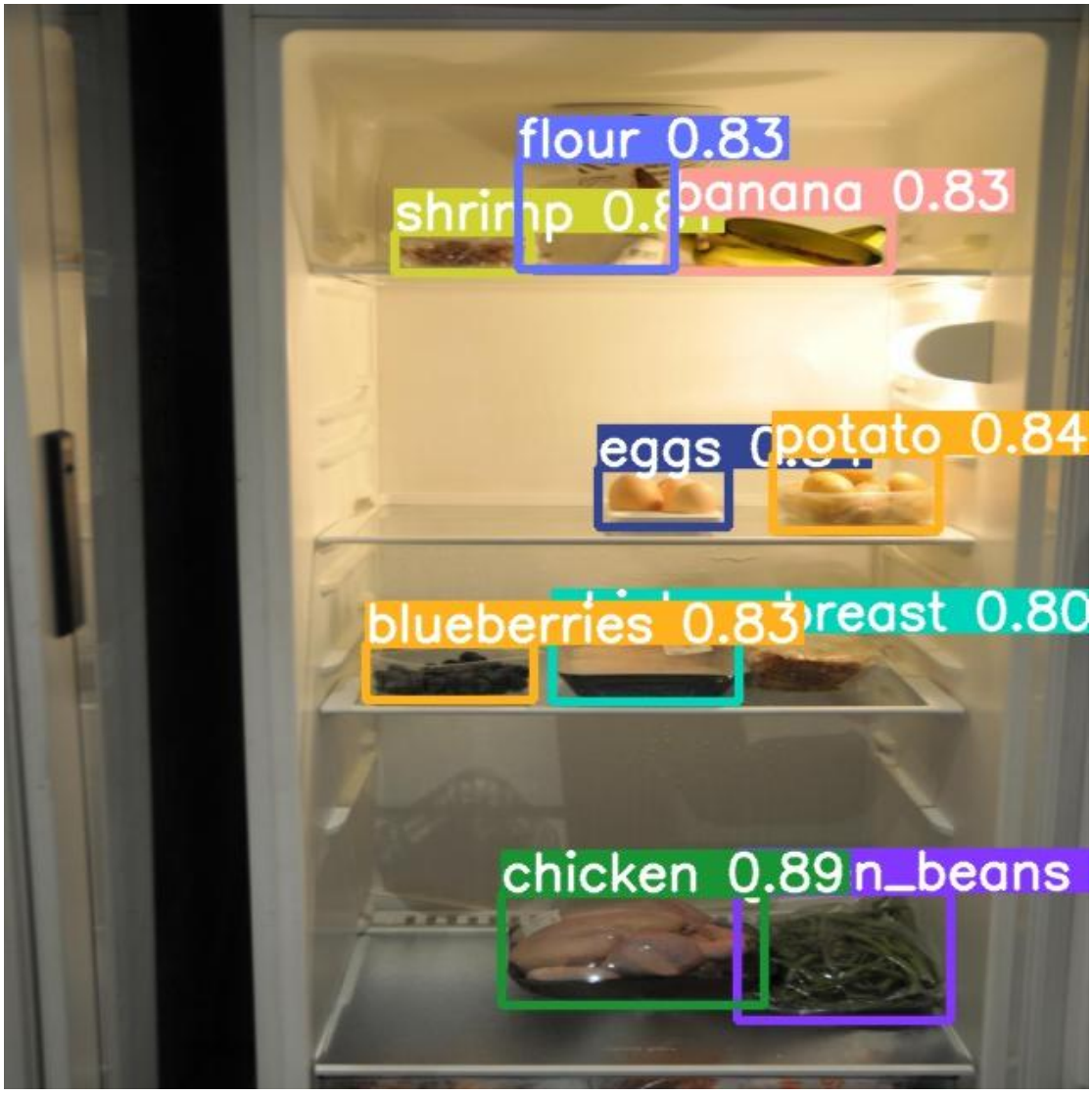
12 0.59375 0.217187 0.15625 0.053125

0 0.533594 0.591406 0.167187 0.0859375

28 0.3875 0.440625 0.128125 0.084375

26 0.405469 0.210938 0.160938 0.06875

7 0.721094 0.598437 0.132812 0.065625



9 0.590625 0.610156 0.171875 0.0640625

24 0.421875 0.228125 0.125 0.0375

1 0.715625 0.217969 0.203125 0.0546875

13 0.545313 0.192969 0.14375 0.101562

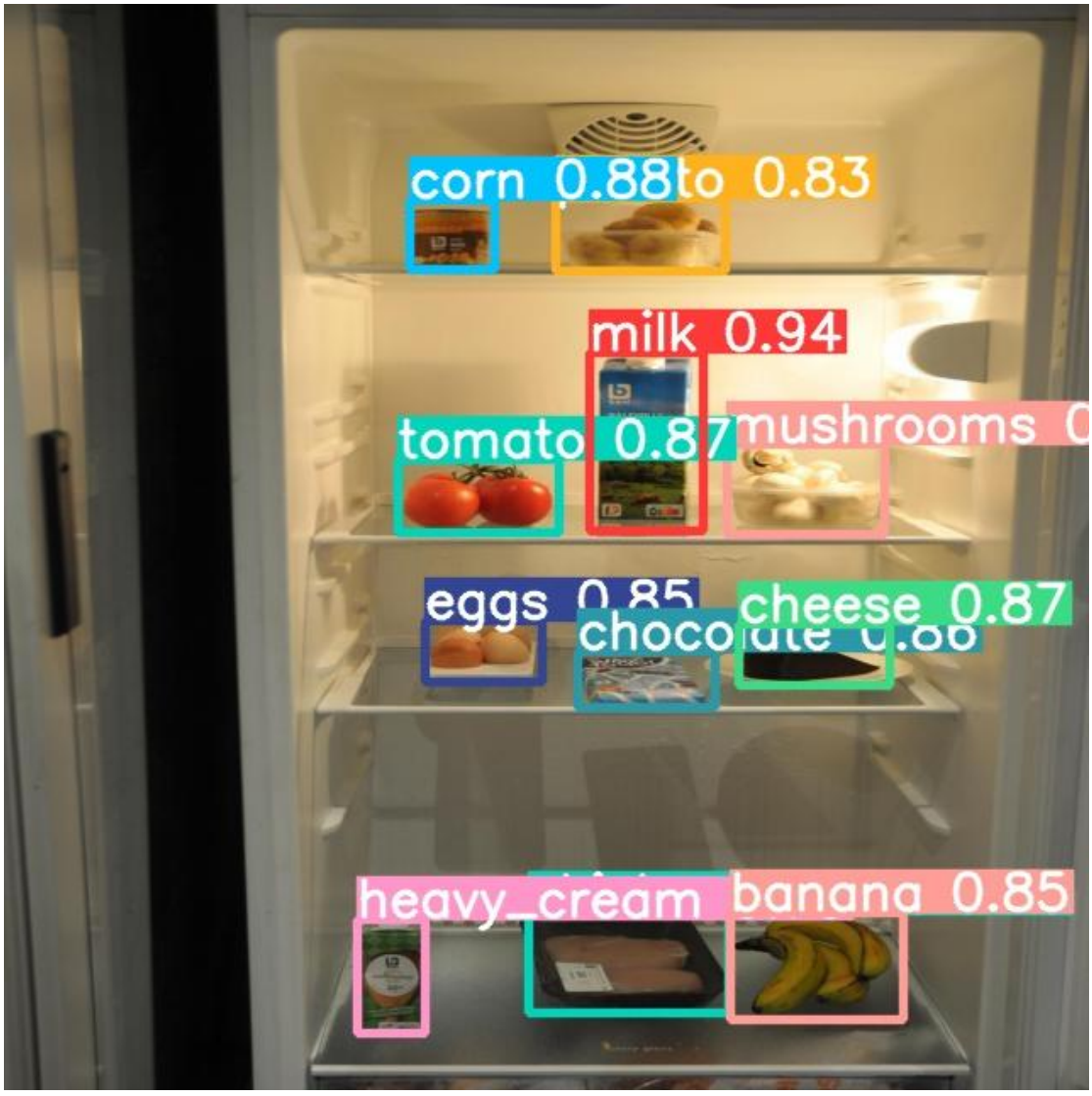
3 0.410156 0.614844 0.154687 0.0515625

12 0.607031 0.453906 0.120312 0.0546875

23 0.784375 0.449219 0.153125 0.0703125

15 0.773438 0.876562 0.196875 0.11875

8 0.578906 0.870313 0.242188 0.103125



9 0.575 0.882812 0.184375 0.090625

18 0.35625 0.894531 0.0625 0.104687

21 0.738281 0.446094 0.145313 0.0828125

23 0.585938 0.211719 0.15625 0.0671875

1 0.748438 0.885156 0.159375 0.0984375

12 0.441406 0.595312 0.107813 0.05625

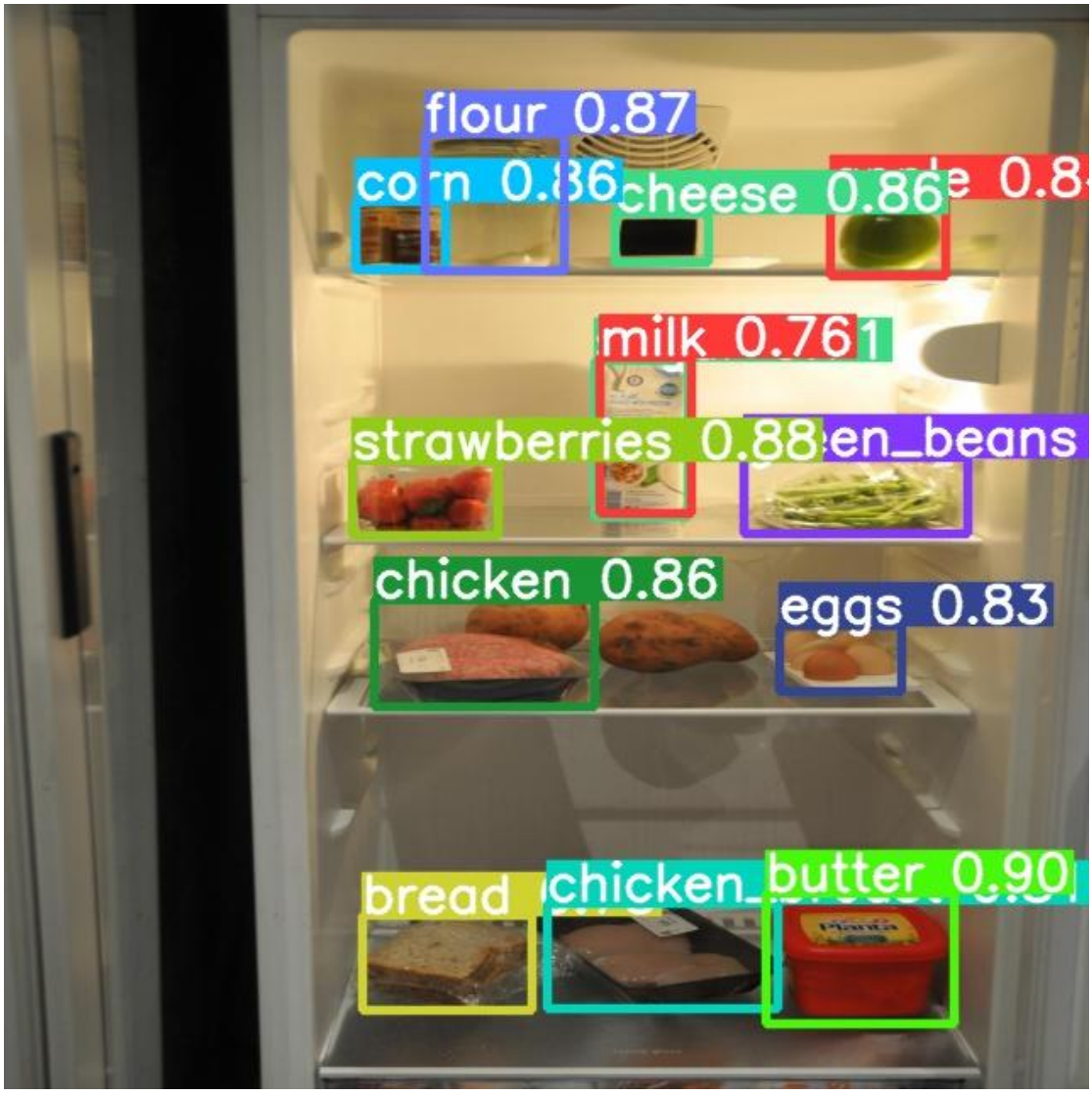
10 0.591406 0.621094 0.126563 0.0515625

29 0.436719 0.452344 0.148438 0.0671875

7 0.746094 0.598437 0.139062 0.05625

11 0.4125 0.211719 0.078125 0.0640625

20 0.591406 0.402344 0.104687 0.164062



27 0.586719 0.4 0.0859375 0.14375

4 0.407813 0.882812 0.15625 0.0875

20 0.590625 0.396875 0.084375 0.14375

9 0.60625 0.876562 0.2125 0.096875

12 0.770312 0.602344 0.1125 0.0609375

0 0.814062 0.214062 0.10625 0.071875

15 0.784375 0.452344 0.20625 0.0703125

7 0.605469 0.214844 0.0859375 0.0453125

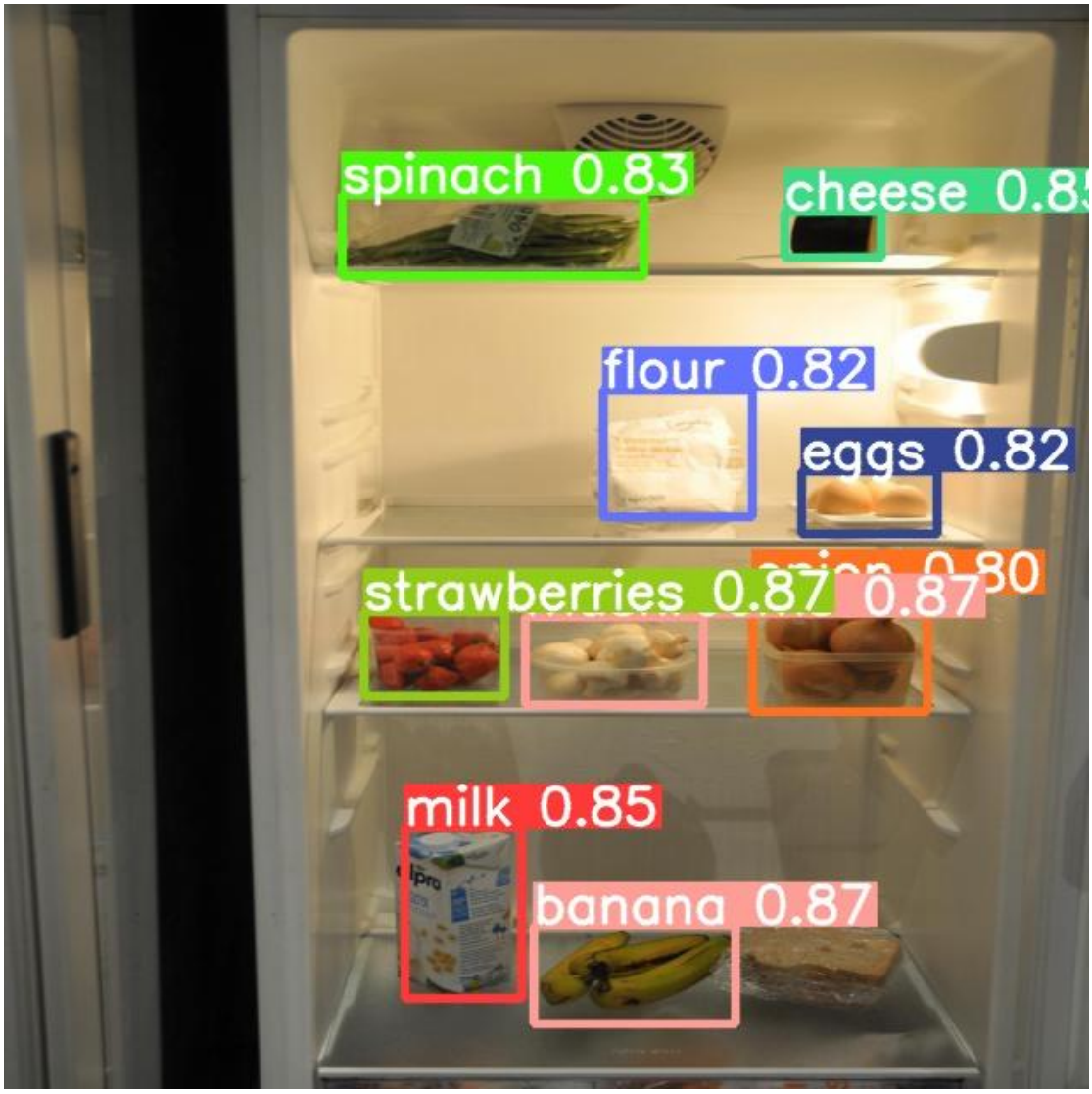
11 0.364844 0.2125 0.0828125 0.0625

8 0.442187 0.597656 0.203125 0.0984375

13 0.451562 0.182031 0.128125 0.126563

26 0.3875 0.454688 0.134375 0.06875

5 0.788281 0.878906 0.173438 0.120312



22 0.769531 0.596875 0.160938 0.109375

12 0.796875 0.458594 0.125 0.0578125

13 0.620313 0.413281 0.1375 0.117188

25 0.45 0.2125 0.278125 0.075

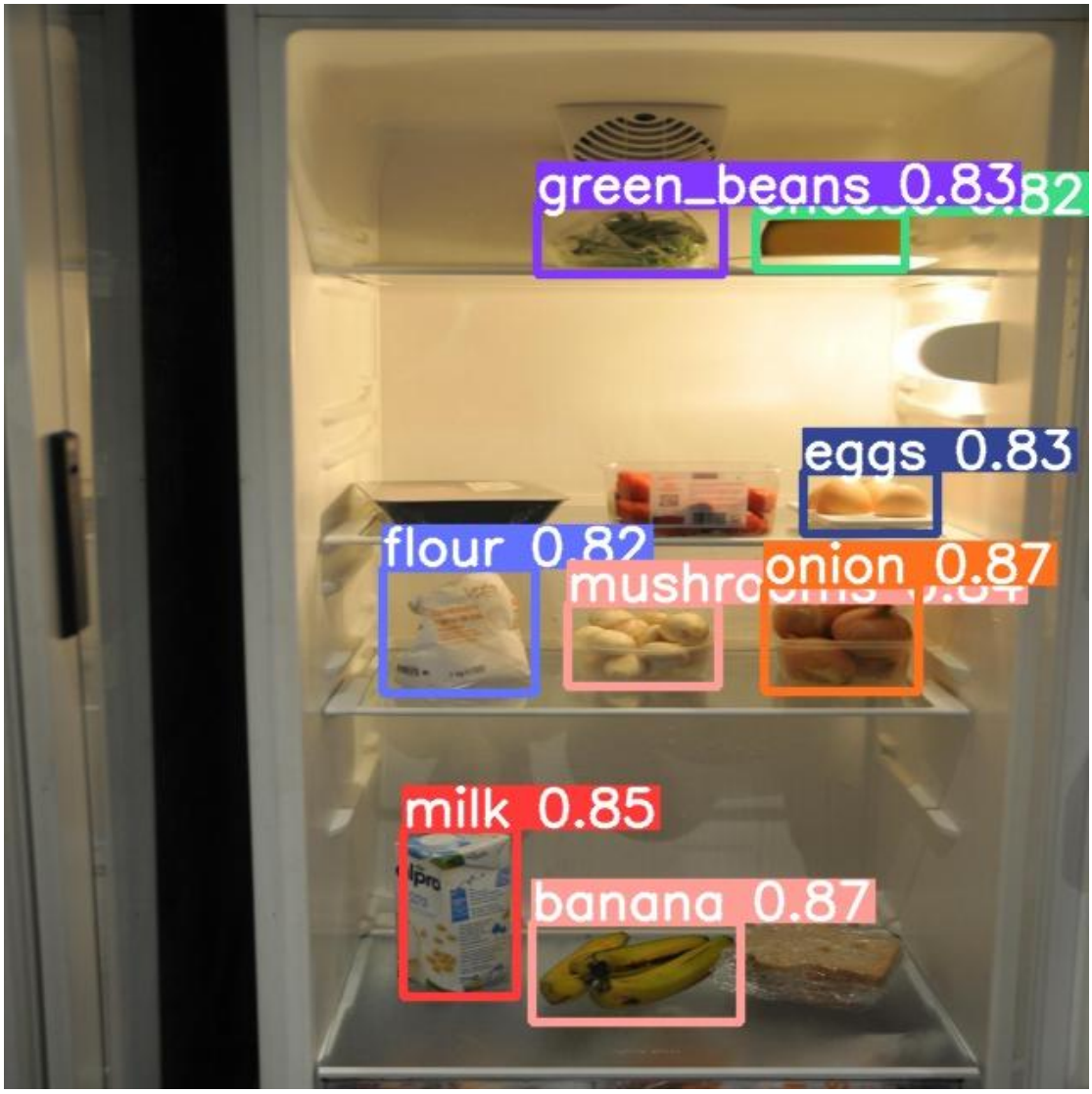
7 0.763281 0.211719 0.0890625 0.0421875

20 0.422656 0.836719 0.107813 0.157813

21 0.561719 0.604688 0.164062 0.08125

1 0.580469 0.89375 0.185937 0.090625

26 0.396094 0.598437 0.129688 0.078125



7 0.760938 0.21875 0.1375 0.05

13 0.41875 0.576563 0.140625 0.115625

12 0.797656 0.458594 0.123438 0.0578125

15 0.576563 0.216406 0.171875 0.0640625

21 0.588281 0.589844 0.139062 0.0765625

20 0.419531 0.836719 0.104687 0.154687

1 0.582031 0.891406 0.192188 0.0921875

22 0.771094 0.583594 0.142188 0.0984375

- total time for prediction 8 sec and 0 to 2 sec for each image
- IOU value is 0.45