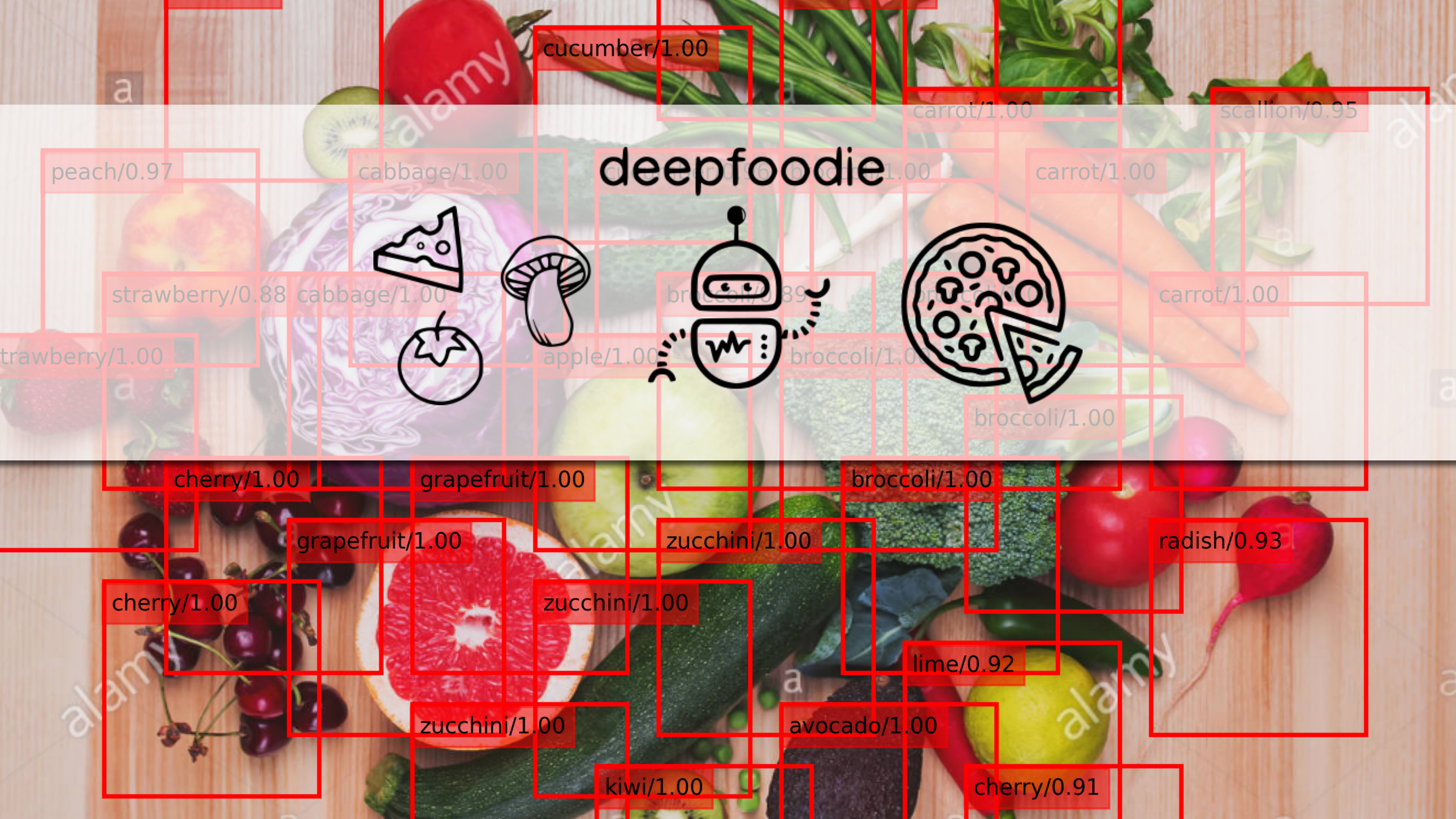
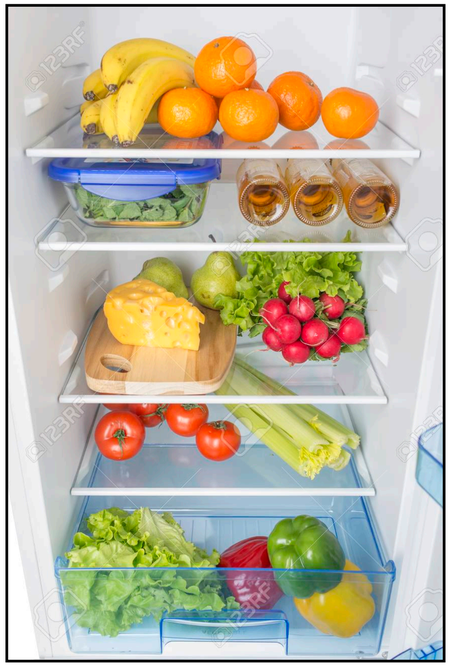

cucumber/1.00
carrot/1.00
scallion/0.95
peach/0.97
cabbage/1.00
carrot/1.00
deepfoodie
strawberry/0.88
cabbage/1.00
carrot/1.00
apple/1.00
broccoli/1.00
broccoli/1.00
cherry/1.00
grapefruit/1.00
broccoli/1.00
grapefruit/1.00
zucchini/1.00
radish/0.93
cherry/1.00
zucchini/1.00
lime/0.92
zucchini/1.00
avocado/1.00
kiwi/1.00
cherry/0.91

Goal

image



ingredients

banana
tangerine
cheese
pear
lettuce
radish
tomato
celery
paprika

recipe

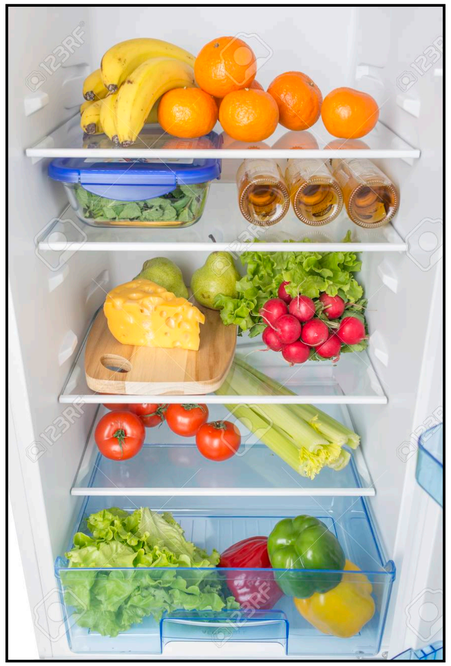
Farmer’s salad

lettuce
tomato
paprika
celery
cheese
radish

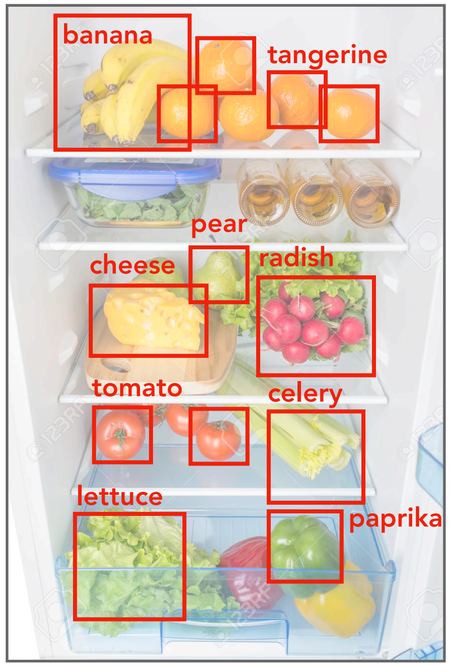
Cut the lettuce leaves crosswise into thin shreds and place in a large bowl. Core and dice the tomatoes; add to the bowl. Dice the celery, and add to the bowl. Cut paprika add to the bowl. Trim and thinly slice the radishes, and add to the bowl. Dice the cheese and add to the bowl. Drizzle with the oil and vinegar. The salad is best served the day it's made.

Food identification

image

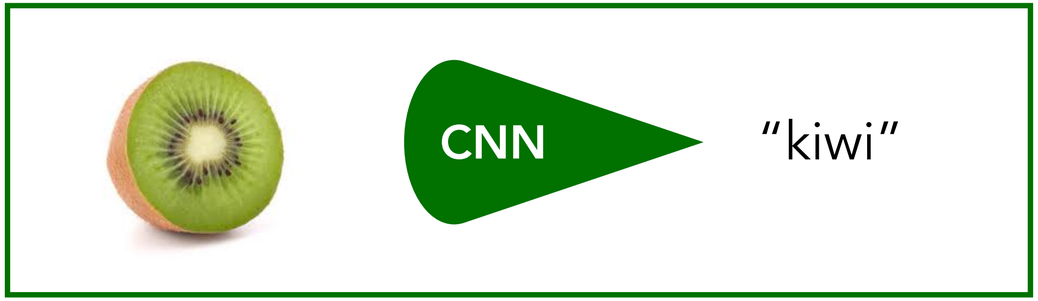


object identification



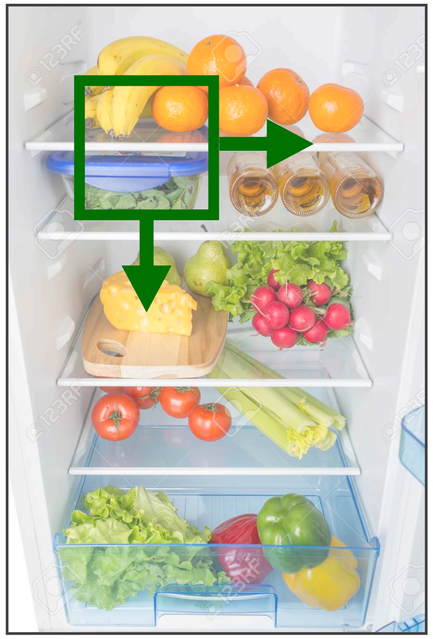
- Deep neural nets for object detection, such as **YOLO** or **SSD**.
- **But:** We don't really need the **where**, only the **what**.
- **Data problem:** There is no dataset of fridge photos with annotated bounding boxes for food items.

Solution: Defining sub-problems



Food classification network

Sliding window algorithm



\+ Tons of labeled images are readily accessible online.

\+ No need for manual labeling

\+ No need to predict bounding boxes

\- Slower than “one-shot” detection

\- No “end-to-end” learning: The model learns the visual features of each object class in isolation of other objects.

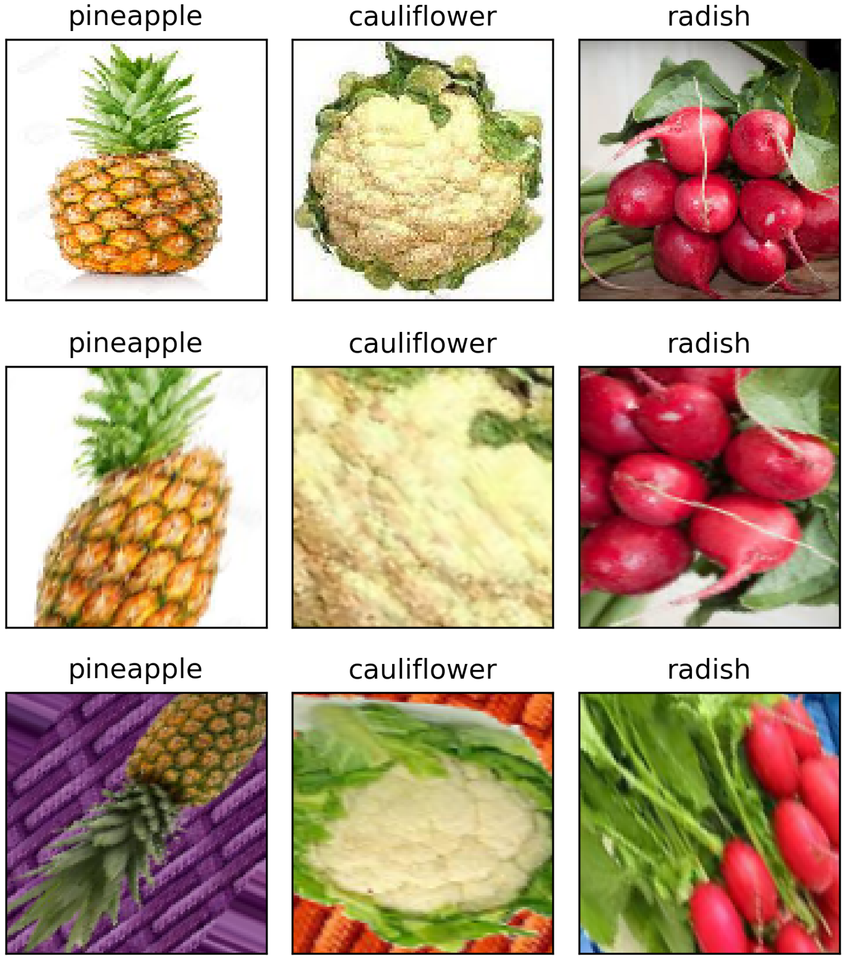
Food classification: Data

~11.000 images from 53 classes (200 images/class)

almond	cauliflower	grapefruit	onion	radish
apple	celery	grapes	orange	raspberry
apricot	cheese	grated_cheese	other	salami
avocado	cherry	kiwi	paprika	scallion
banana	chicken_breast	lemon	passionfruit	strawberry
beef	chocolate	lettuce	peach	tomato
blackberry	corn	lime	pear	watermelon
blueberry	cucumber	mango	pineapple	whole_chicken
broccoli	egg	melon	plum	zucchini
cabbage	eggplant	mushroom	pomegranate	
carrot	fig	olive	pork	

- **Data augmentation**: Random transformations such as flipping, shearing, zooming

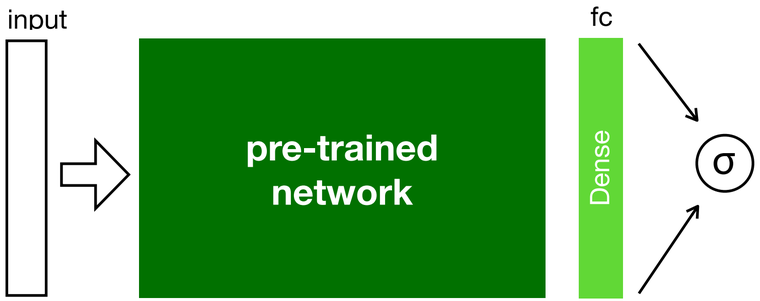
- **Data generation**: We synthesized pictures of items on random backgrounds (fridges, tables, or random patterns/ textures with different colors)



Food classification: Model

Google Cloud

Transfer learning on pre-trained deep CNNs & fine-tuning

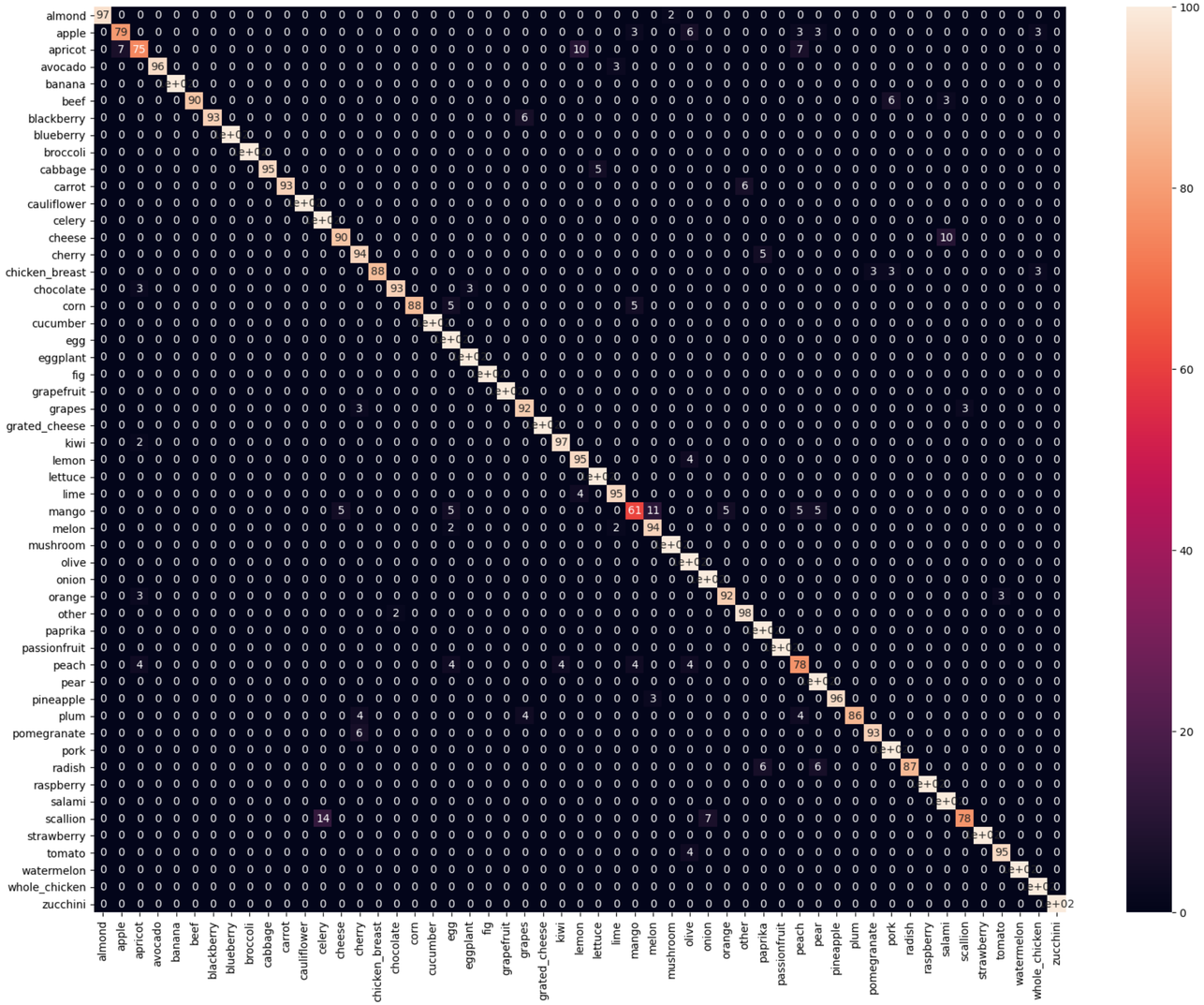


Tested hyper-parameters:

- Resnet50, MobileNetV2, InceptionV3
- Different data augmentations
- Number of epochs
- Target size

Best model: **InceptionV3** / **~95% accuracy** on hold-out test set (10% of the data)

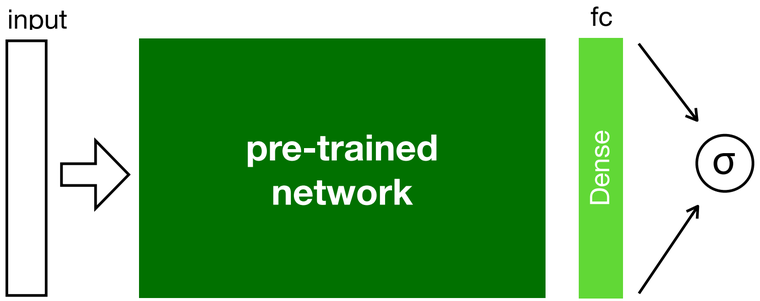
Confusion matrix



Food classification: Model

Google Cloud

Transfer learning on pre-trained deep CNNs

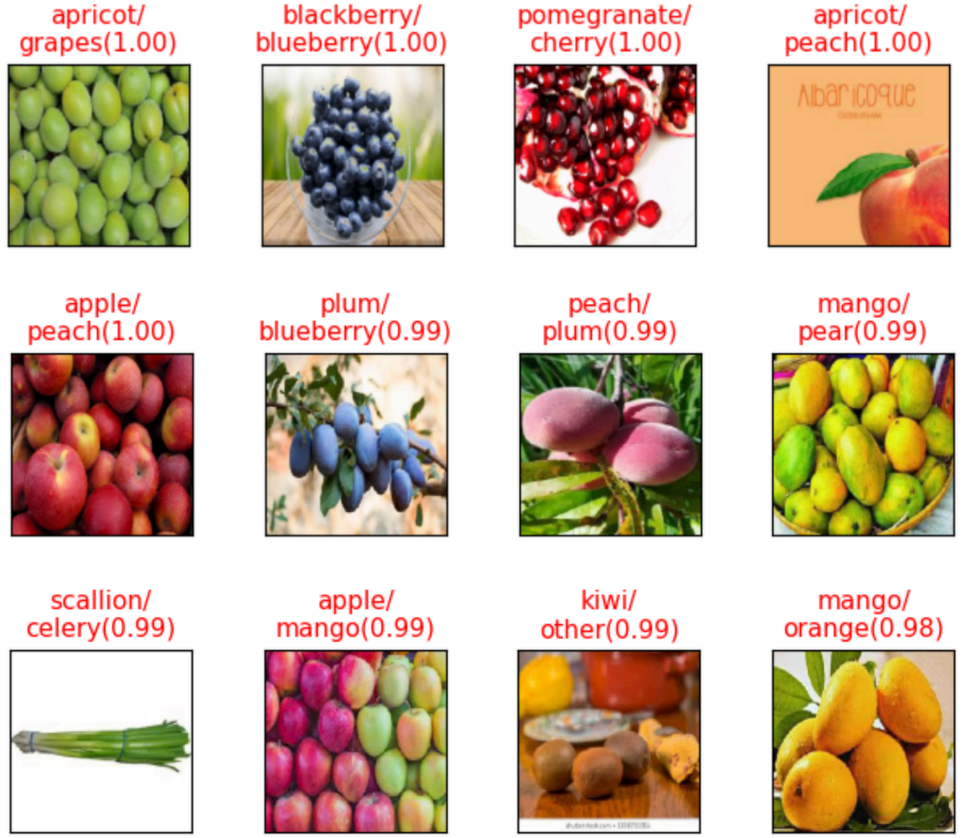


Tested hyper-parameters:

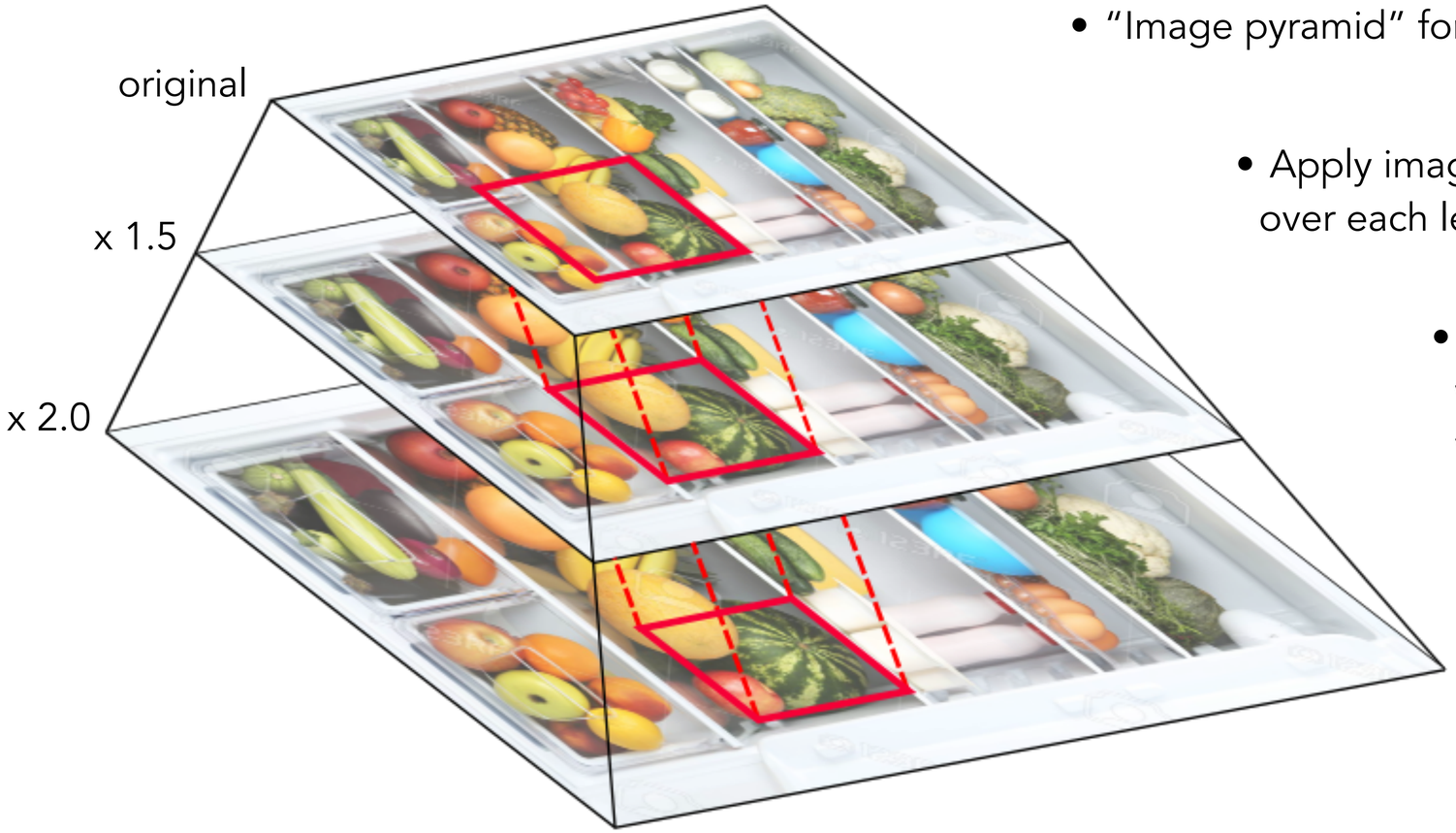
- Resnet50, MobileNetV2, InceptionV3
- Different data augmentations
- Number of epochs
- Target size

Best model: **InceptionV3** / ~95% accuracy on hold-out test set (10% of the data)

Worst misclassifications

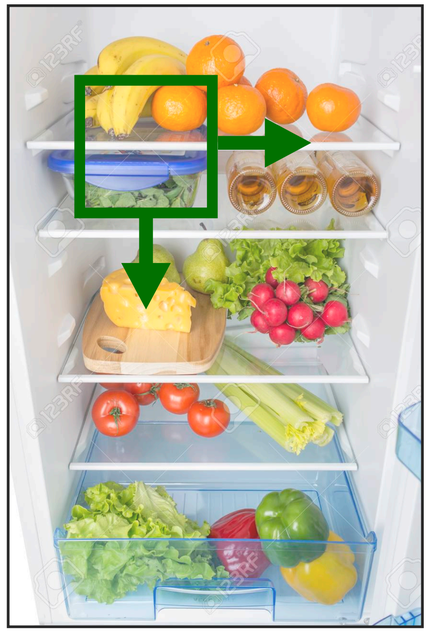


Object detection

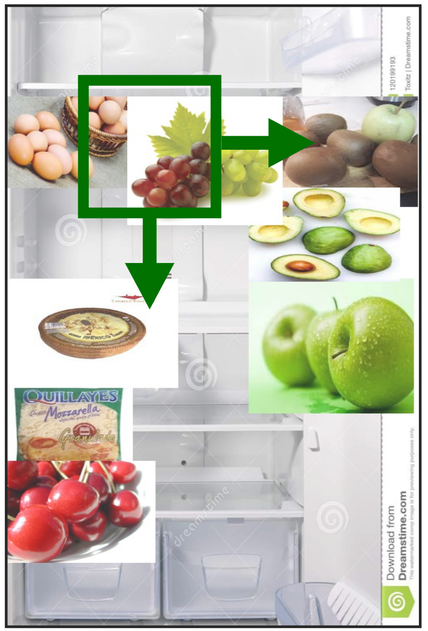


- "Image pyramid" for different object sizes
- Apply image classification on a sliding window over each level of the pyramid.
- Minimize running time by using a "strided" window, optimizing array slicing and vectorizing all operations as much as possible.

Tuning the object detection algorithm



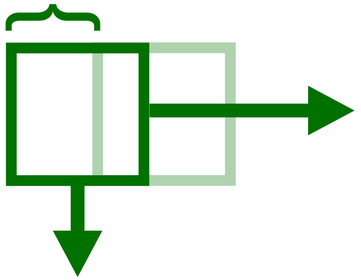
“Real” object set
35 downloaded images
manually labelled



“Artificial” object set
100 images from test set
randomly generated

Scaling factors

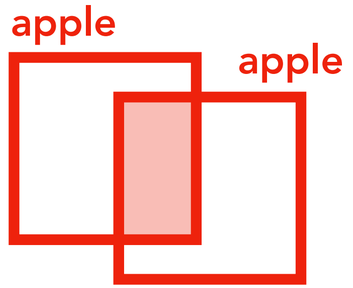
Stride length



Decision threshold

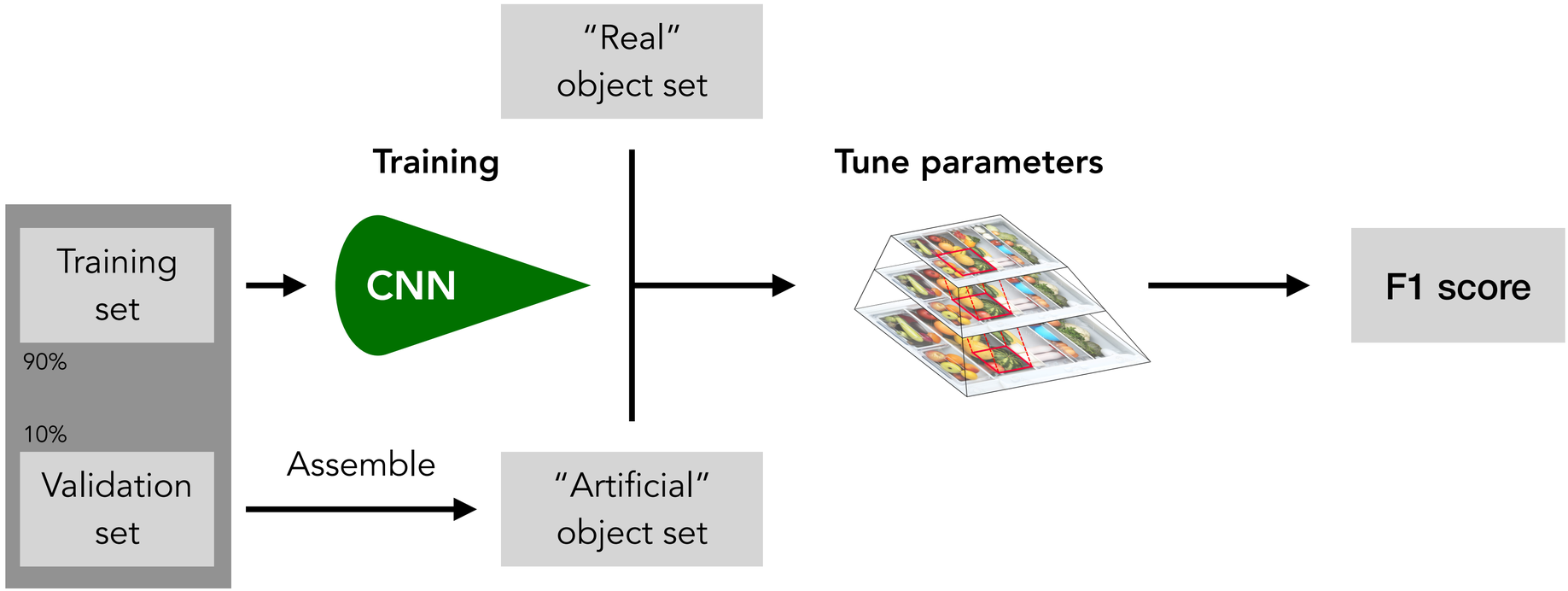
90%

Maximum overlap
(non-max suppression)



Google Cloud

End-to-end training



Final model

Best model:

- **InceptionV3** / ~95% accuracy
- Decision threshold 87%
- Only one object size with scaling factor 1.5 (for input image of size 1024)
- Maximum overlap 20%

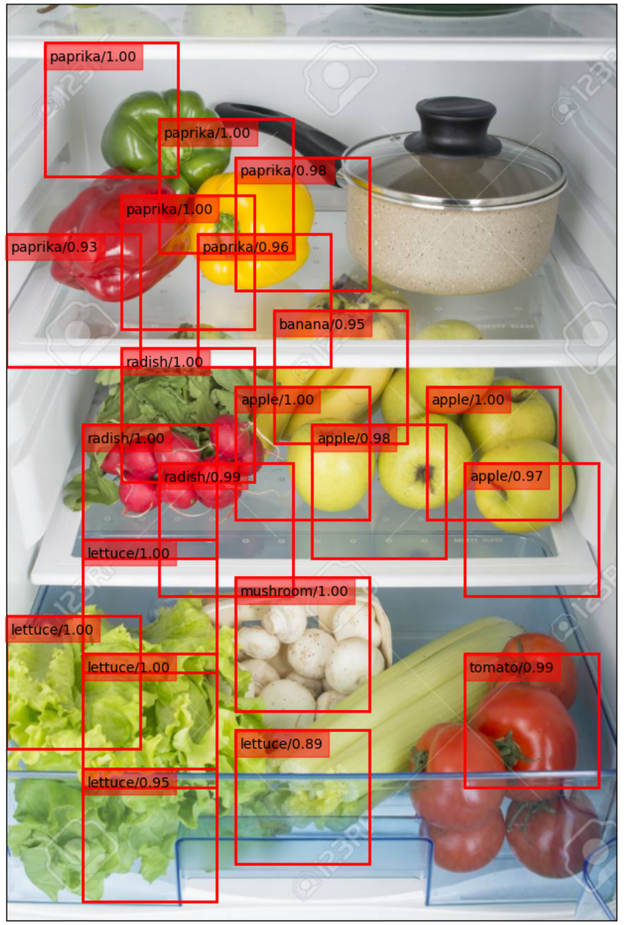
Performance metrics on training set:

F1: 77.2 %

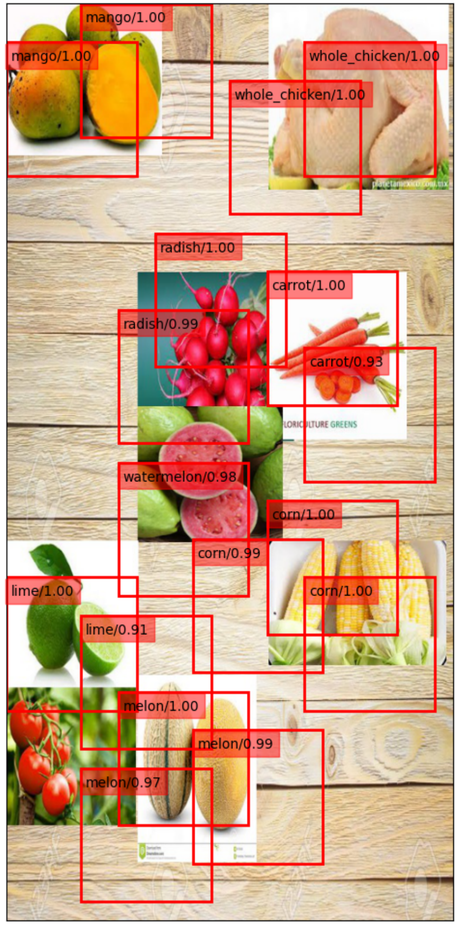
Precision: 76.1 %

Sensitivity: 79.1 %

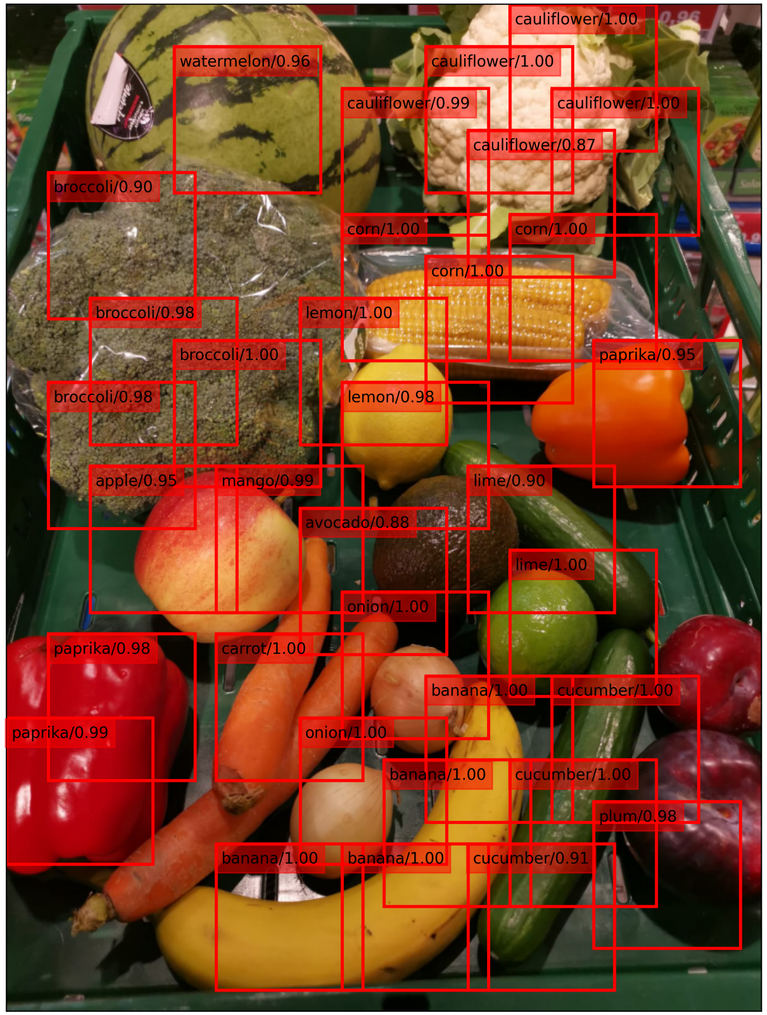
"Real" training example



"Artificial" example



Test set

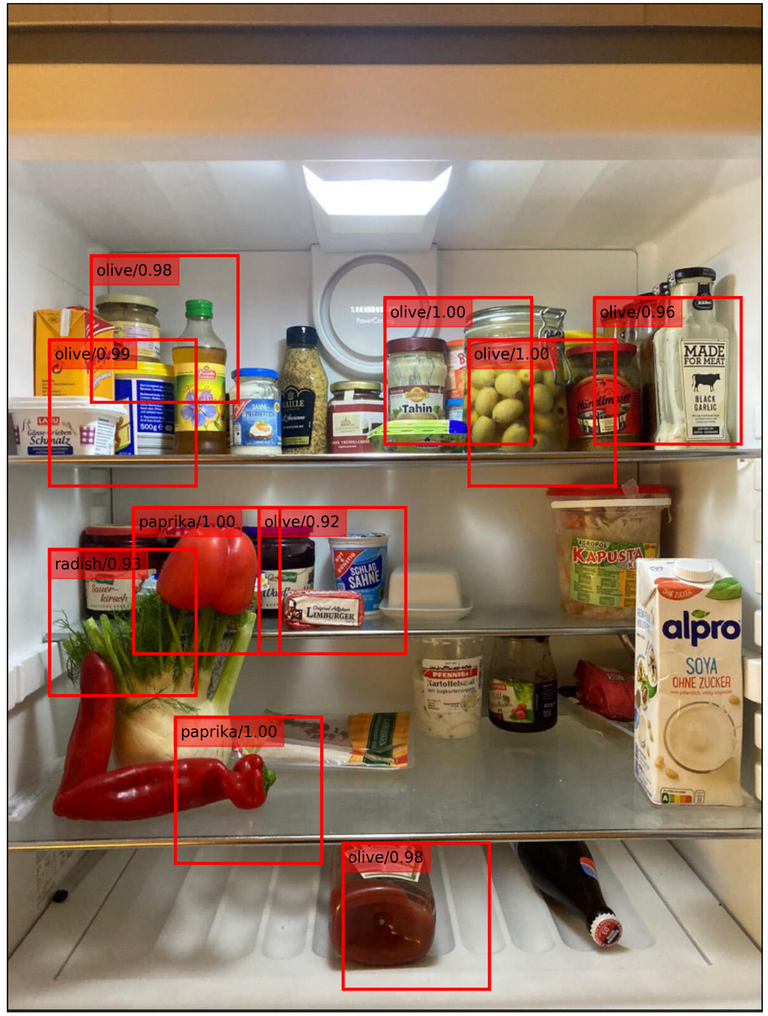


43 photos from EDEKA

F1: 81.3 %

Precision: 77.7 %

Sensitivity: 87.0 %



Recipe suggestion

Recipe 1M+ Dataset (1 million recipes with URLs):
https://github.com/schmidtdominik/RecipeNet

Fridge

banana
tangerine
cheese
pear
lettuce
radish
tomato
celery
paprika

Suggested Recipes

Farmer's salad

lettuce
tomato
paprika
celery
cheese
radish
oil
vinegar

Cut the lettuce leaves crosswise into thin shreds and place in a large bowl. Core and dice the tomatoes; add to the bowl. Dice the celery, and add to the bowl. Cut paprika add to the bowl. Trim and thinly slice the radishes, and add to the bowl. Dice the cheese and add to the bowl. Drizzle with the oil and vinegar. The salad is best served the day it's made.

Bulgarian Shopska salad

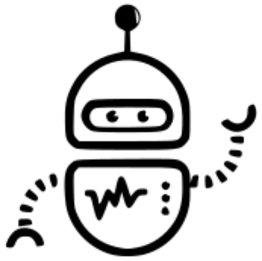
tomato
cucumber
paprika
onion
parsley
oil
vinegar
salt
pepper
cheese

Gather the ingredients. In a large bowl, place tomatoes, cucumber, peppers, onion and parsley, and toss.
Place oil, vinegar, salt, and pepper to taste in a screw-top jar. Cover and shake until well blended.
Toss dressing with vegetables, turn into a serving bowl, and refrigerate until ready to serve.
Top with crumbled cheese and portion on chilled plates. Serve with hearty bread and a glass of rakia.

Future work

Improve deepfoodie's vision

- More data
  - More classes, e.g. also packaged items
  - More images per class
  - Scraping
  - **Data generation**



- Algorithm is vulnerable to **image quality**
  - Use data "augmentations" to simulate bad lighting conditions or image quality

- Different algorithm (**YOLO** or **SSD**)
  - Sliding window algorithm is vulnerable to object occlusions and cannot count objects.
  - **But:** Need to generate a new dataset

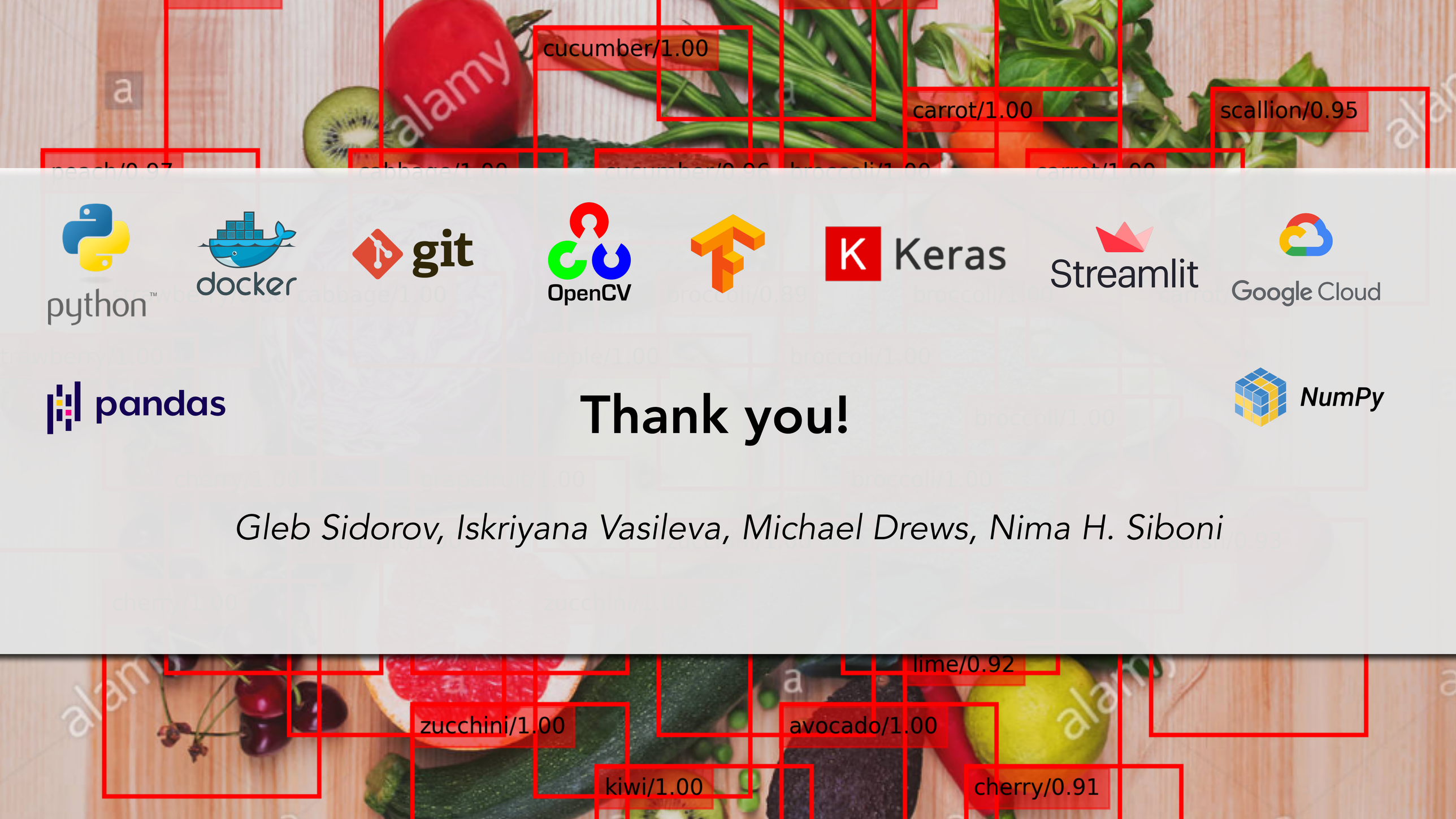
Cooking assistant

- Publishing the recipes database with Co2 emission, H2O-consumption, and nutrition info.

- Suggest recipes based on …
  - Nutrients / Calories
  - Environmental aspects
  - Cost
  - Preparation time

- **Generation of new recipes** instead of look-up table
  - Use generative networks to define similarity of recipes in a latent space

Live Demonstration



python™ docker git OpenCV Keras Streamlit Google Cloud

pandas NumPy

Thank you!

Gleb Sidorov, Iskriyana Vasileva, Michael Drews, Nima H. Siboni