



Collecting user annotations for training neural network-based object detectors

◆-----◆
Presentazione della prova finale di:

Tommaso Redaelli
Matricola 830442

Relatore: *Prof. Gianluigi Ciocca*

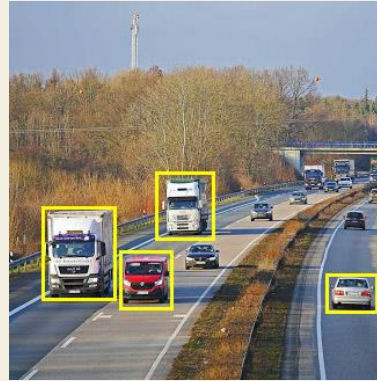
Co-relatore: *Dr. Luigi Celona*

Anno Accademico 2019 – 2020

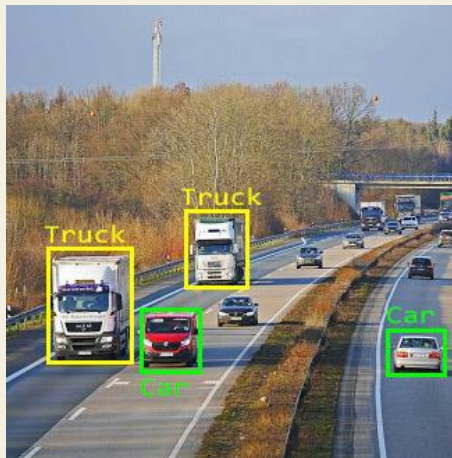
What is object detection



Classification



Localization







Detection

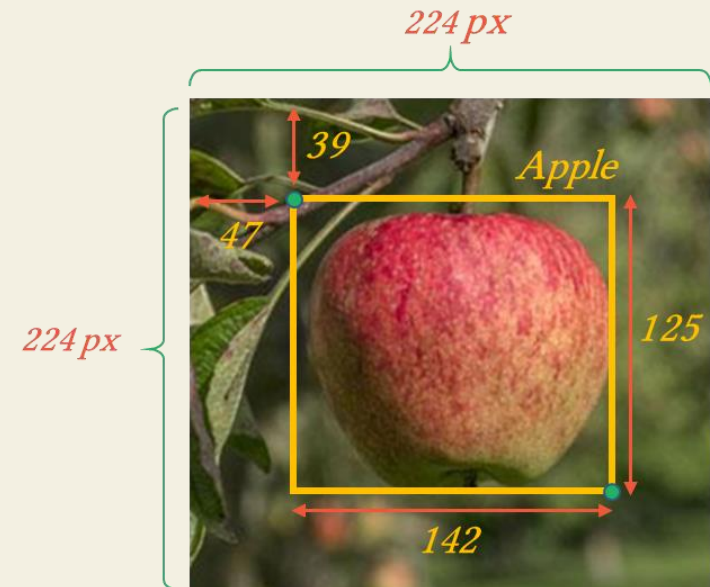
WHAT and WHERE the objects are

Neural network approach

“Deep learning neural network models learn to map inputs to outputs given a training dataset of examples.”

Single example of training:

- ▶ Image  
- ▶ Ground-truth:
 - Bounding box  $[47, 39, 142, 125]$
 - Label  «Apple»



Idea and goal

1

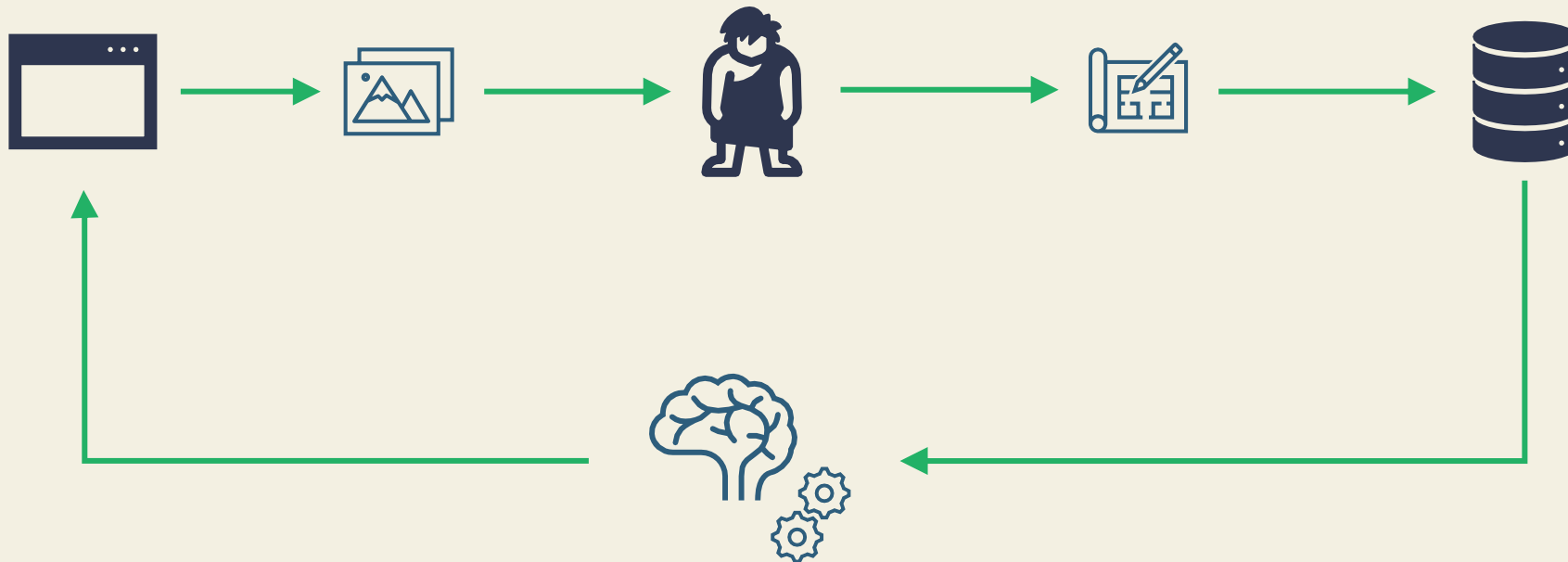
Provide suggestions
on depicted objects

2

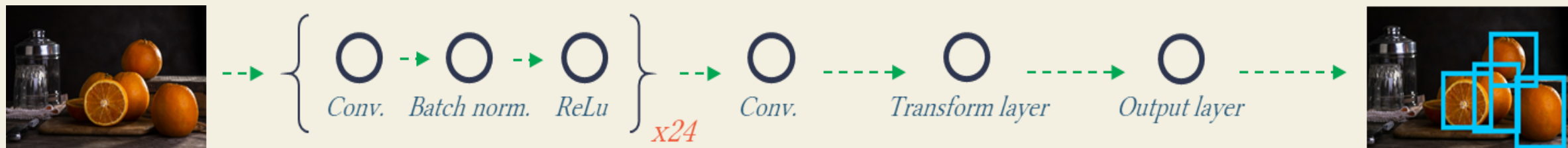
Allow user to validate
and add new detections

3

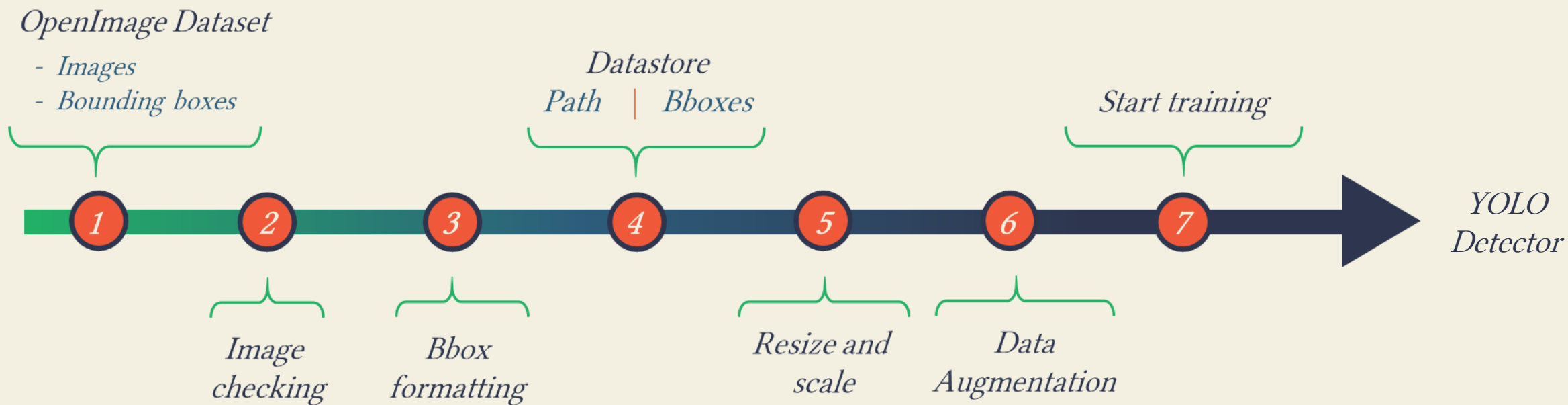
Use annotations to retrain or
build new neural networks



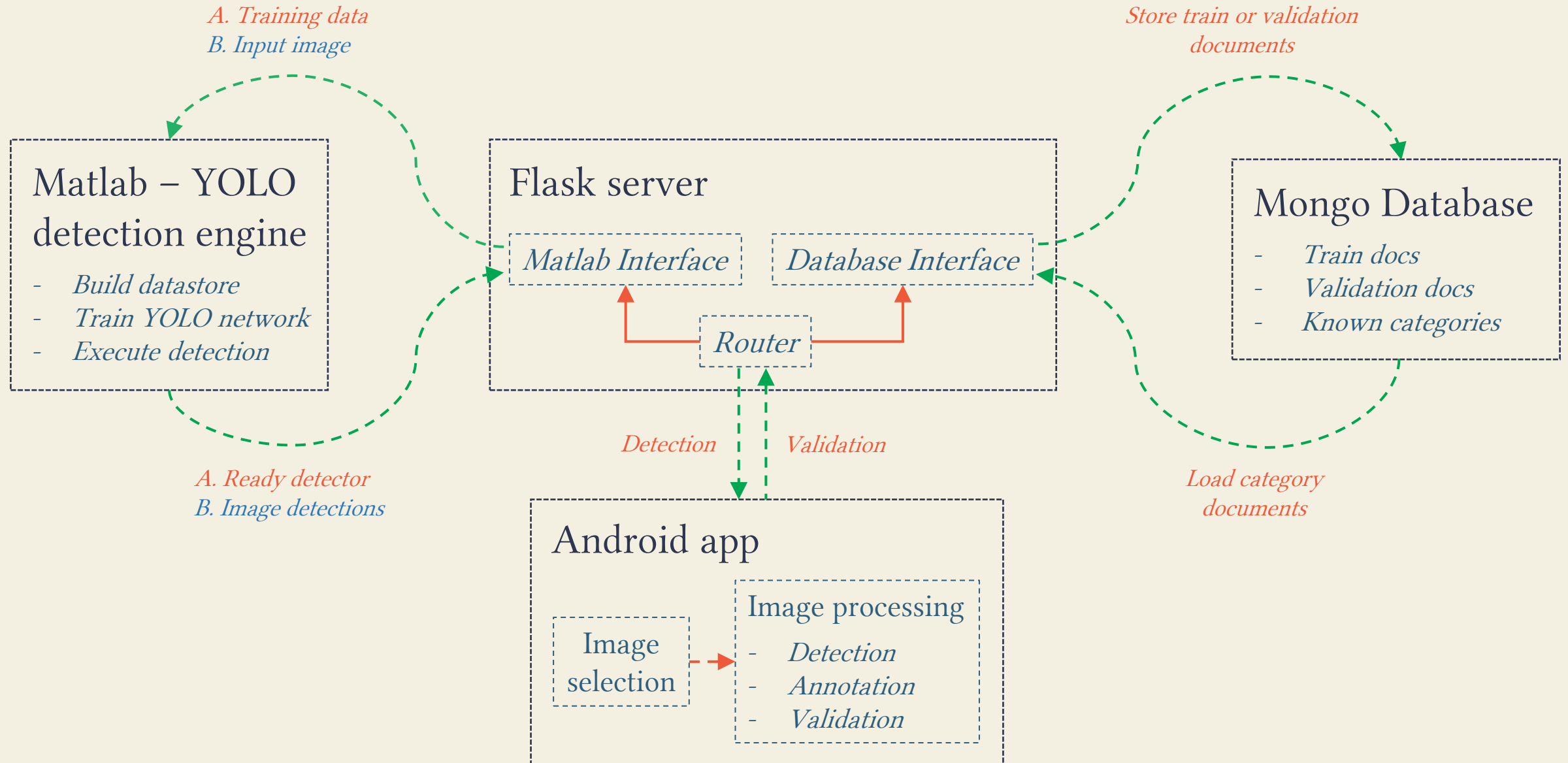
YOLO neural network



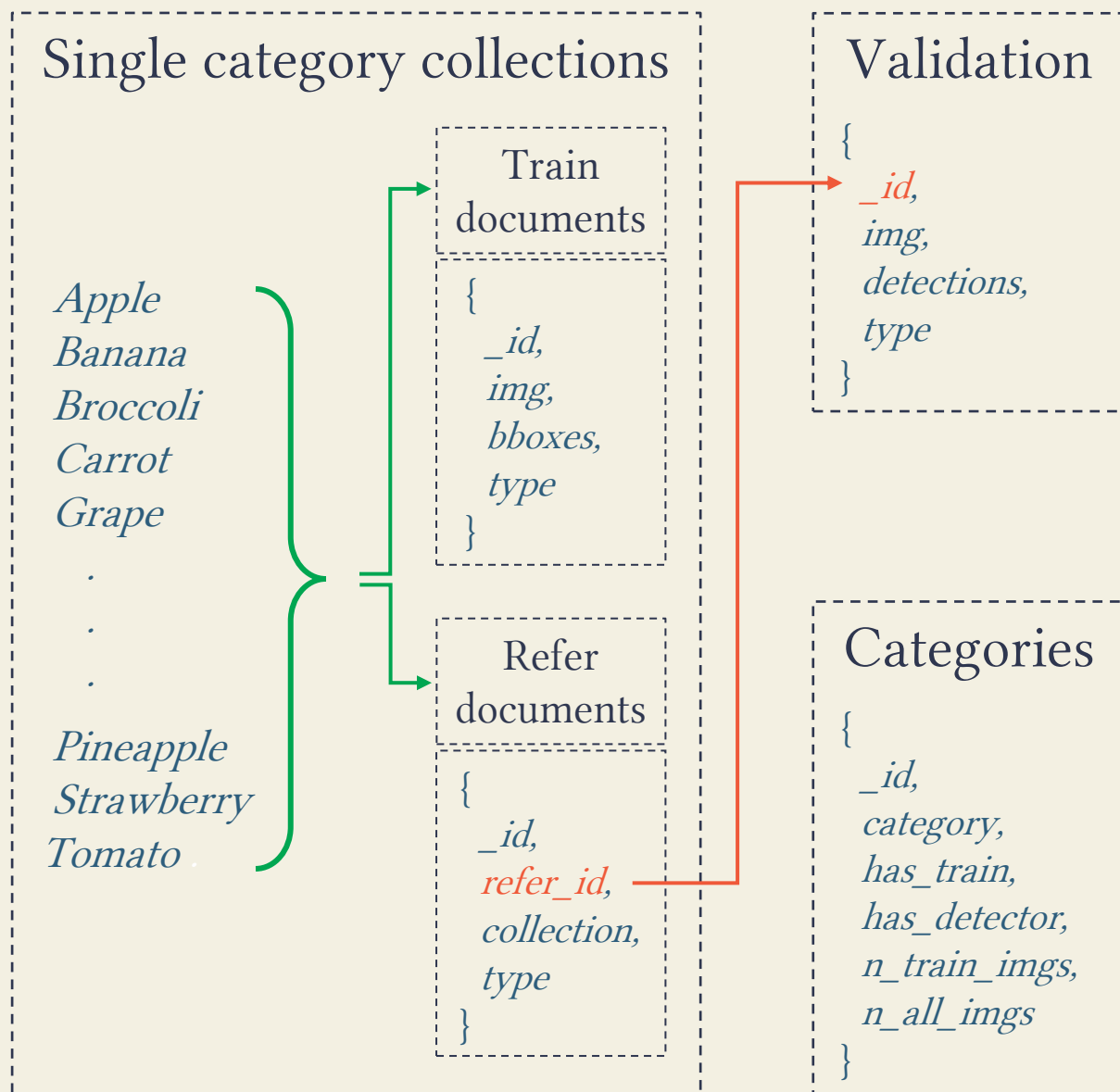
Training process



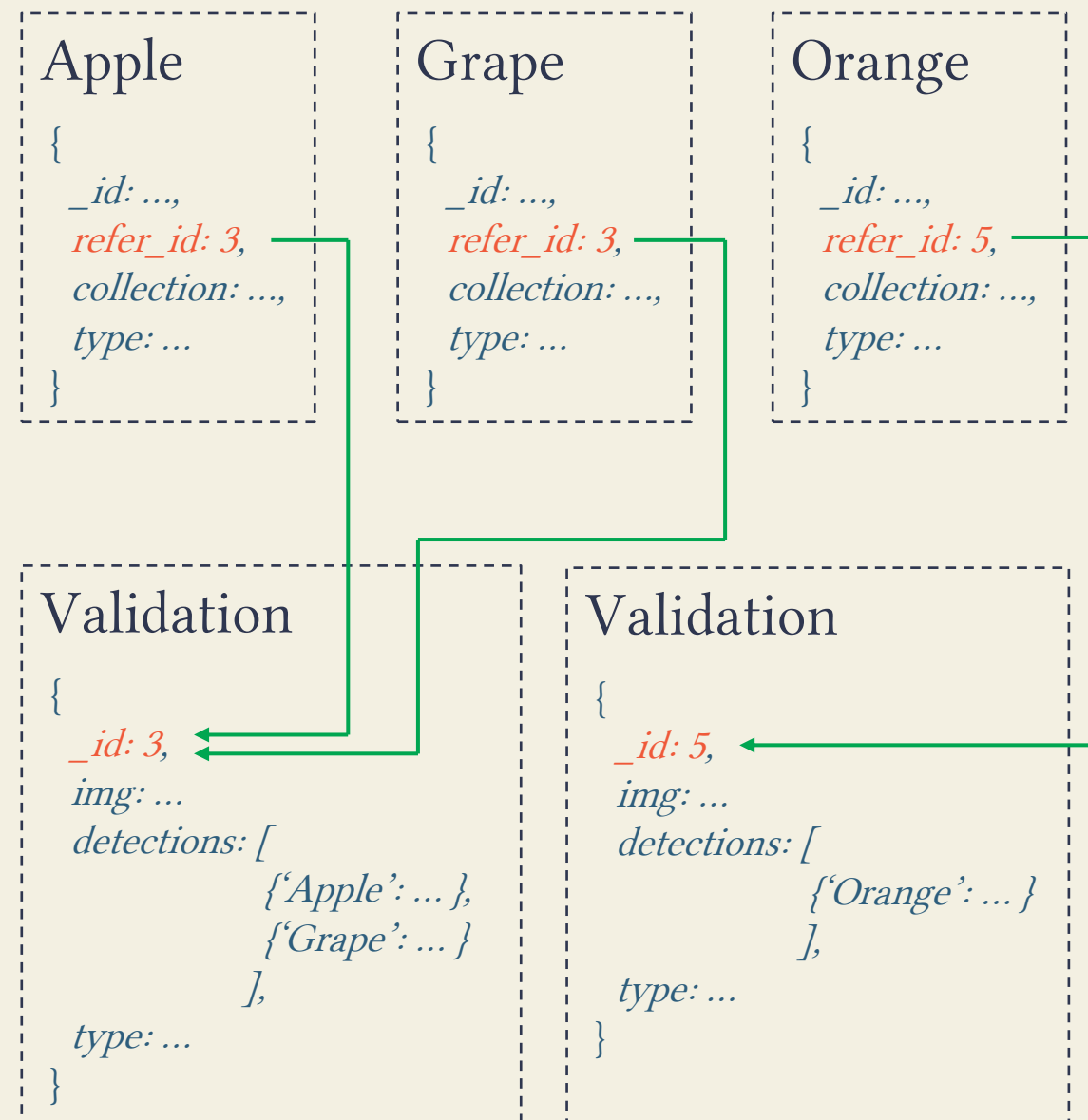
System structure



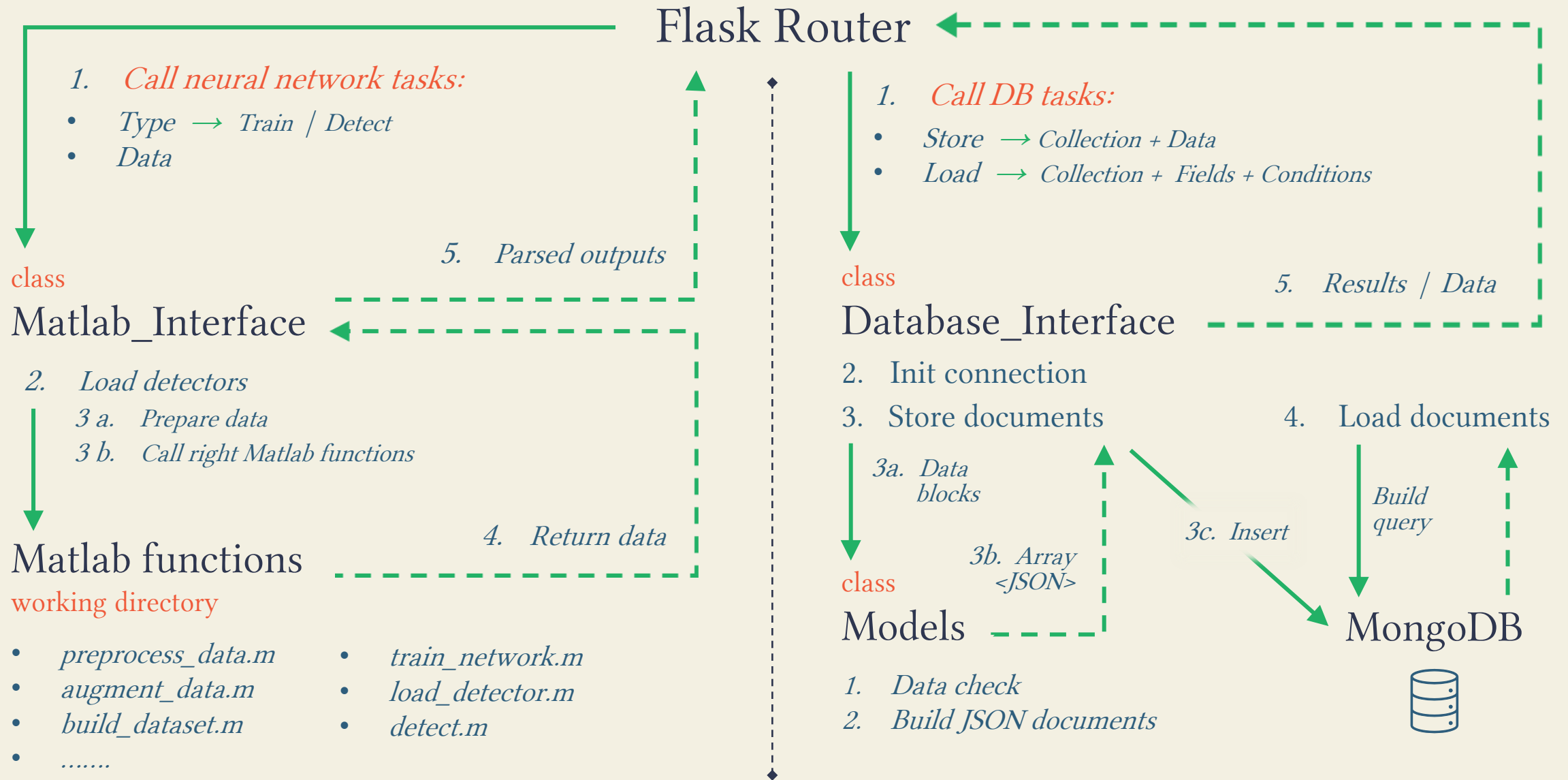
Mongo database



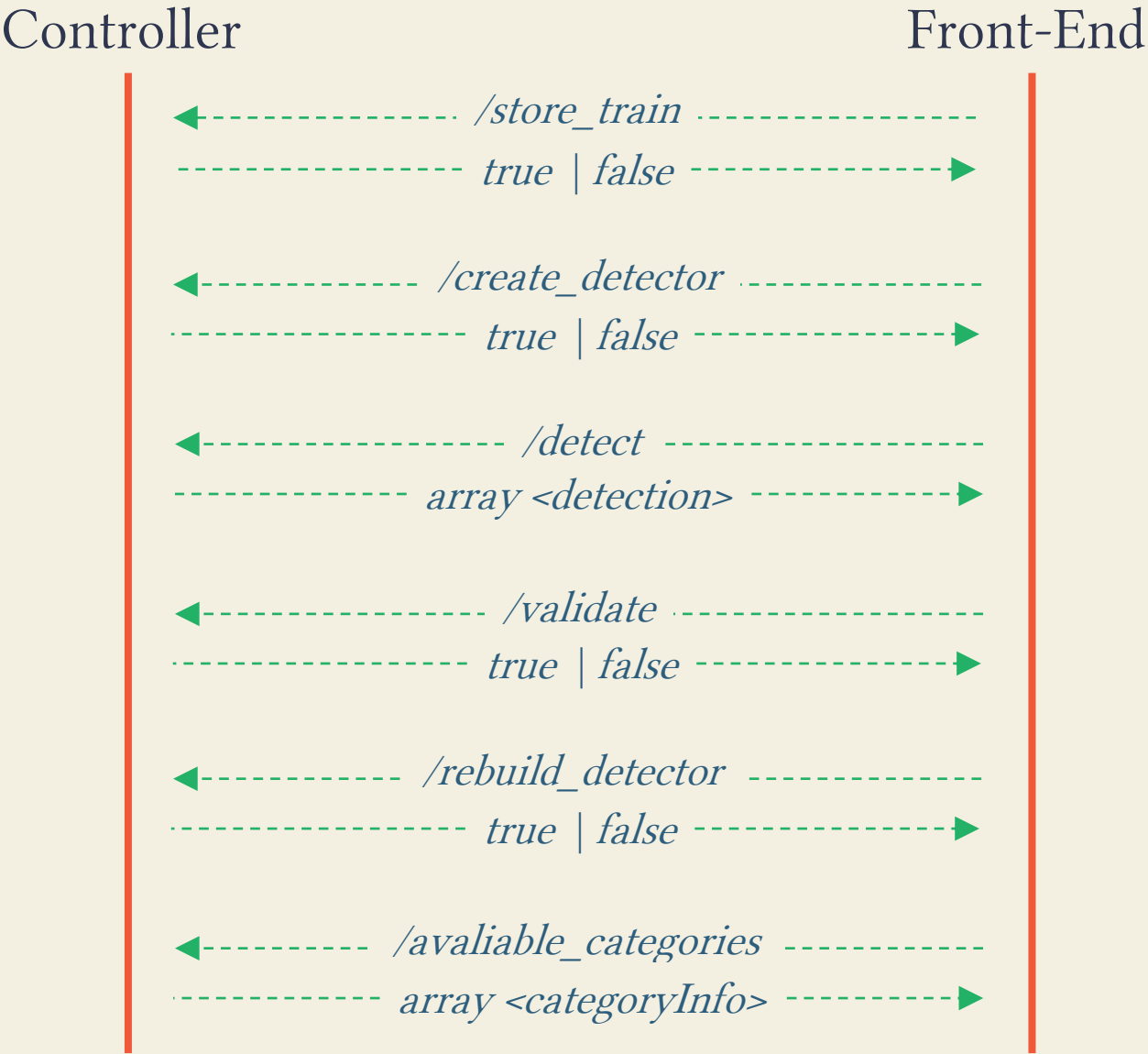
Reference mechanism



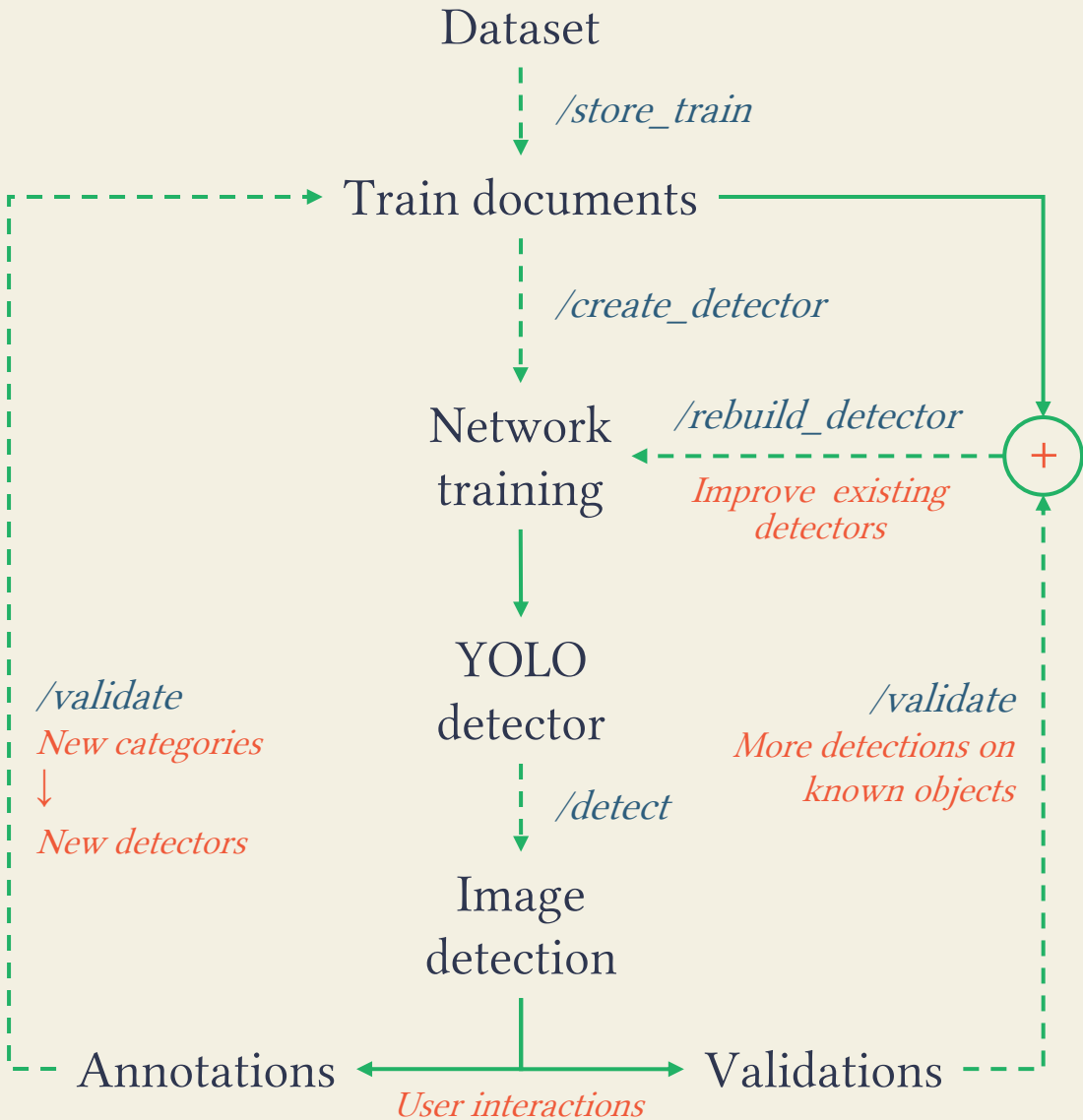
Server – Components integration



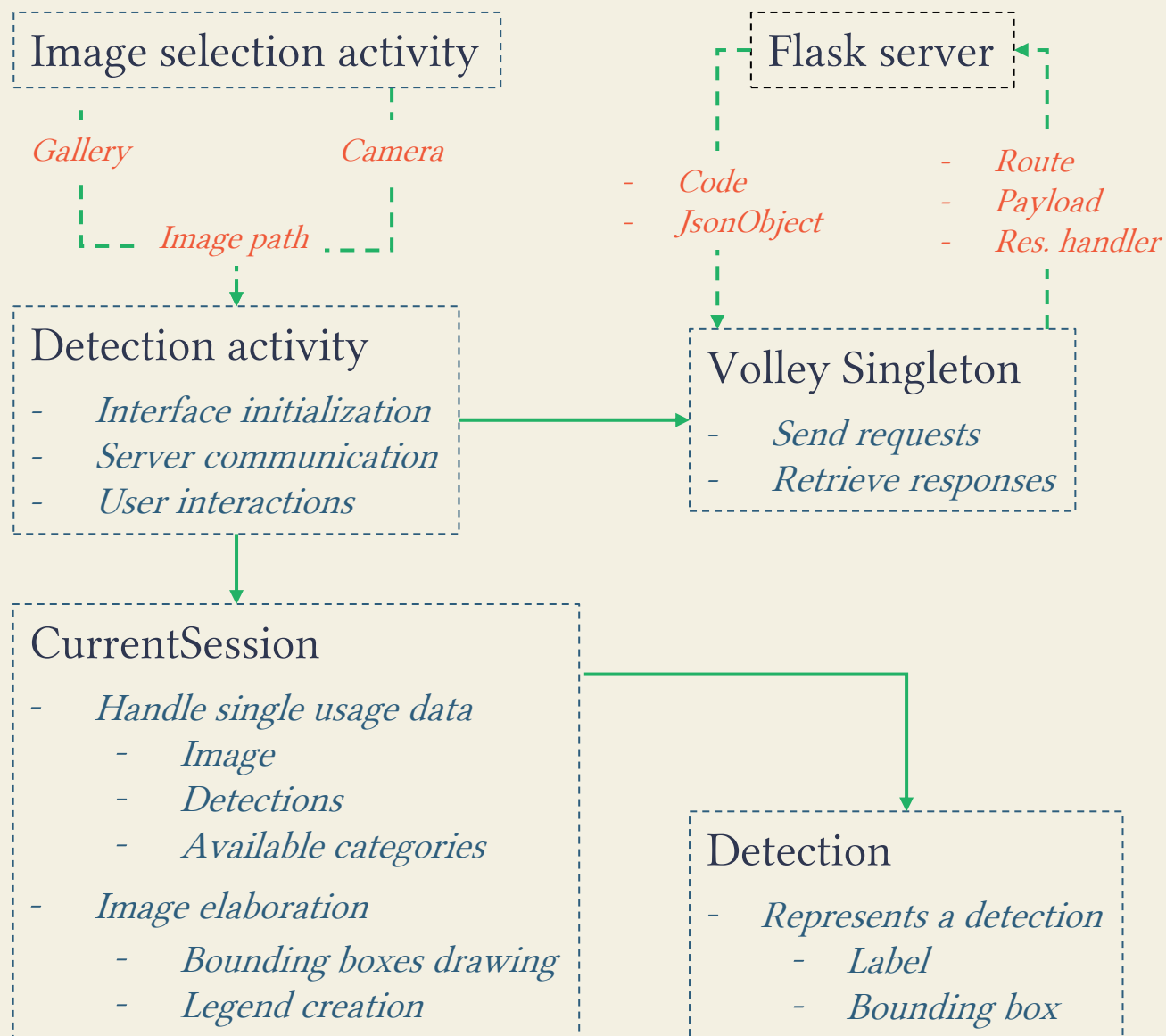
Server – Routing



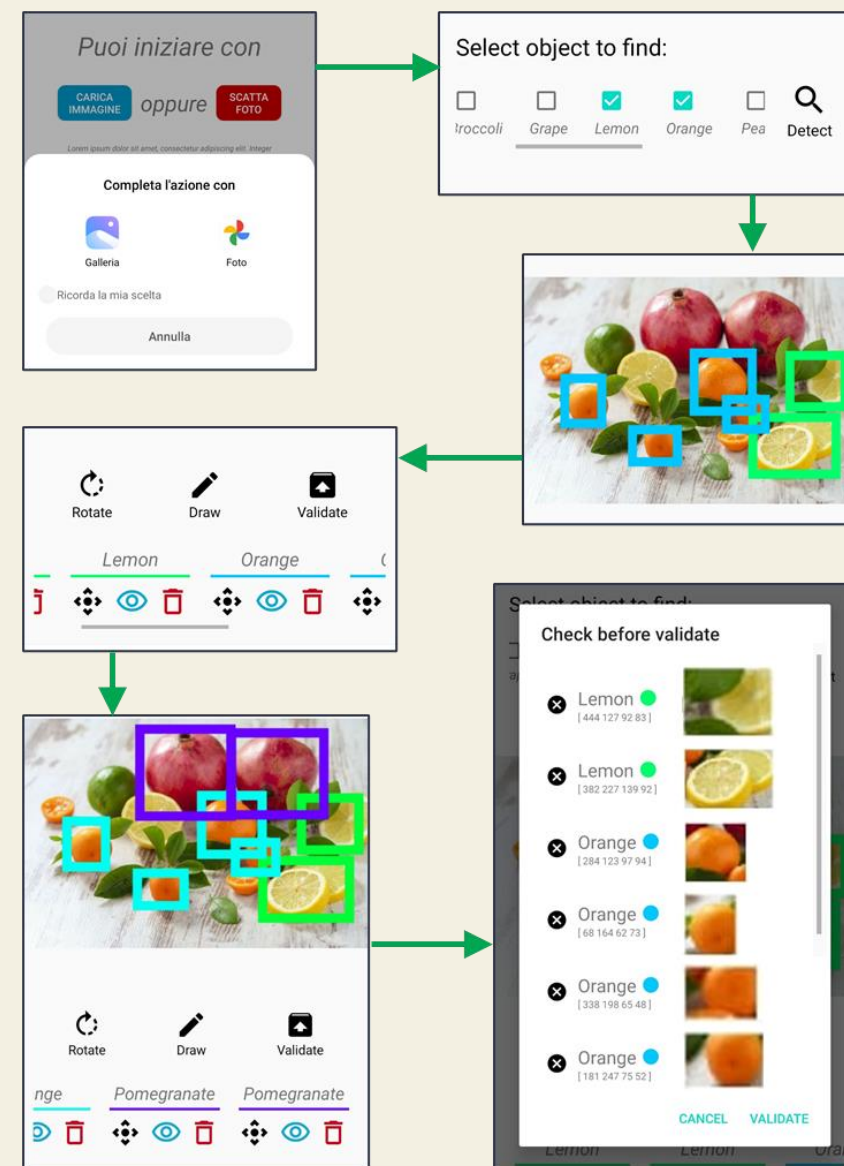
Complete use cycle



Android app



Typical usage flow



Detections

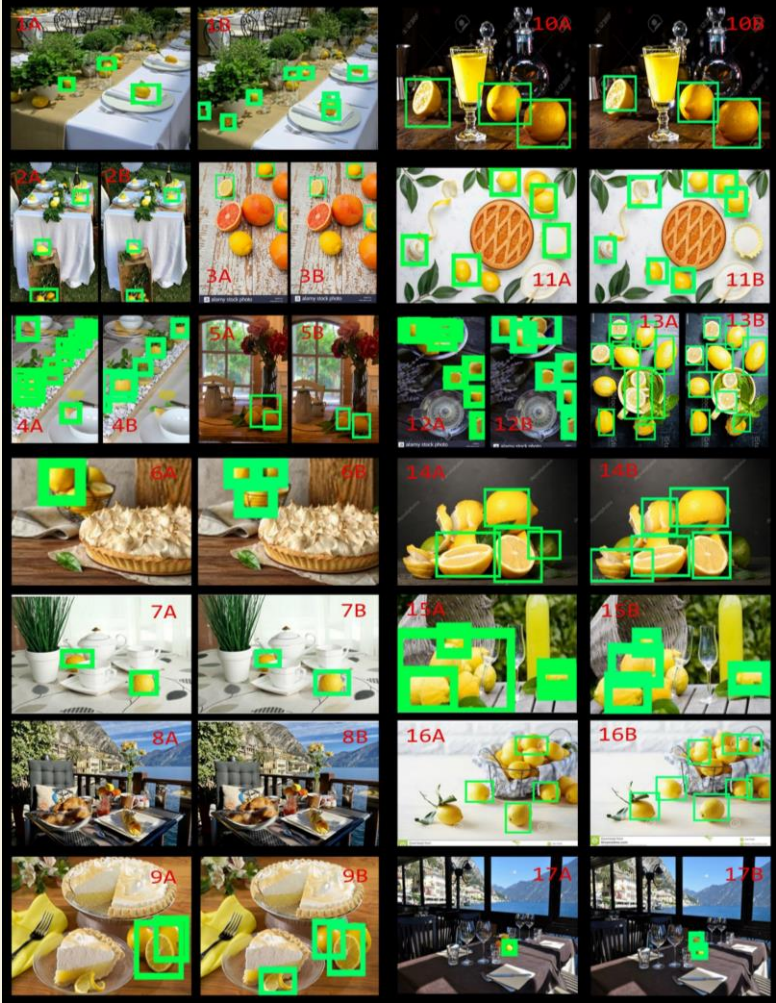


- Apple
- Banana
- Broccoli
- Grape
- Lemon
- Orange
- Pear
- Pineapple
- Tomato

Accuracy



Retrain

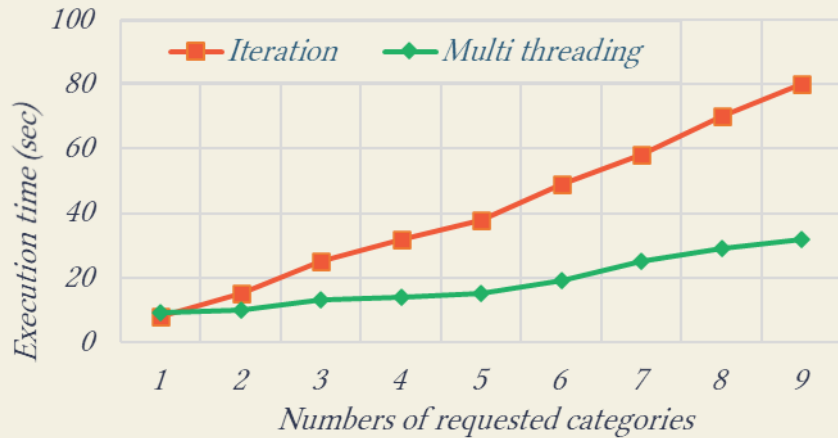


Starting Detector [A]
Trained on 209 imgs -> 90 more annotated imgs
Accuracy: 44%

Retrained Detector [B]
Accuracy: 70%

Limits

- ▶ Detection time (already optimized thorough *multi-threading*)



- ▶ No universal format for category names

- ▶ No check on the honesty of the user annotations

- ▶ Different and unsorted detections for a single object

Possible improvements

- ▶ GPU computing
 - ▶ Real-time
 - ▶ Video detection

- ▶ Guide user through a hierarchy of categories

- ▶ Queue of to-be-annotated images
- ▶ Multiple user validation

- ▶ Sort multiple detection by score
- ▶ Non-max suppressor