Object recognition

1. Train classifier
   • Several different algorithms
   • Annotated images
   • Feature extraction etc.

2. Suggest proposal areas
   • Several different algorithms

3. Classify proposal areas
   • Several different algorithms
Example (first try)

- Transfer learning (Alexnet)
  - Deep convolutional neural network
  - Trained on millions of annotated images in thousands of categories
  - Retrain the last layers for specific classification task
1. Train classifier

- Class(es) of interest
- Not class(es) of interest
2. Proposal areas

Mask uninteresting areas

Here
Based on red background of the trap
2. Proposal areas

Apply algorithm for proposal areas

Here
Selective Search
2. Proposal areas

Select only interesting boxes

Here based on size and aspect ratio
2. Proposal areas
3. Classify boxes

Apply previously trained classifier per box

Here
Classifier trained on 4 species and background
3. Classify boxes

Clean predictions

Here
Based on overlap of boxes and probability of class
4. Compare classification to ground truth

- Correctly classified as Suzukii (true positive)
- Fly missed
- Suzukii classified as other species
- Species mixed up
Some observations

- Miss classifications
  - Mixed up classes (flies, trap)
  - More training data / classes

- Flies missed
  - Selective search algorithm not good / not good parameterized
Second try

1. get proposal areas

- Red filter
- Inverse
- Select pixel clusters
- Find cluster centers
- Draw boxes
Still some problematic areas:

- edges
- overlap
- shadow

(for later)
Precision = \#correctly classified / \#classified as Suzukii
Recall = \#correctly classified / \#ground truth Suzukii

Green box = Ground truth (location of Suzukii); Blue box = False predicted Suzukii; Green circle (green and blue overlapping box) = correctly predicted Suzukii
What’s better?

• Almost always just 2 classes
Sometimes it works, sometimes it doesn’t

• Difficult lightning conditions
• Distance to trap
• Image at angle
Photo requirements

• Resolution (Now: 5472 x 3648 px)
• Sensor target distance
• View angle
• Lightning conditions (shadows, sun glint)
• Sharpness
• Background (trap)
• ...

...
Data requirements

- Annotated images (Johannes, 2018)
  - Static images
  - Drone-based images
- Images by David
- Images by WUR
- ...
Planning

• Other cameras:
  • Mavic: 4000x3000 pix
  • DJI Zenmuse X4s: 5472×3648 pix

• Other than RGB: Hyperspectral / NIR / ...

• Publications

• DSS

• ...