Mask R-CNN in GRASS GIS

Ondřej Pešek

Czech Technical University in Prague
Faculty of Civil Engineering
Department of Geomatics

29. 8. 2019

Table of contents

Motivation

Situation

- higher density of the satellite monitoring systems
- higher density of aerial imagery
- vectorization of analogue maps

Motivation

Situation

- higher density of the satellite monitoring systems
- higher density of aerial imagery
- vectorization of analogue maps
- higher quality

Ondřej Pešek

Motivation

Situation

- higher density of the satellite monitoring systems
- higher density of aerial imagery
- vectorization of analogue maps
- higher quality
- open data

Motivation

Situation

- higher density of the satellite monitoring systems
- higher density of aerial imagery
- vectorization of analogue maps
- higher quality
- open data
- data standardization

Motivation

Common classification methods

- manual classification

GRASS GIS

- manual classification
  - GRASS Digitizing tool

Motivation

Common classification methods

- manual classification
- supervised classification

GRASS GIS

- manual classification
  - GRASS Digitizing tool
- supervised classification
  - g.gui.iclass + i.maxlik

Motivation

Common classification methods

- manual classification
- supervised classification
- unsupervised classification

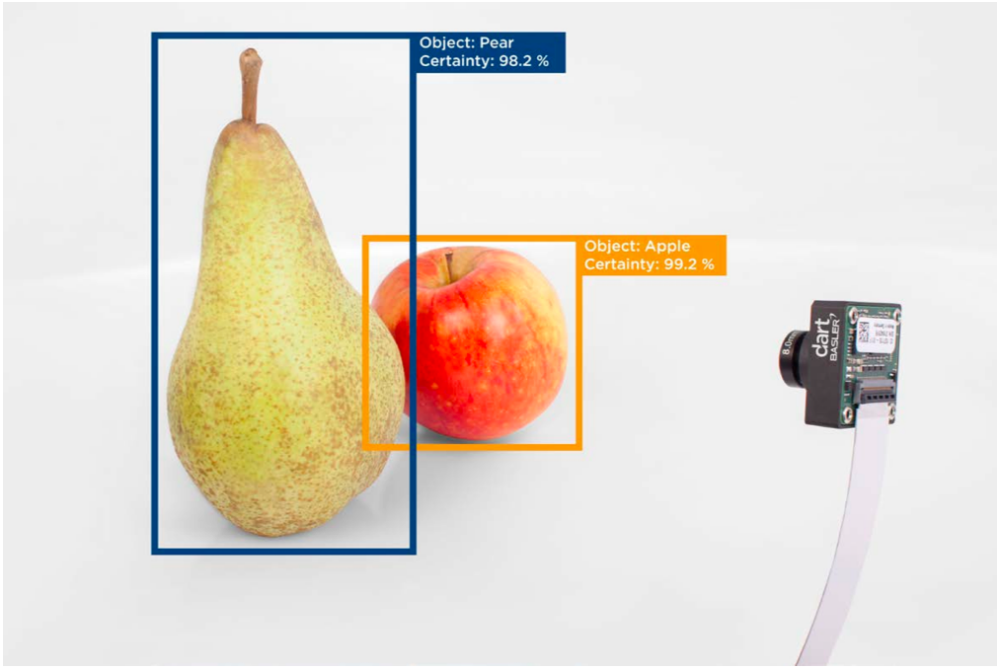
GRASS GIS

- manual classification
  - GRASS Digitizing tool
- supervised classification
  - g.gui.iclass + i.maxlik
- unsupervised classification
  - i.cluster + i.maxlik

Motivation

Why neural networks?

- human brain is the most powerful tool we know
- we are trying to get human-understandable results



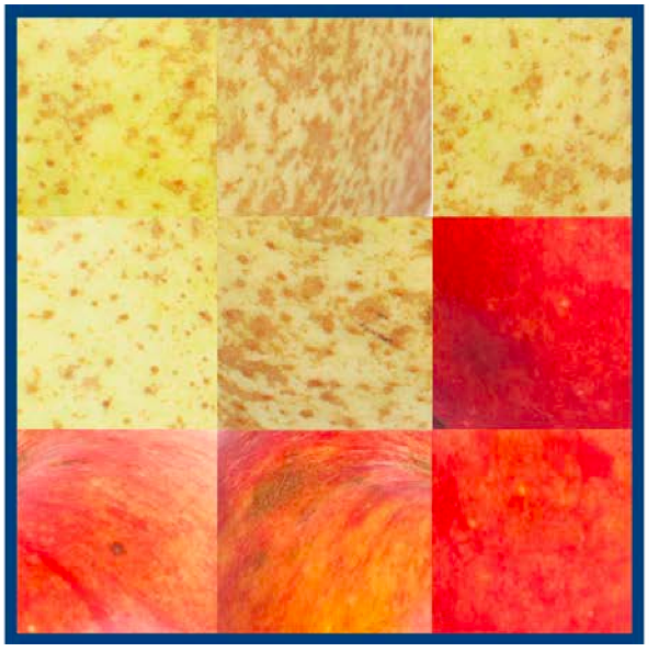
Source: [1]

Motivation

Ondřej Pešek

Why neural networks?

- human brain is the most powerful tool we know
- we are trying to get human-understandable results



Source: [1]

Motivation

Why neural networks?

- human brain is the most powerful tool we know
- we are trying to get human-understandable results

Source: [1]

Convolutional neural networks

105	102	100	97	96
103	99	103	101	102
101	98	104	102	100
99	101	106	104	99
104	104	104	100	98

Kernel Matrix

0	-1	0
-1	5	-1
0	-1	0

89		

105	102	100	97	96
103	99	103	101	102
101	98	104	102	100
99	101	106	104	99
104	104	104	100	98

Kernel Matrix

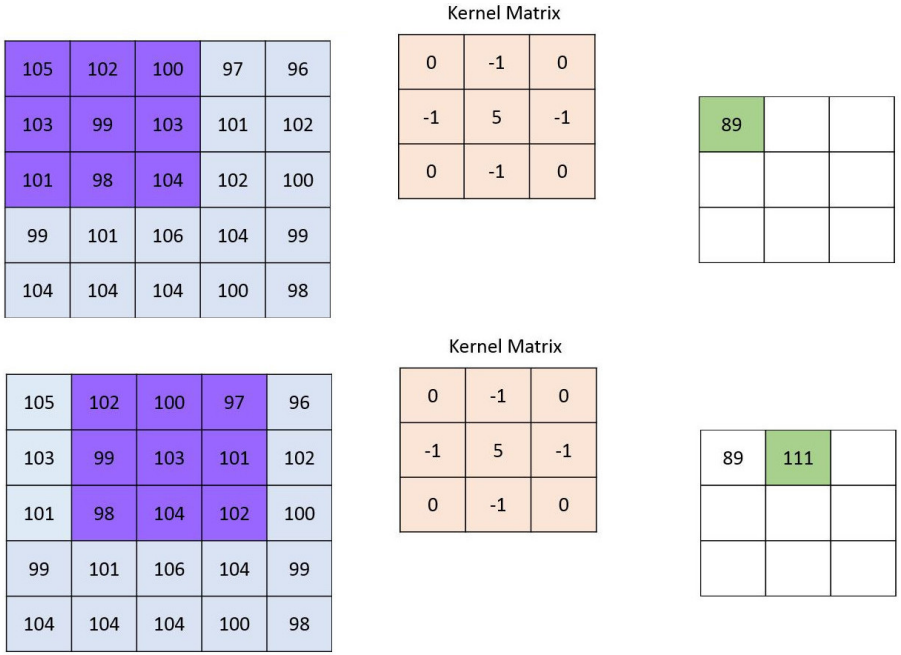
0	-1	0
-1	5	-1
0	-1	0

89	111	

Ondřej Pešek

Convolutional neural networks

Ondřej Pešek



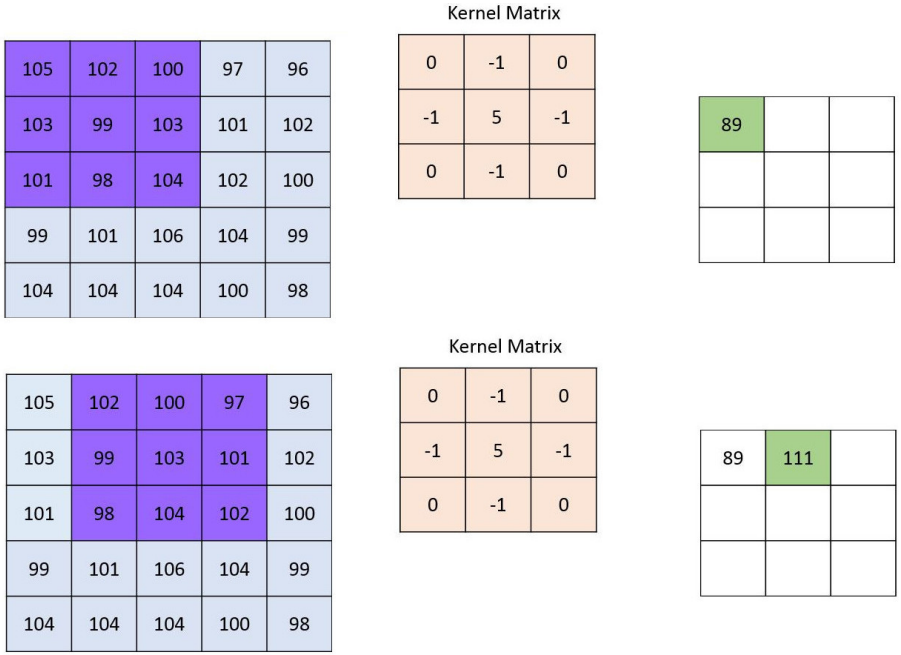
Why convolutional neural networks?

- ResNet got in ILSVRC 2016 top 5 error of 3.6 %

Source: [1]

Convolutional neural networks

Ondřej Pešek



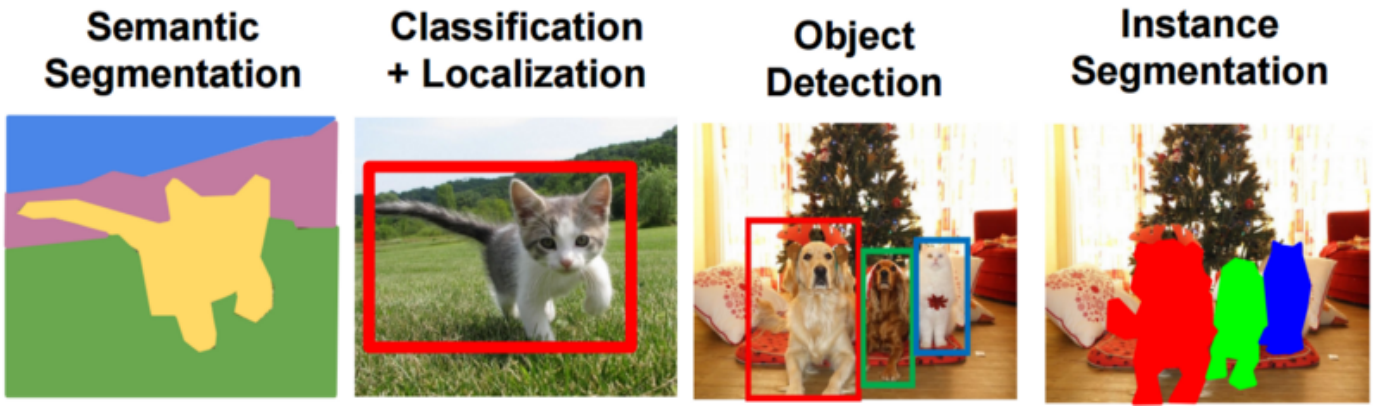
Why convolutional neural networks?

- ResNet got in ILSVRC 2016 top 5 error of 3.6 %
- human 8 %

Source: [1]

Mask R-CNN

Ondřej Pešek

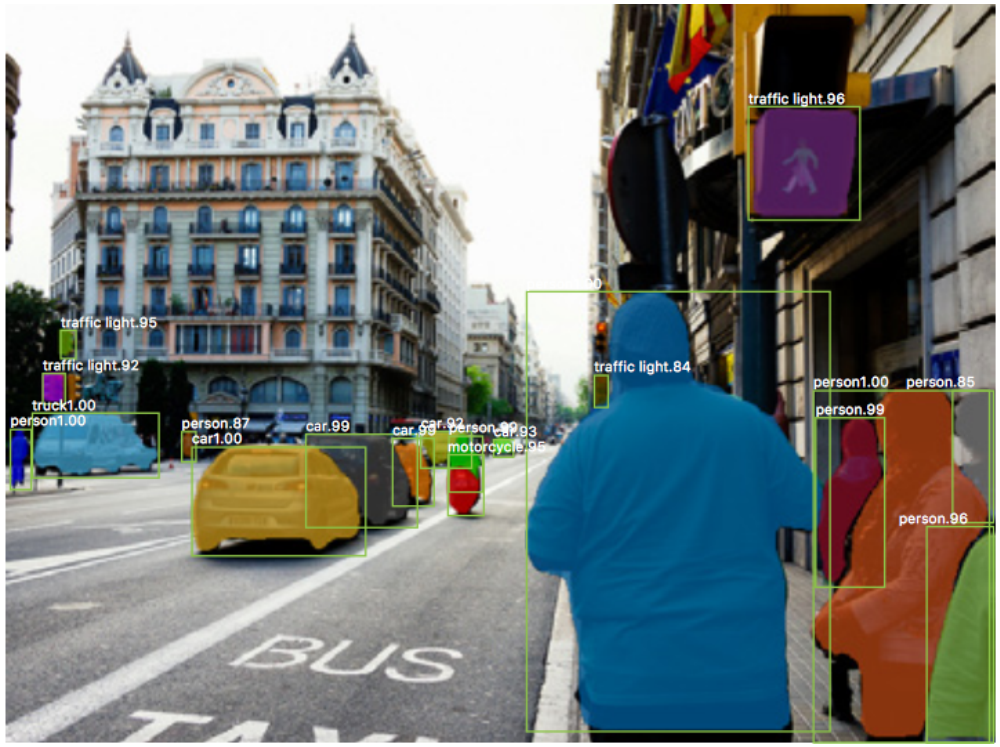


Source: [2]

Mask R-CNN

Ondřej Pešek

Instance segmentation



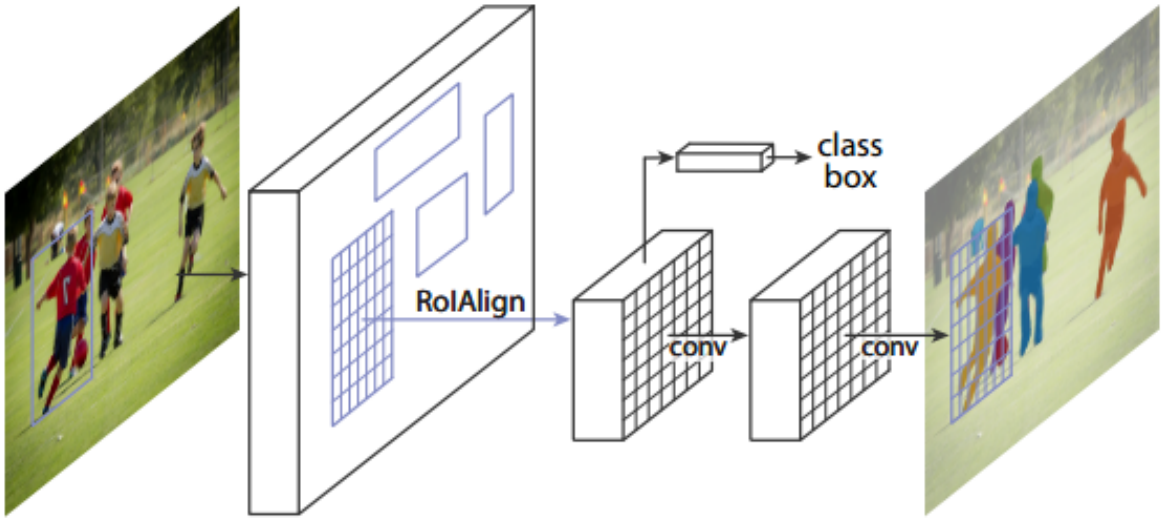
Source: [2]

Mask R-CNN

Ondřej Pešek

Two parts:

- backbone
- head

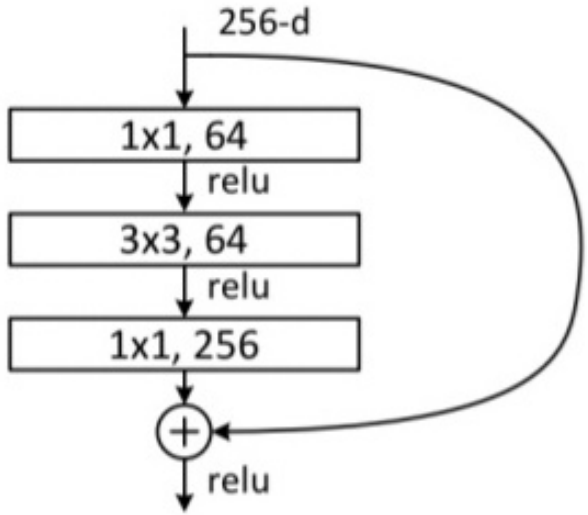


Source: [3]

Mask R-CNN

Backbone architecture:

- ResNet

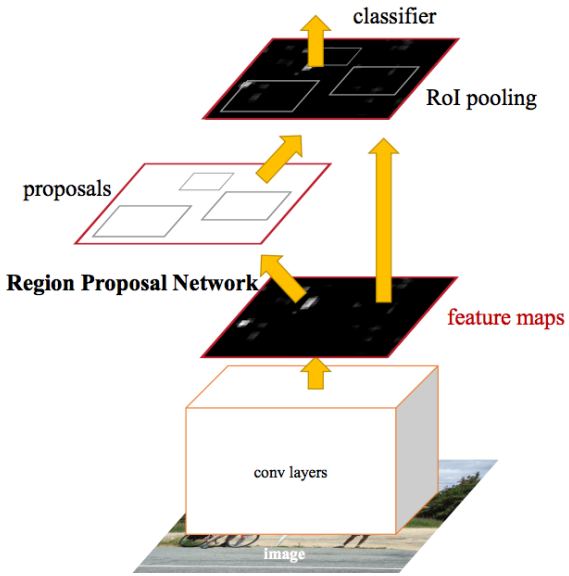


Source: [3]

Mask R-CNN

Backbone architecture:

- ResNet
- RPN

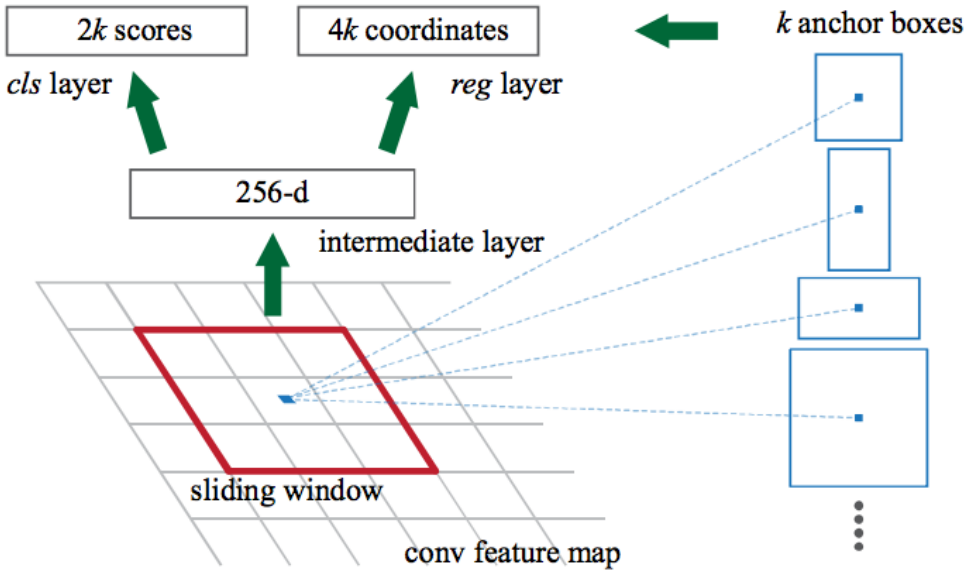


Source: [3]

Mask R-CNN

Backbone architecture:

- ResNet
- RPN

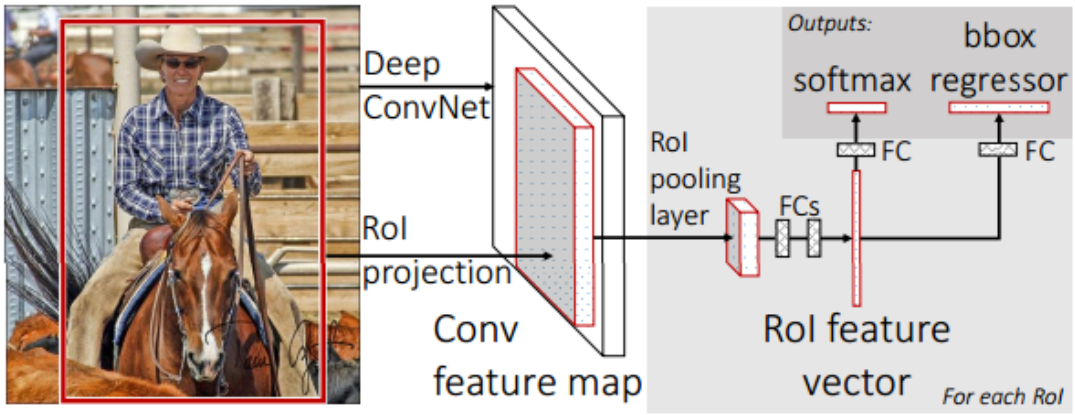


Source: [3]

Mask R-CNN

Head architecture:

- softmax $\rightarrow$ class

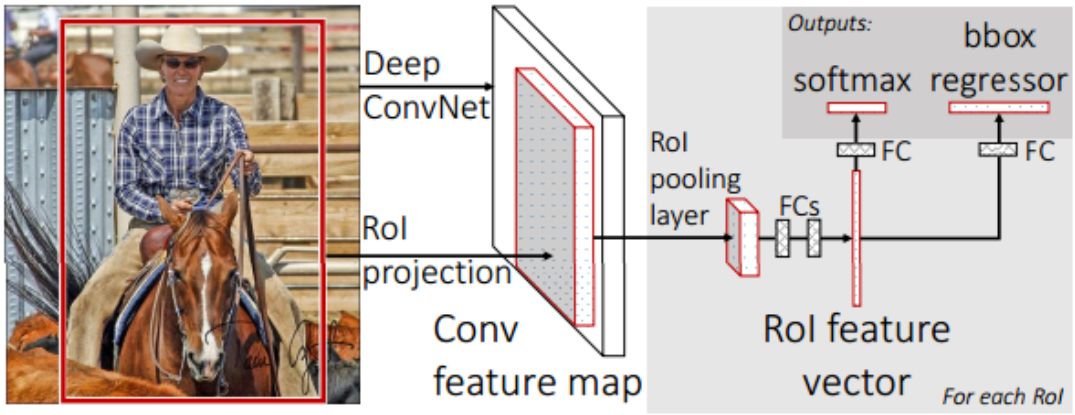


Source: [4]

Mask R-CNN

Head architecture:

- softmax → class
- regression → bounding box

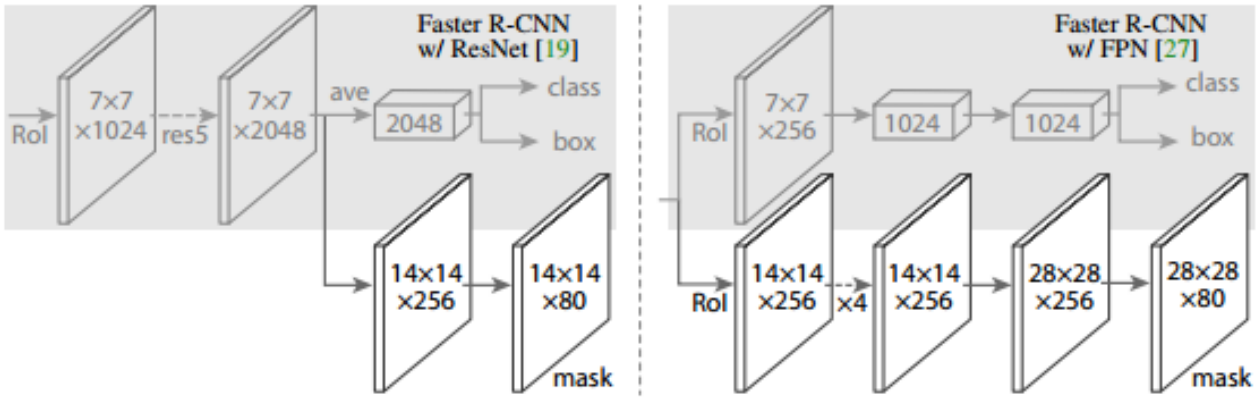


Source: [4]

Mask R-CNN

Head architecture:

- softmax → class
- regression → bounding box
- FCN → mask



Source: [4]

Ondřej Pešek

Usage

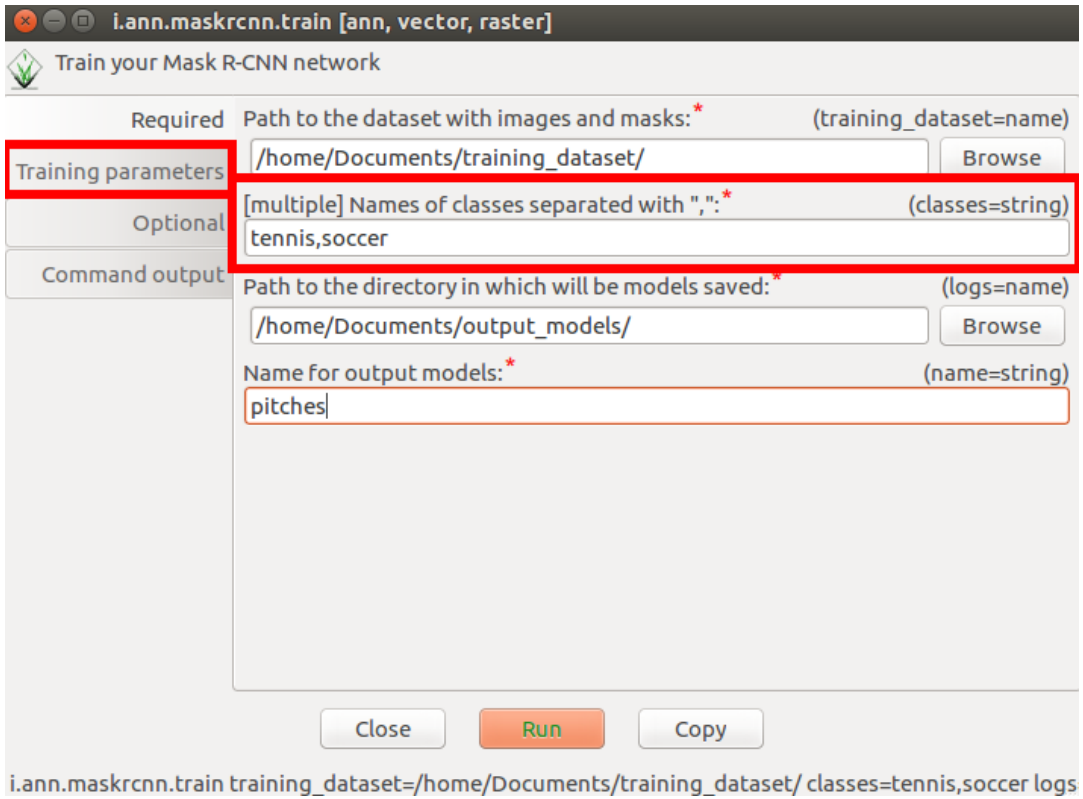
Workflow:

- i.ann.maskrcnn.train
- i.ann.maskrcnn.detect

i.ann.maskrcnn.train

Workflow behind

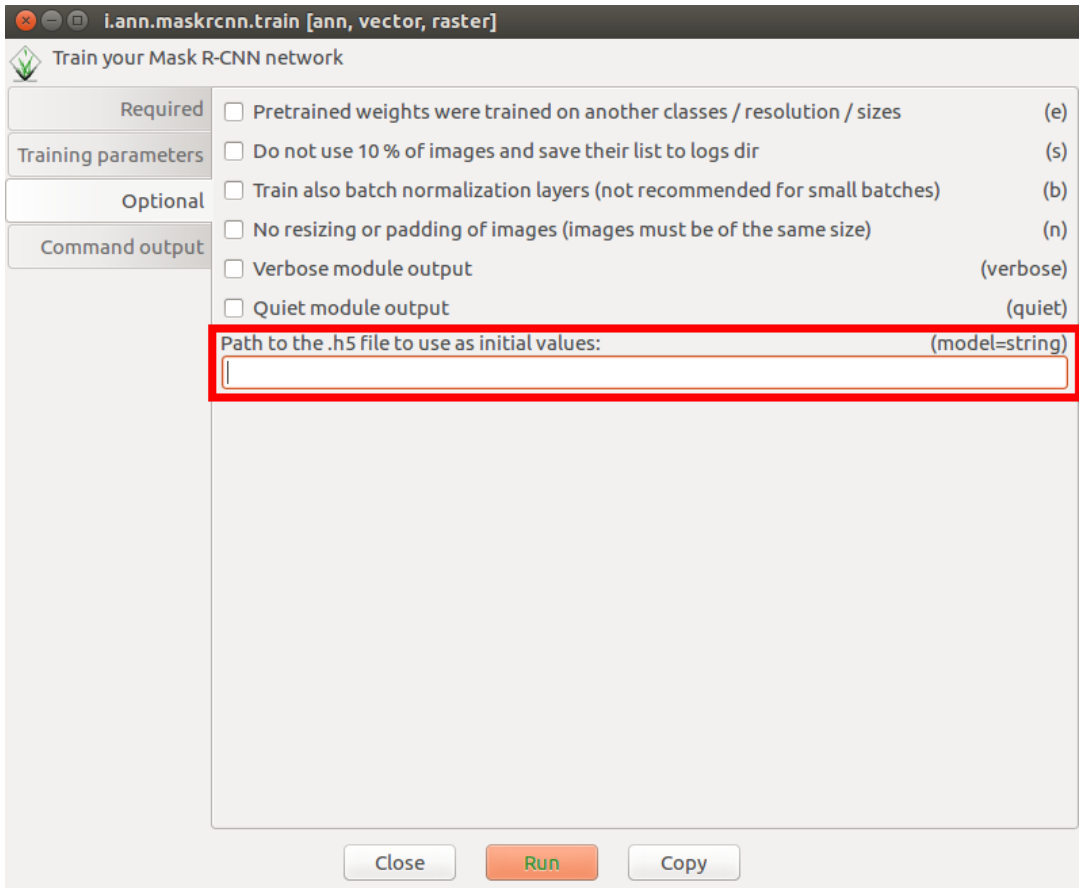
- configuration of the model



i.ann.maskrcnn.train

Workflow behind

- configuration of the model
- feed the model with pre-trained weights

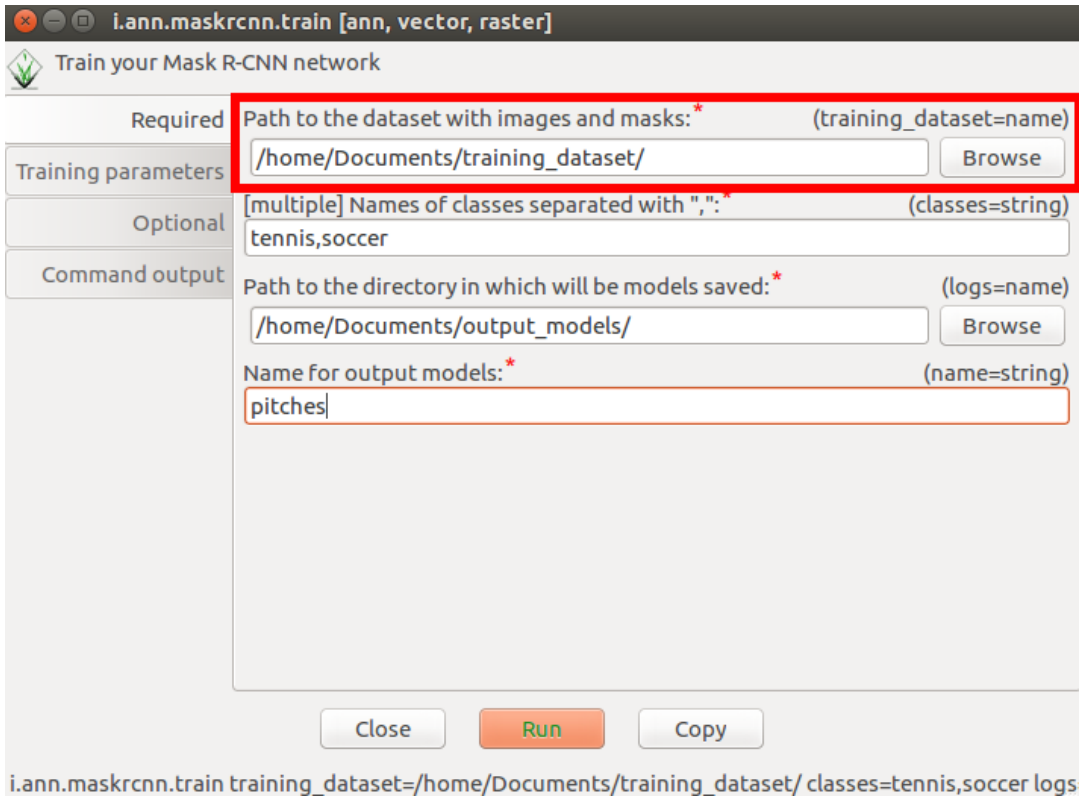


Ondřej Pešek

i.ann.maskrcnn.train

Workflow behind

- configuration of the model
- feed the model with pre-trained weights
- read the training dataset

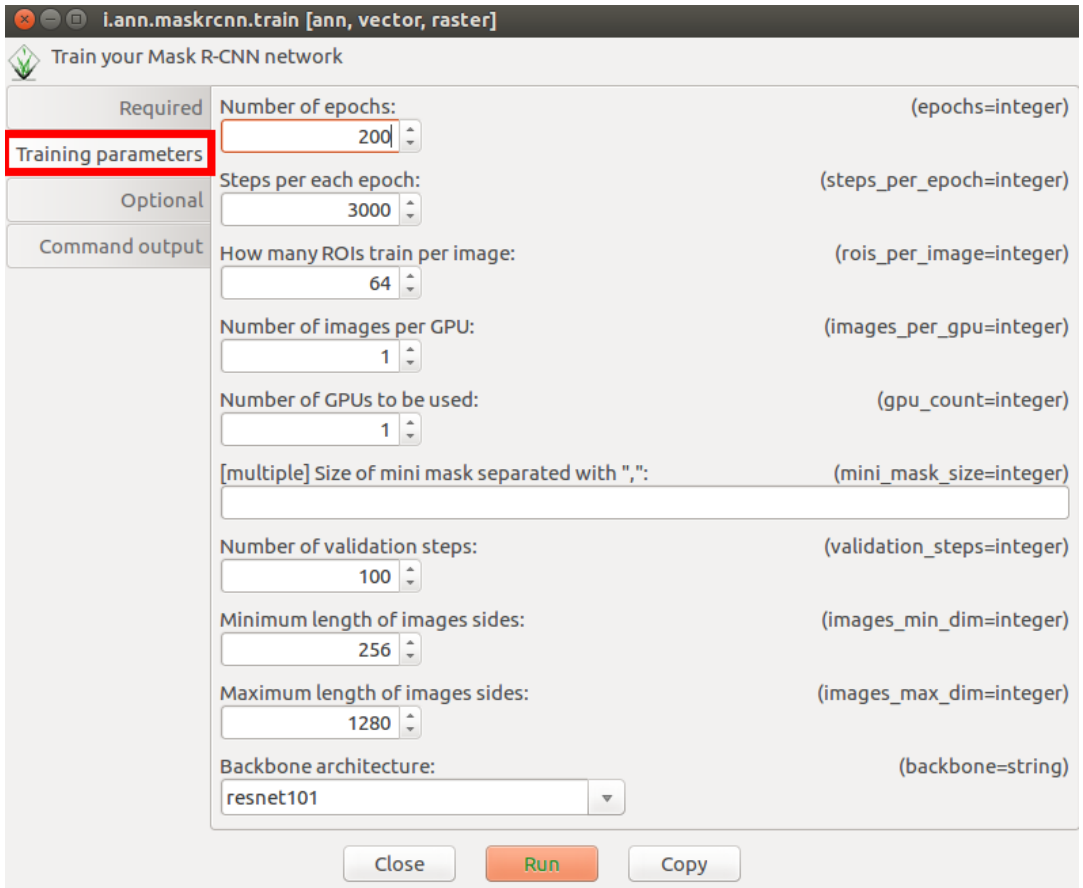


Ondřej Pešek

i.ann.maskrcnn.train

Workflow behind

- configuration of the model
- feed the model with pre-trained weights
- read the training dataset
- train

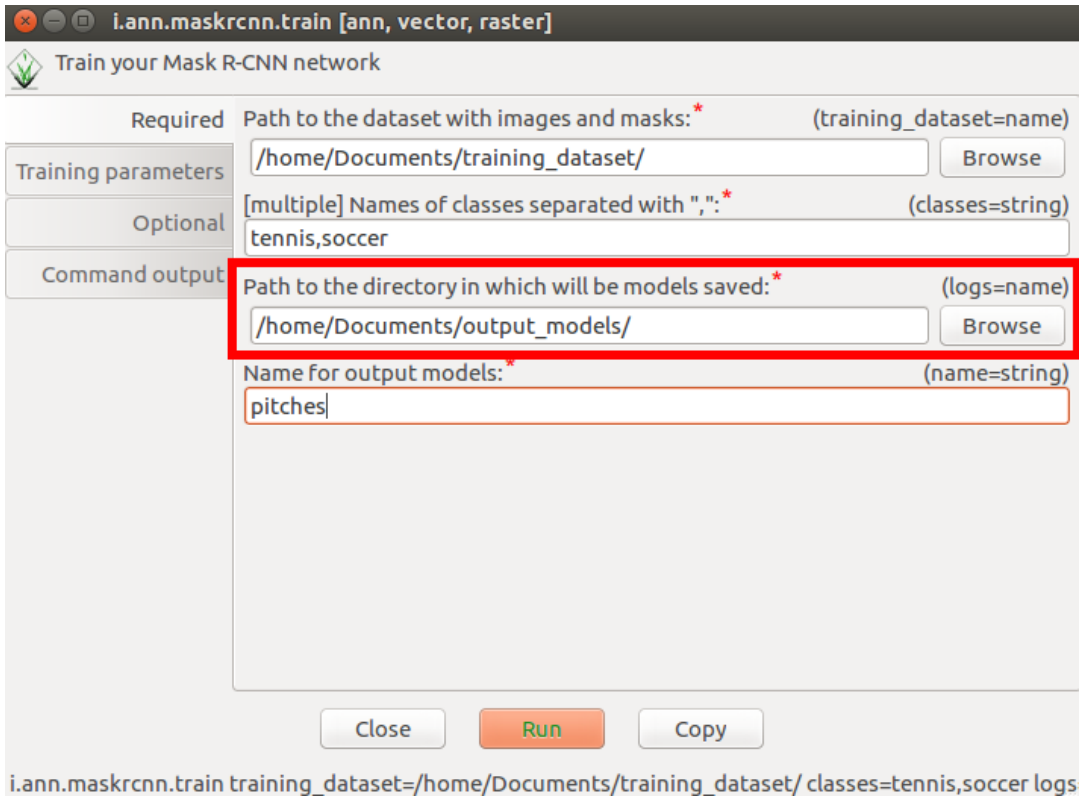


Ondřej Pešek

i.ann.maskrcnn.train

Workflow behind

- configuration of the model
- feed the model with pre-trained weights
- read the training dataset
- train
- save the model

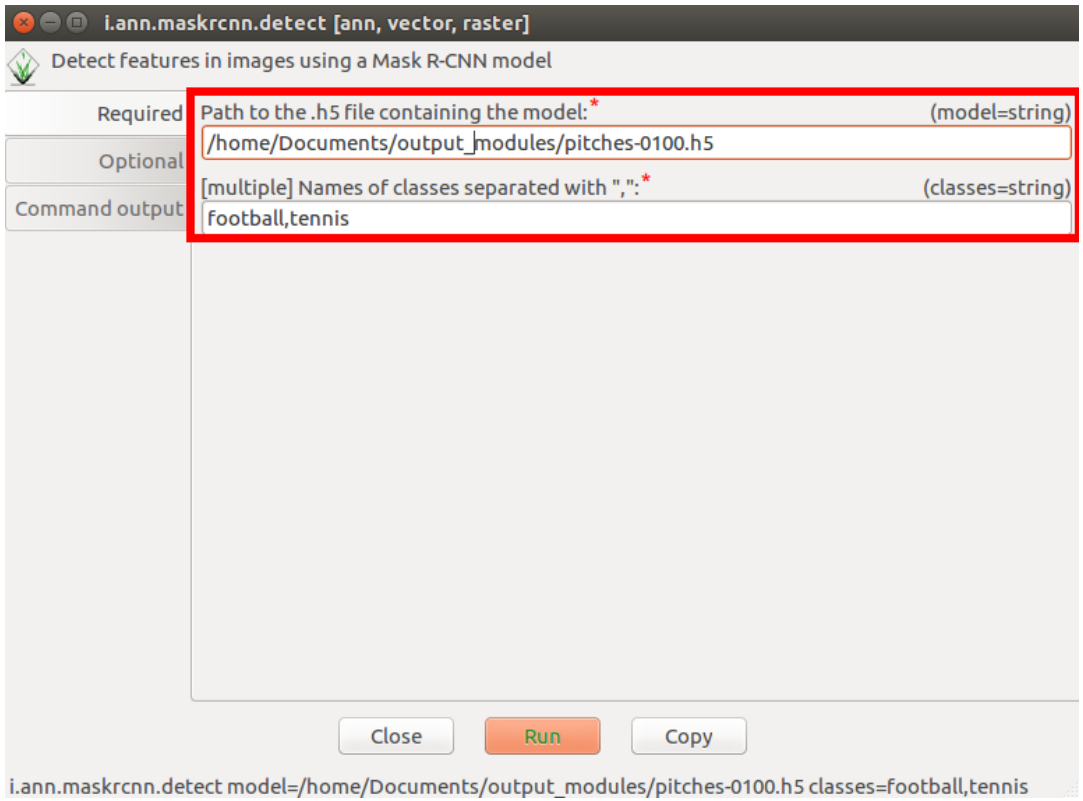


Ondřej Pešek

i.ann.maskrcnn.detect

Workflow behind

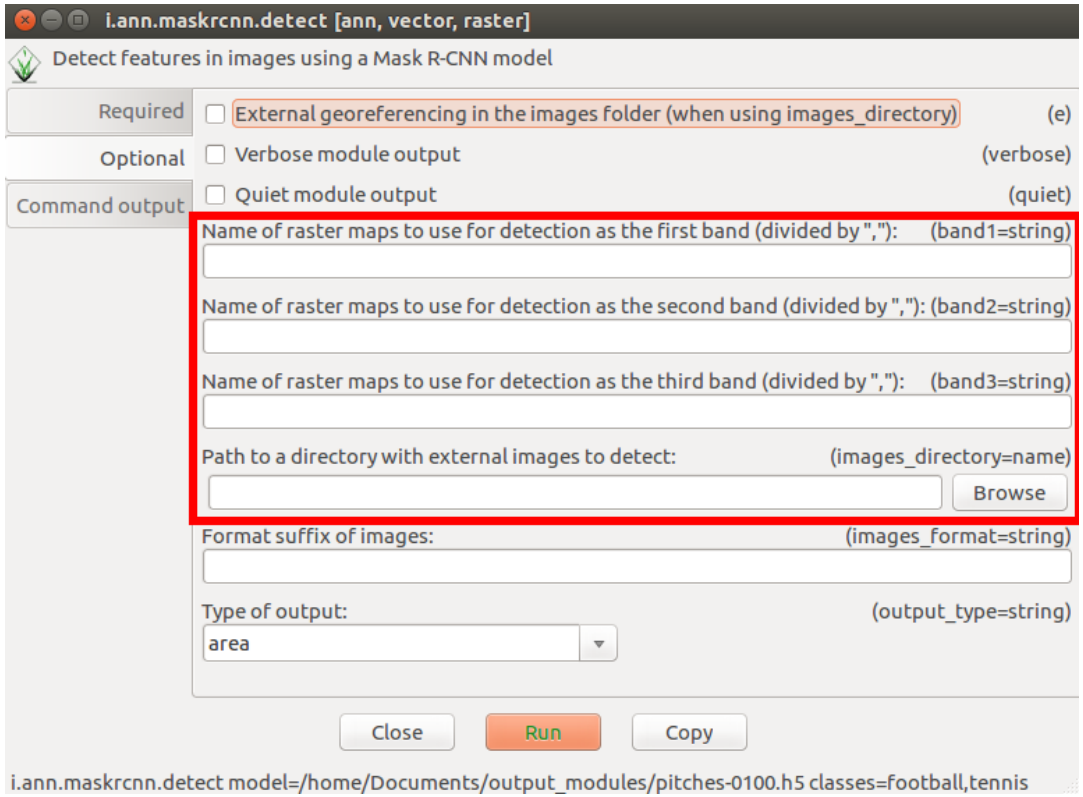
- load the model



i.ann.maskrcnn.detect

Workflow behind

- load the model
- detection for each raster

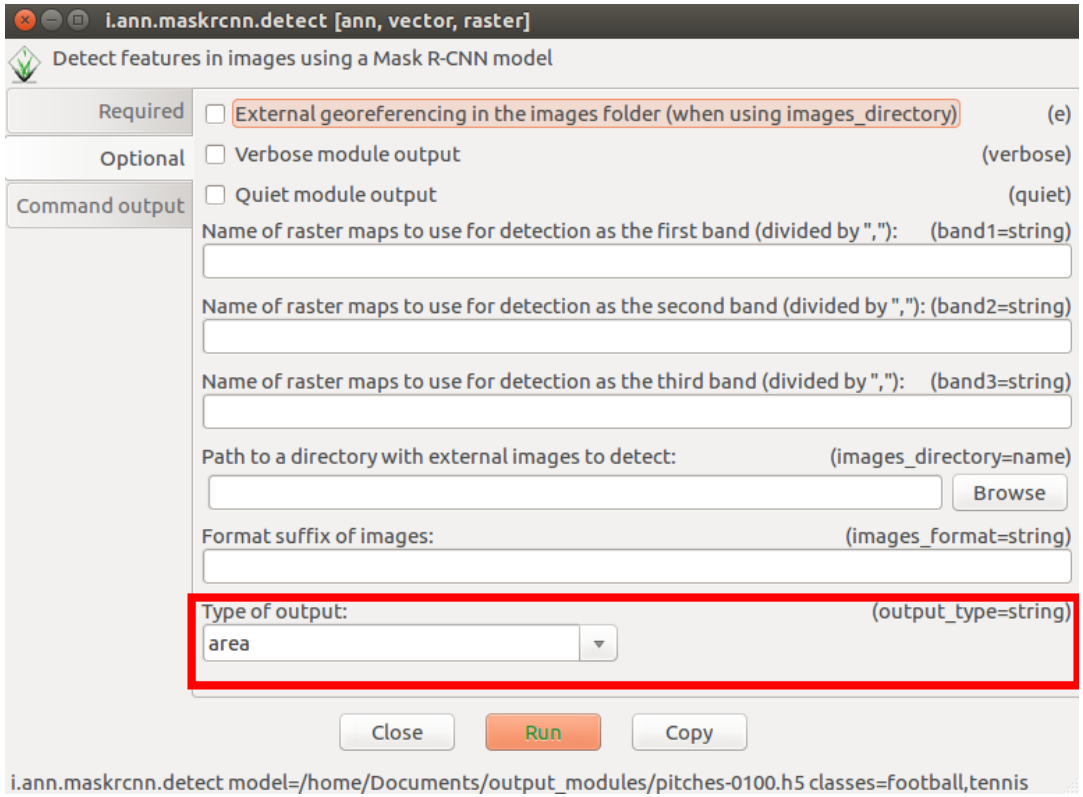


Ondřej Pešek

i.ann.maskrcnn.detect

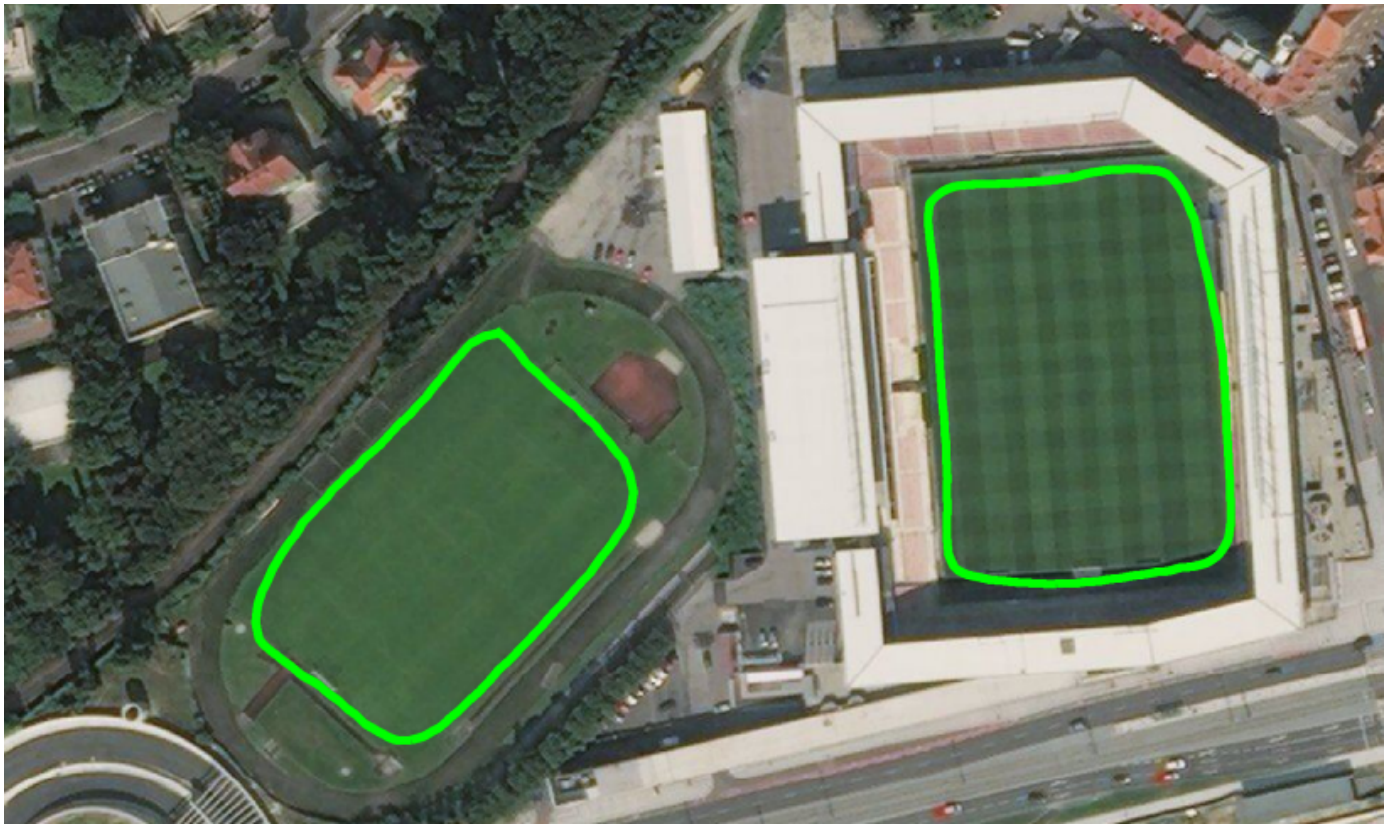
Workflow behind

- load the model
- detection for each raster
- vectorization



Ondřej Pešek

Results



Loss function 0.96, 54000 training images

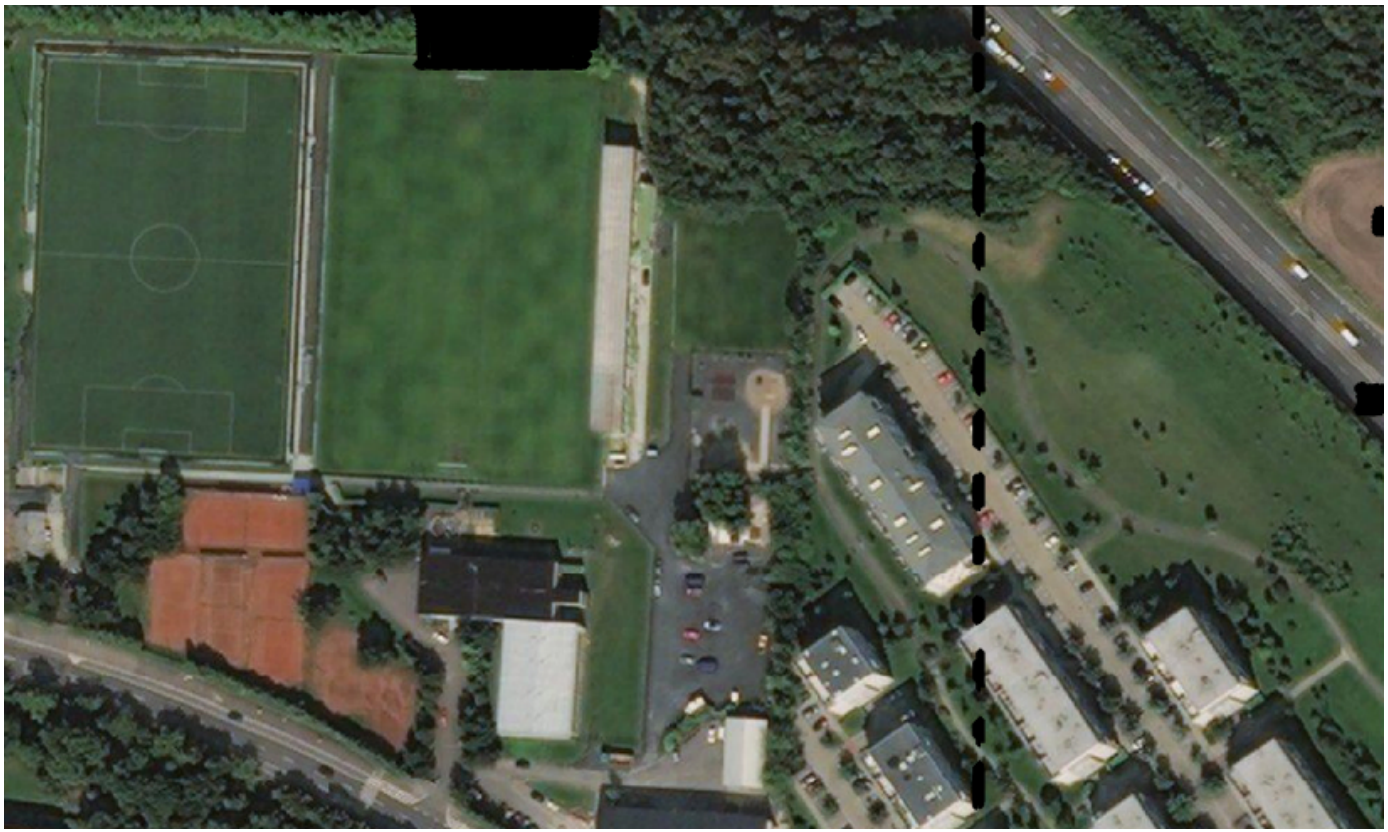
Results

Loss function 0.96, 54000 training images

Results

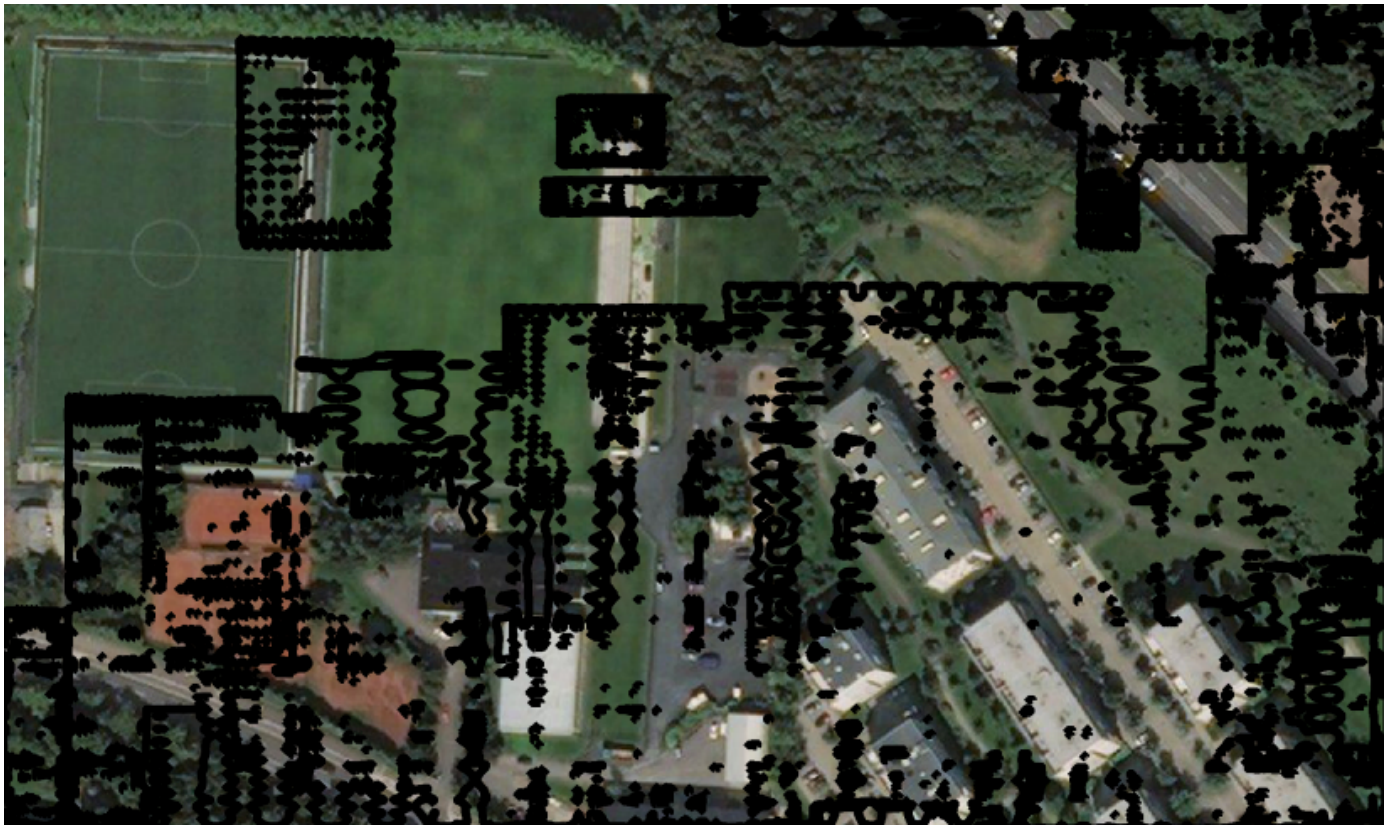
Loss function 0.96, 54000 training images

Results



Epoch 1, loss function 35.01, 2400 training images

Results



Epoch 10, loss function 5.87, 2400 training images

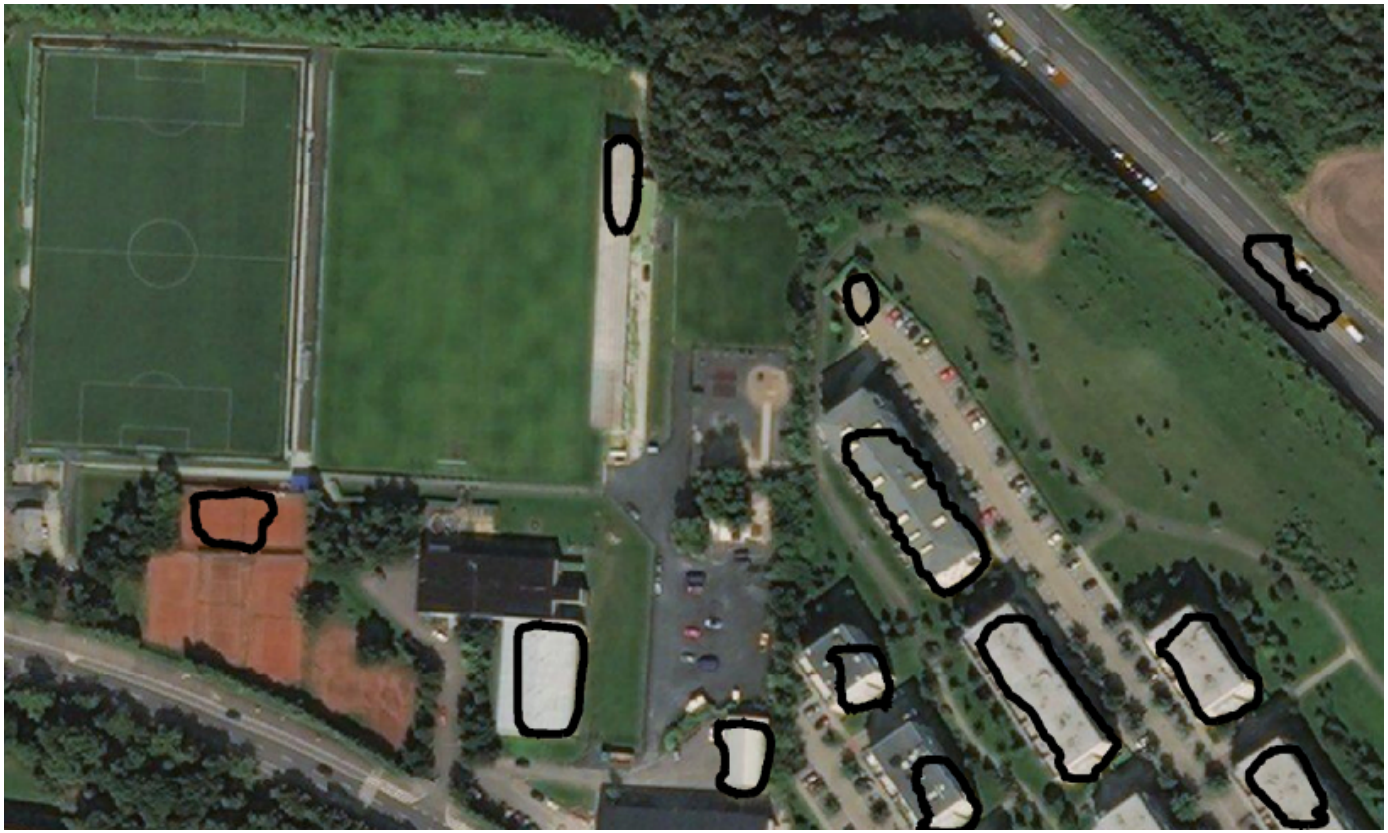
Results

Epoch 50, loss function 1.36, 2400 training images

Results

Epoch 150, loss function 0.63, 2400 training images

Results



Epoch 180, loss function 0.50, 2400 training images

Conclusion

- source code
  - `https://github.com/OSGeo/grass-addons/tree/master/grass7/imagery/i.ann.maskrcnn`
- installation using command

  `g.extension extension=i.ann.maskrcnn`
- next steps
  - multispectral rasters
  - training on rasters and vectors imported in GRASS
  - more architectures
  - (TensorFlow 2.0)

Sources

[1] RUSSAKOVSY, Olga et al. ImageNet Large Scale Visual Recognition Challenge. International Journal of Computer Vision IJCV. 2015, 115, n. 3, pp. 211–252.
[2] http://cs231n.stanford.edu/
[3] HE, Kaiming et al. Mask R-CNN. In: International Conference on Computer Vision (ICCV). 2017.
[4] GIRSHICK, Ross. Fast R-CNN. In: International Conference on Computer Vision (ICCV). 2015.

Thank you for your attention.