

Компьютерное зрение

Практический курс
Савельева Юлия Олеговна
i.o.saveleva.kpfu@gmail.com
2-й семестр, 14.04.2022 г.



Segmentation

Dataset

Статья с методом:

<https://arxiv.org/pdf/1505.04597.pdf>

Ссылка на датасет:

<https://cocodataset.org/#download>

2017 Train images [118K/18GB]

2017 Val images [5K/1GB]

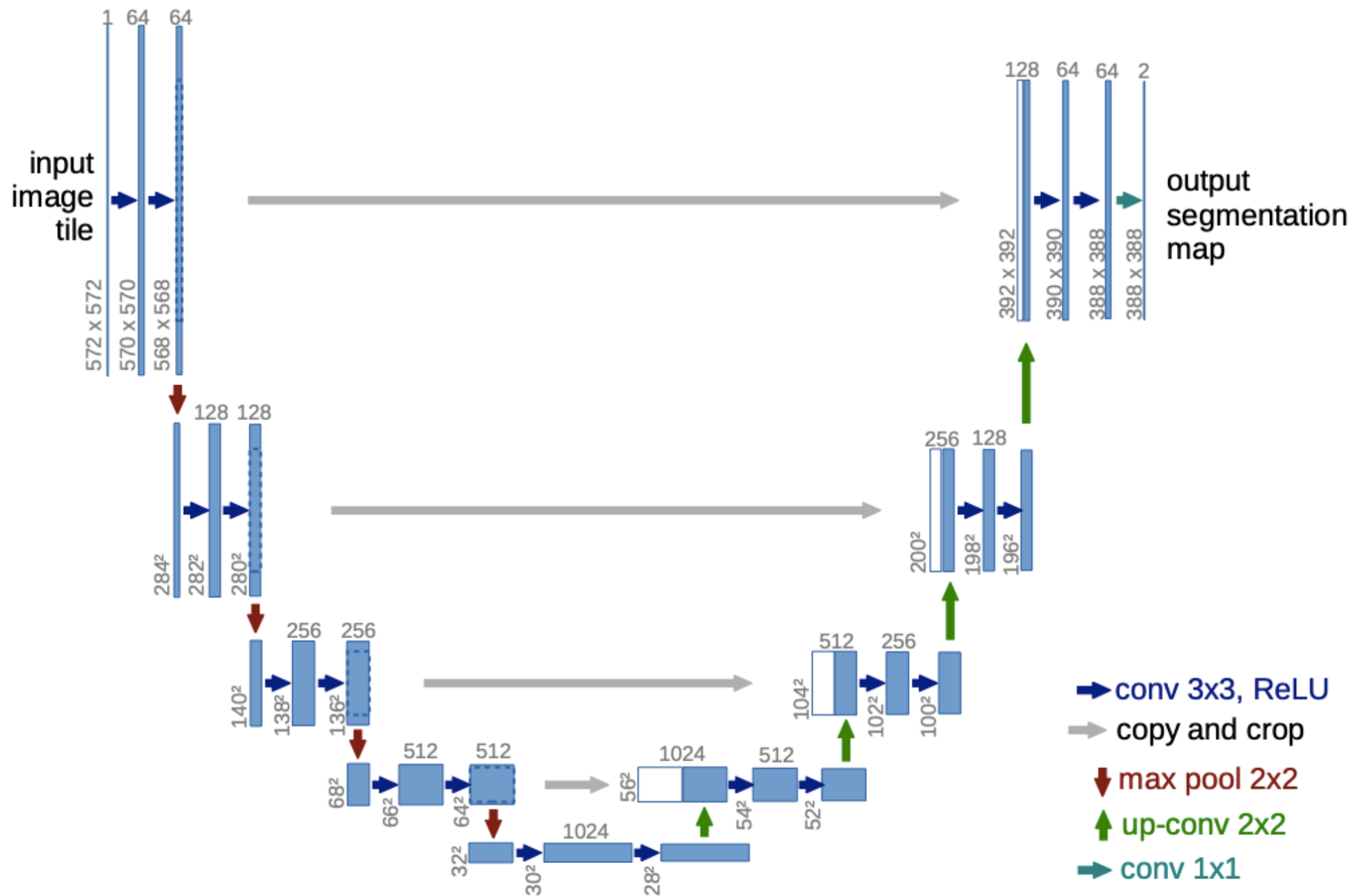
2017 Train/Val annotations [241MB]

COCO API: <https://github.com/cocodataset/cocoapi/blob/master/PythonAPI/pycocoDemo.ipynb> - демо для чтения картинок и аннотаций

- 1) Скачать датасет и настроить даталоудер для обучающей и валидационной выборки на чтение и предобработку картинок.
- 2) Реализовать архитектуру UNet.
- 3) Реализовать целевую функцию и обеспечить проход картинки через пайплайн.

Segmentation

Architecture



`nn.ConvTranspose2d(in_channels=in_channels, out_channels=out_channels, kernel_size=2, stride=2)`

Segmentation

Preprocessing, Loss, Accuracy

Preprocessing.

Представить лейблы сегментации как одноканальные картинки, где в каждом пикселе находится номер класса.

Preprocess image and label:

RandomCrop/CenterCrop (388, 388)

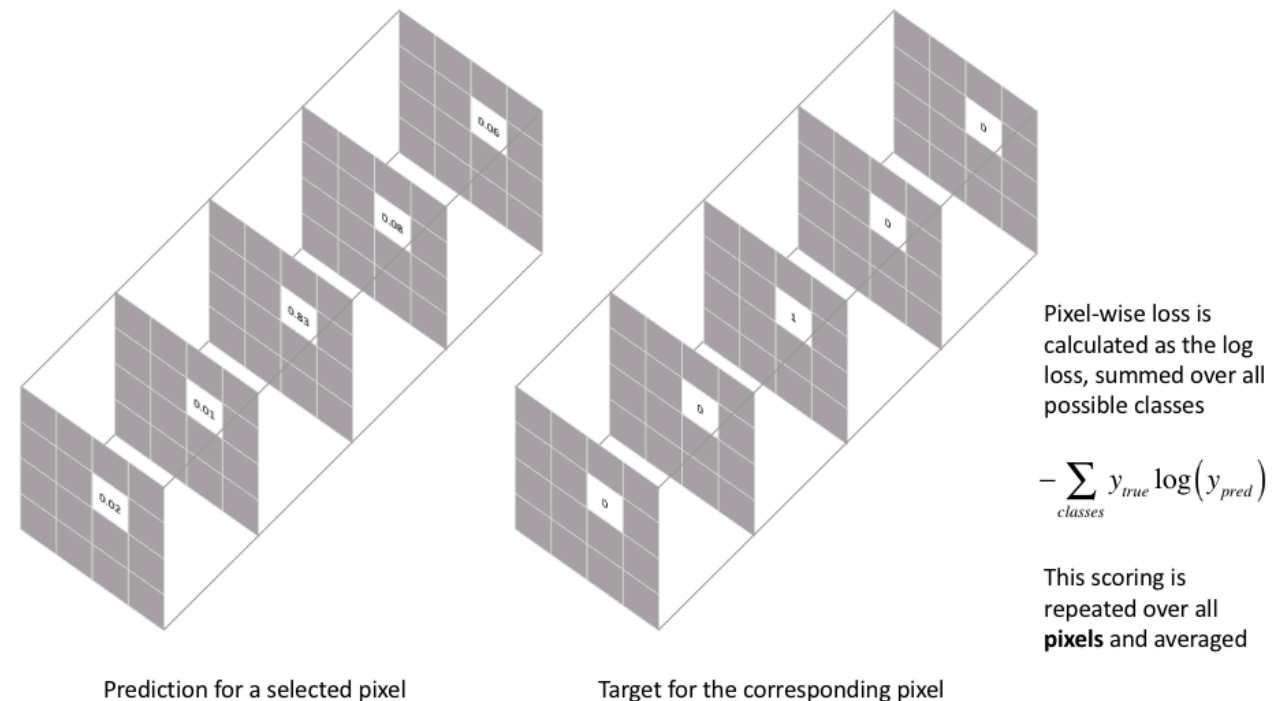
Preprocess image:

MirrorPad (92, 92, 92, 92)

Normalize (red_(μ , σ), green_(μ , σ), blue_(μ , σ))

Loss.

Использовать стандартный Cross Entropy.



Accuracy. **Mean Pixel Accuracy (MPA):** Average ratio of correctly classified pixels per-class basis

$$PA = \frac{1}{k+1} \sum_{i=0}^k \frac{p_{ii}}{\sum_{j=0}^k p_{ij}}$$

- $k + 1$ classes including void or background
 - p_{ij} is the amount of pixels of class i inferred to belong to class j
- p_{ii} represents the number of true positive

Дедлайн 28.04.2022 23:59