D·Labs

Introduction to Face Recognition

12.04.2021

Tomasz Iżycki, Machine Learning Engineer

Overview

Introduction

Business case

Project overview

Computer Vision intro

Segmentation net

Face Recognition net

Q&A session

Business Case

01

D·Labs

Business case

Problem to solve:

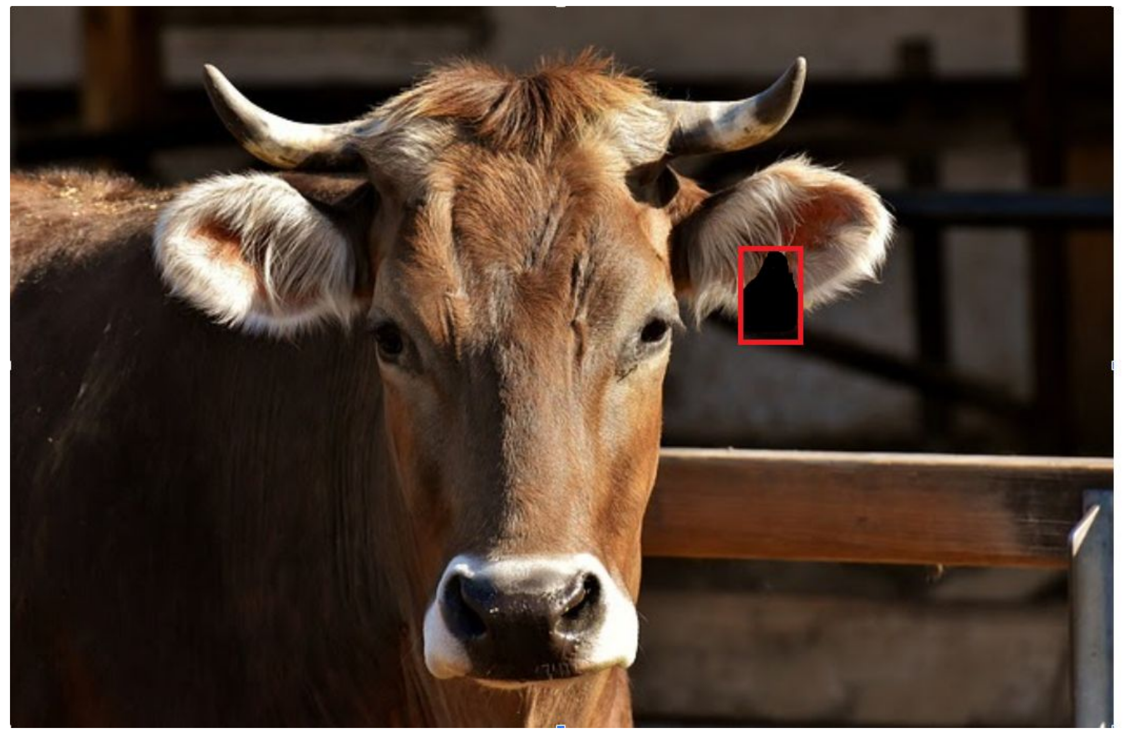
- Proper handling cattle identification using image taken by farm workers (without using ear markings)
  - Monitor and manage cattle across one particular or many farms to help farm workers
  - Keep tracking medical history of cattle

D·Labs

Project overview

02

Project overview



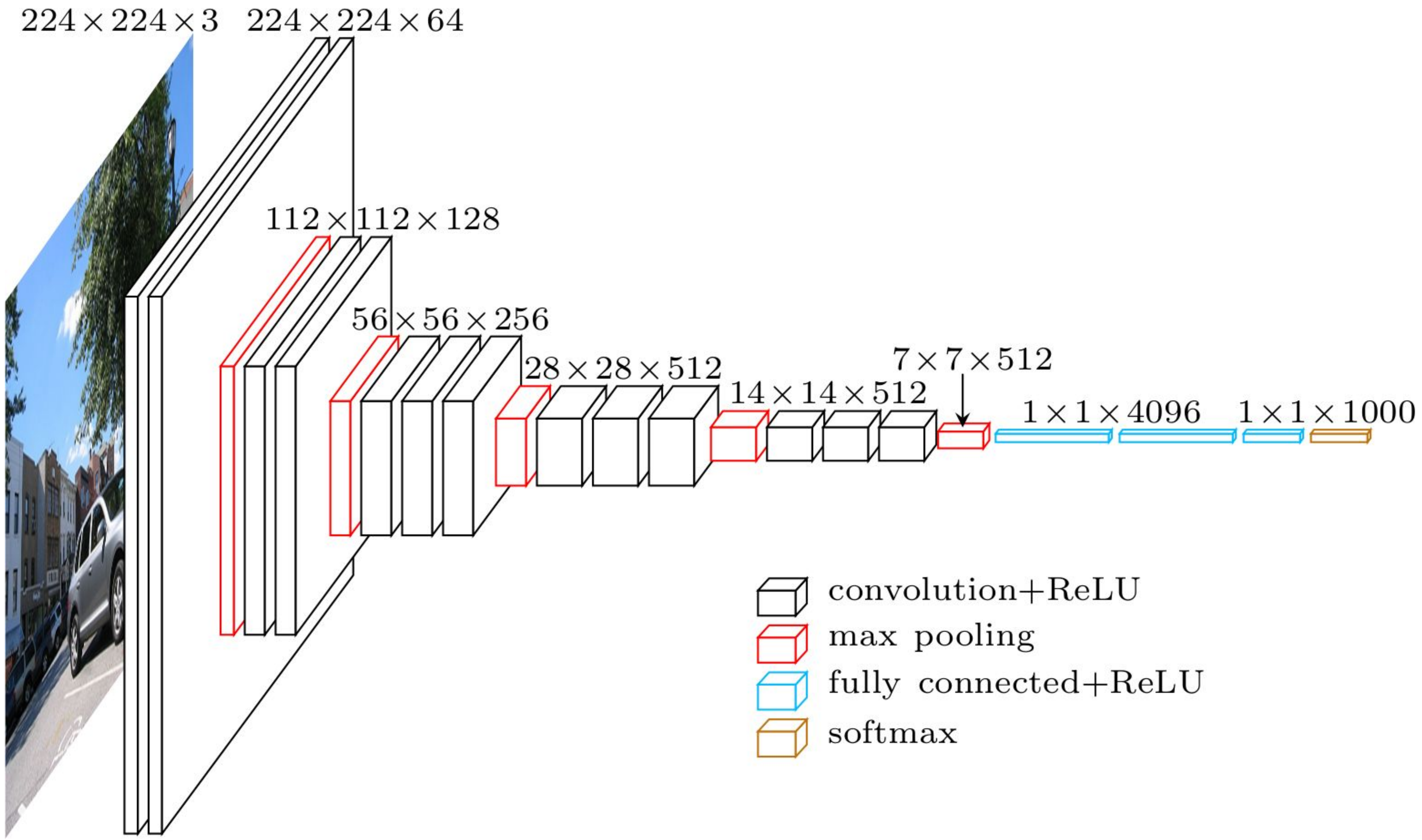
- Gathering data
- Labeling data - ear marking removal
- Segmentation net - detect ear marking
- Face recognition net
- Backend deployment
- App deployment
- Further Data gathering
- Net recalibration

Computer Vision intro

03

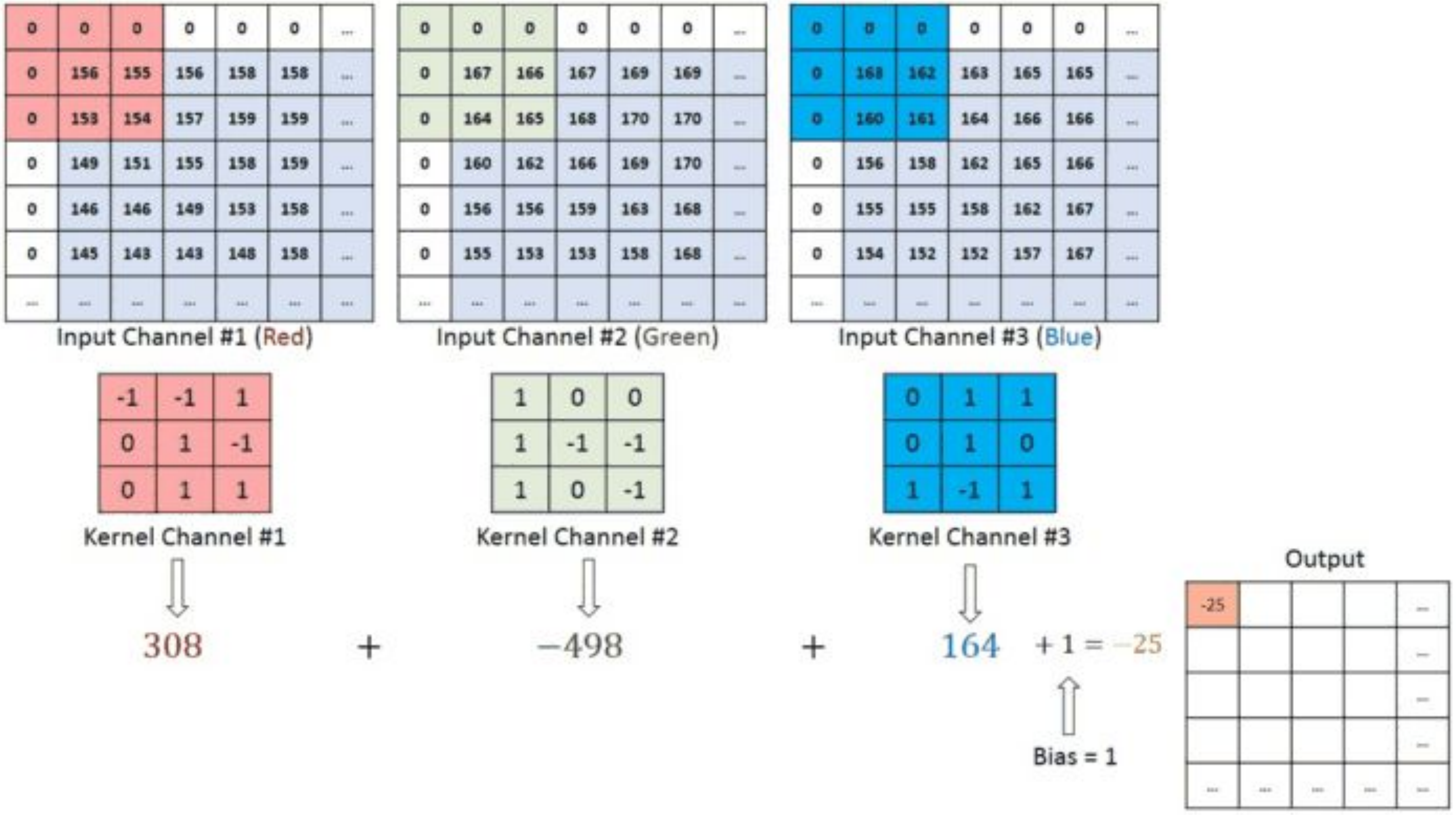
Computer Vision catch-up

CNN networks - VGG example



Computer Vision catch-up

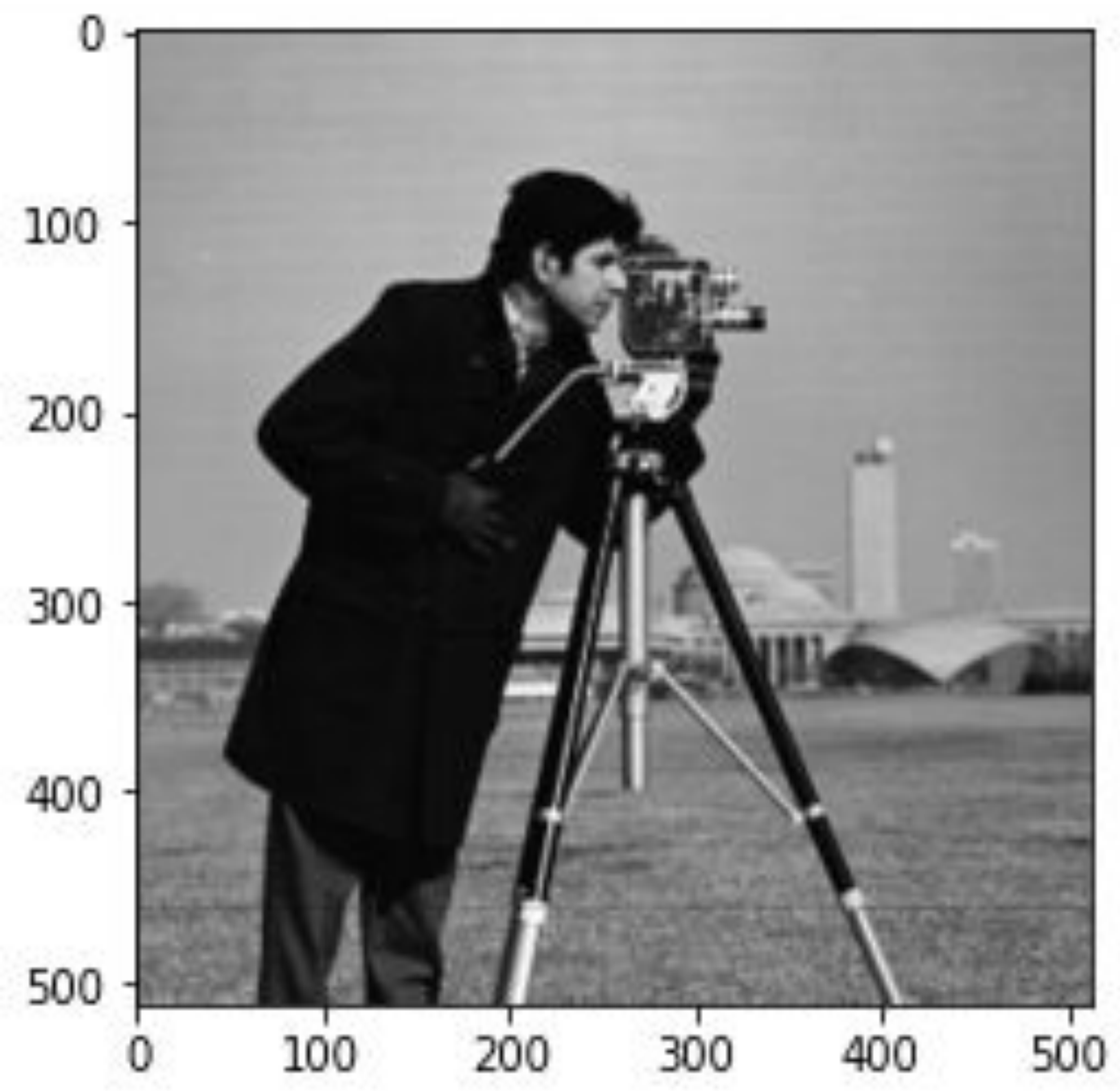
CNN kernel operations



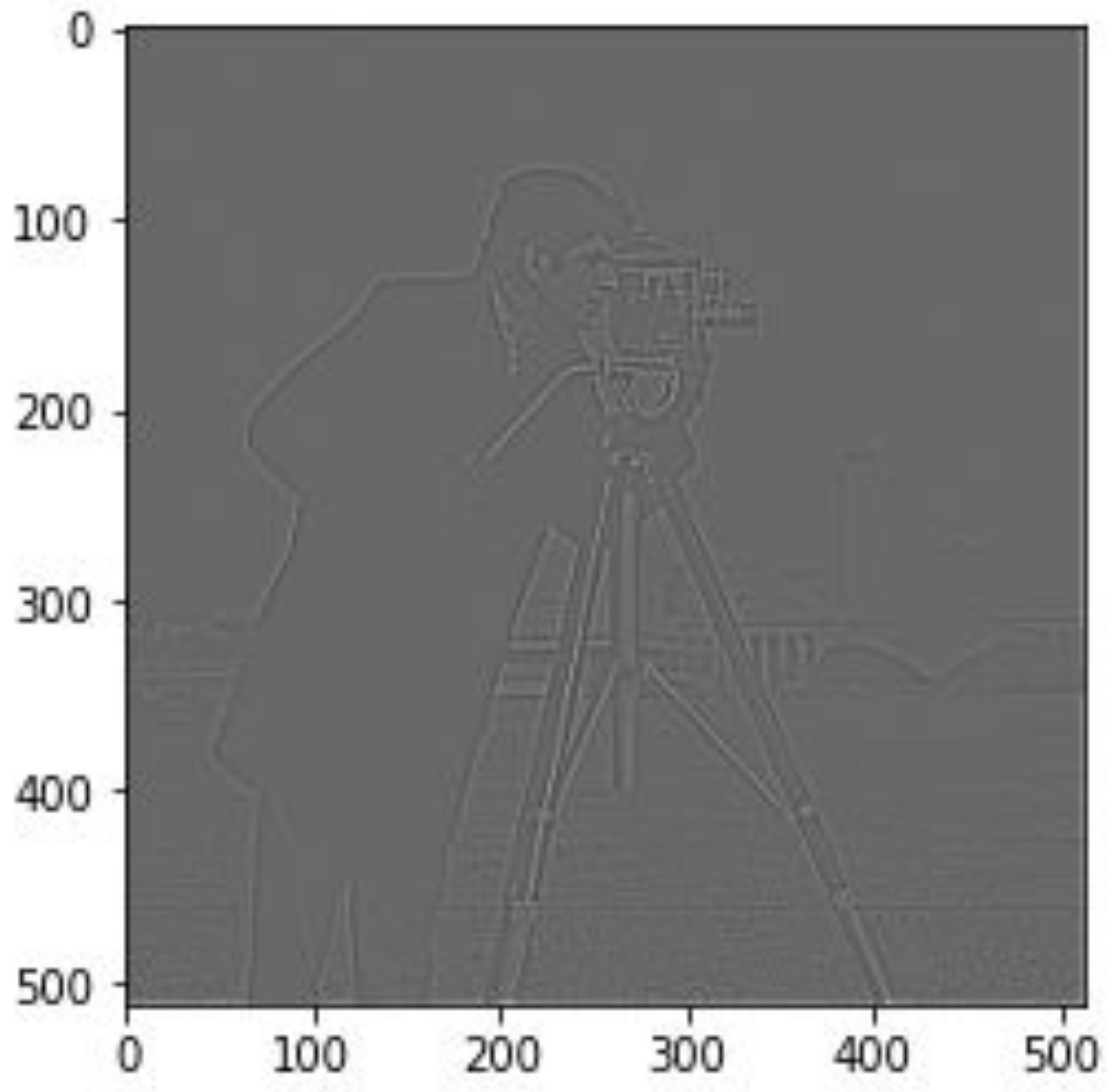
D·Labs

Computer Vision catch-up

CNN kernel examples

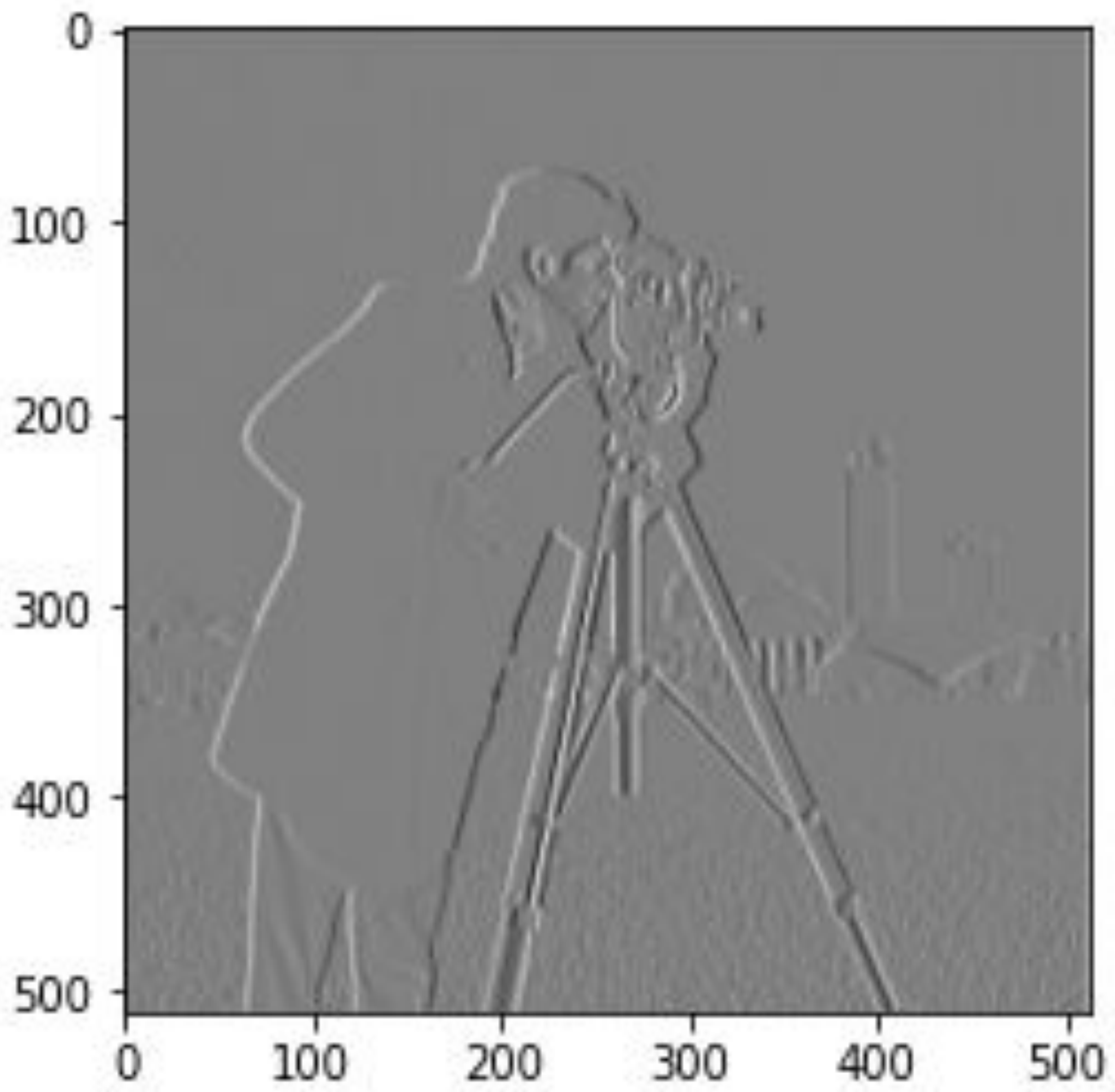


Original image



$$\begin{pmatrix} -1 & -1 & -1 \\ -1 & 8 & -1 \\ -1 & -1 & -1 \end{pmatrix}$$

Edge filter



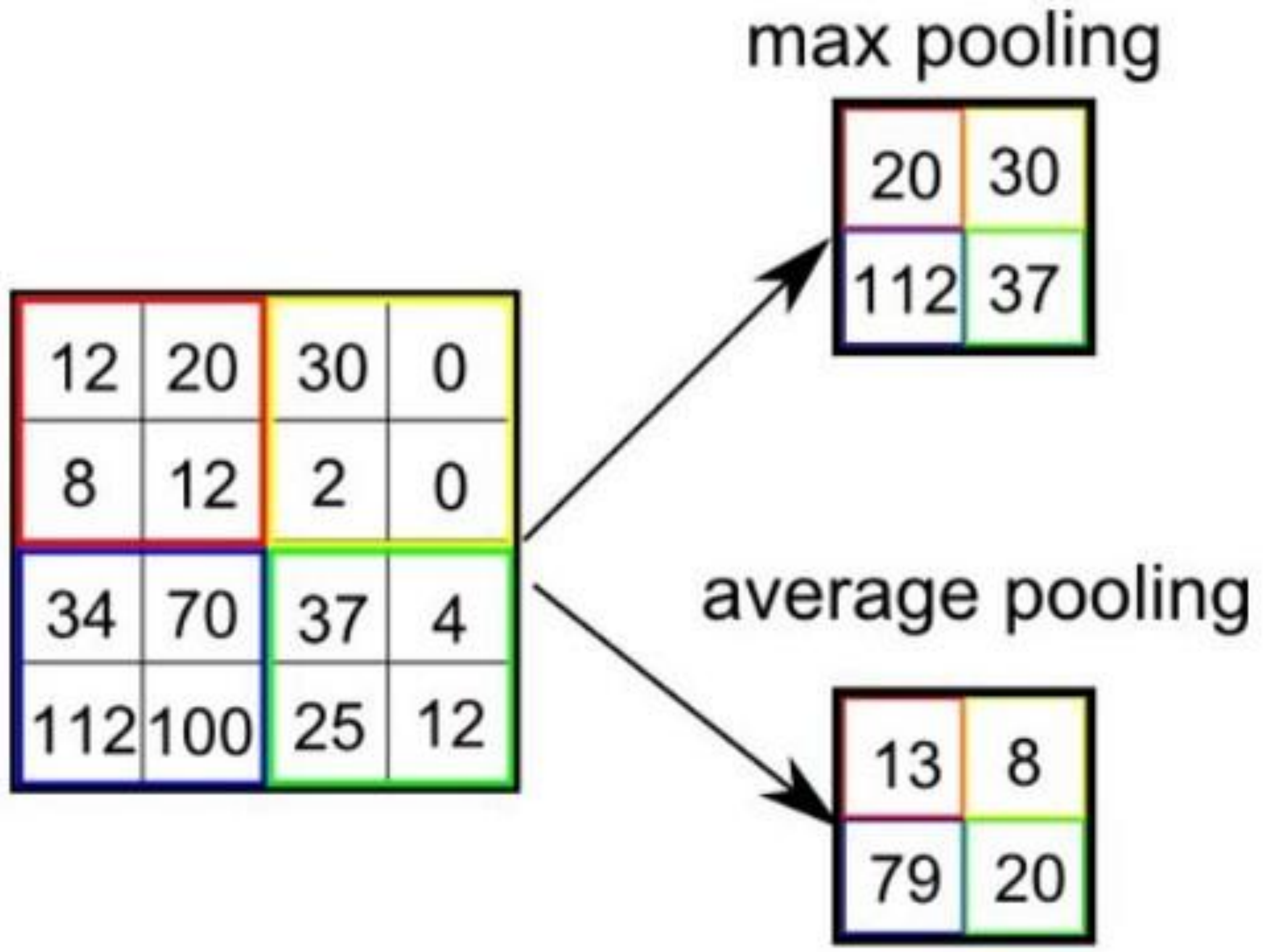
$$\begin{pmatrix} 1 & 0 & -1 \\ 1 & 0 & -1 \\ 1 & 0 & -1 \end{pmatrix}$$

Left sobel filter

Computer Vision catch-up

CNN pooling

- One of the downsampling method
  - reduce number of feature-map coefficients to process
  - induce spatial-filter hierarchies by making successive convolution layers look at increasingly large windows (in terms of fraction of the original input they cover)
- Alternative to using higher strides in prior convolutional layers
- Max pooling tends to work better than using higher strides or average pooling because features tend to encode the spatial presence of some pattern or concept over the different tiles of the feature map and it's more informative to look at the maximal presence of different features than at their average presence

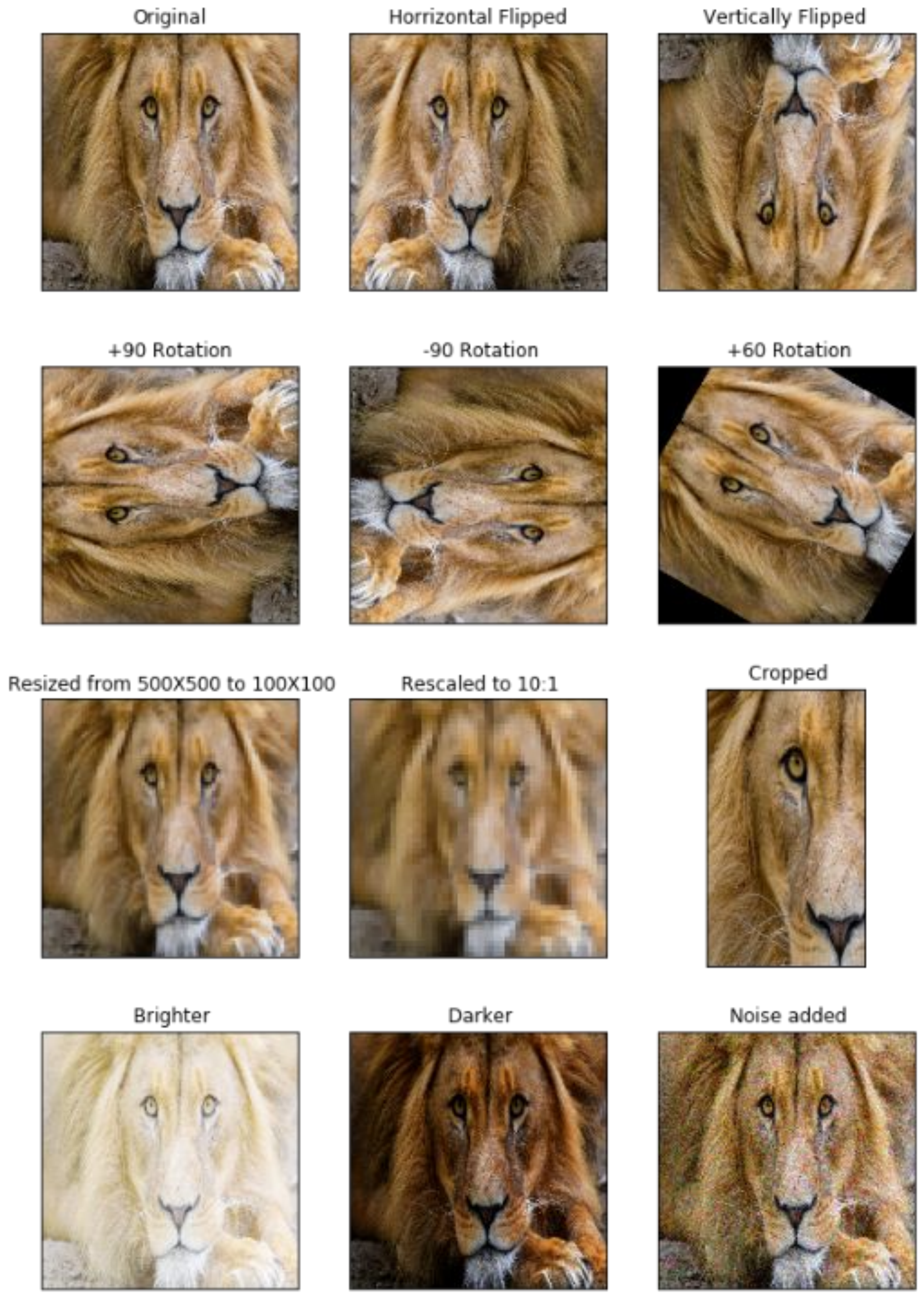


D·Labs

Computer Vision catch-up

Data augmentation

- One of the regularization methods
- Way to resolve an issue of limited amount of data
- Reduce overfitting
- Improve the generalization of our model by introducing more diversity to our data

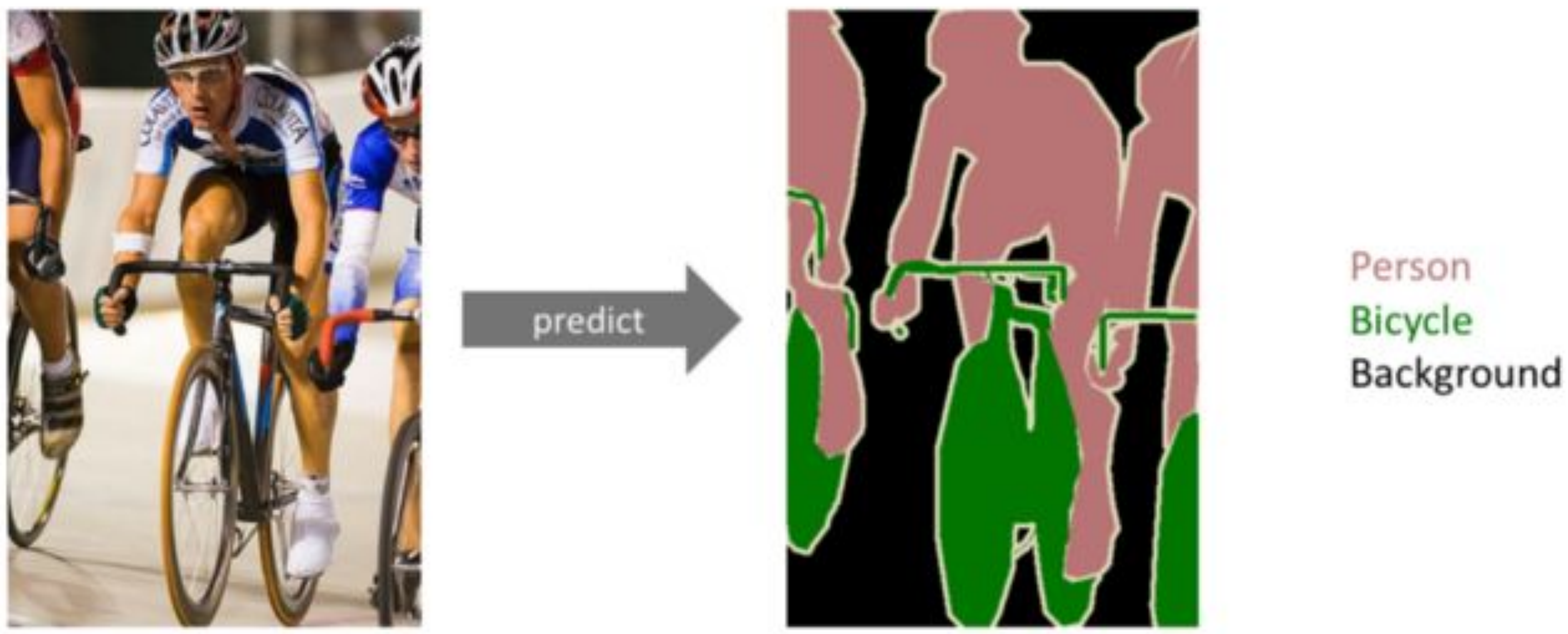


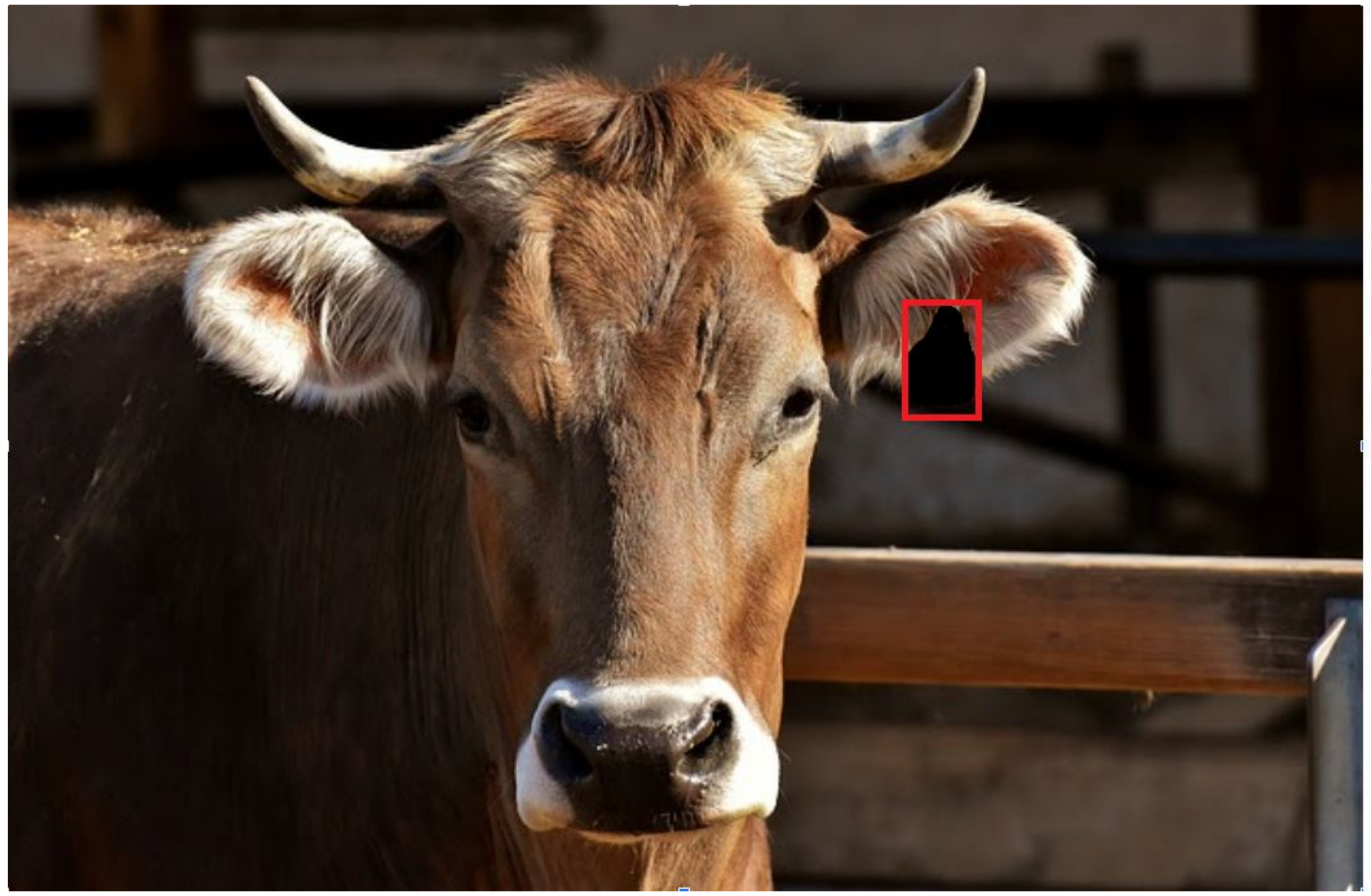
Segmentation net

04

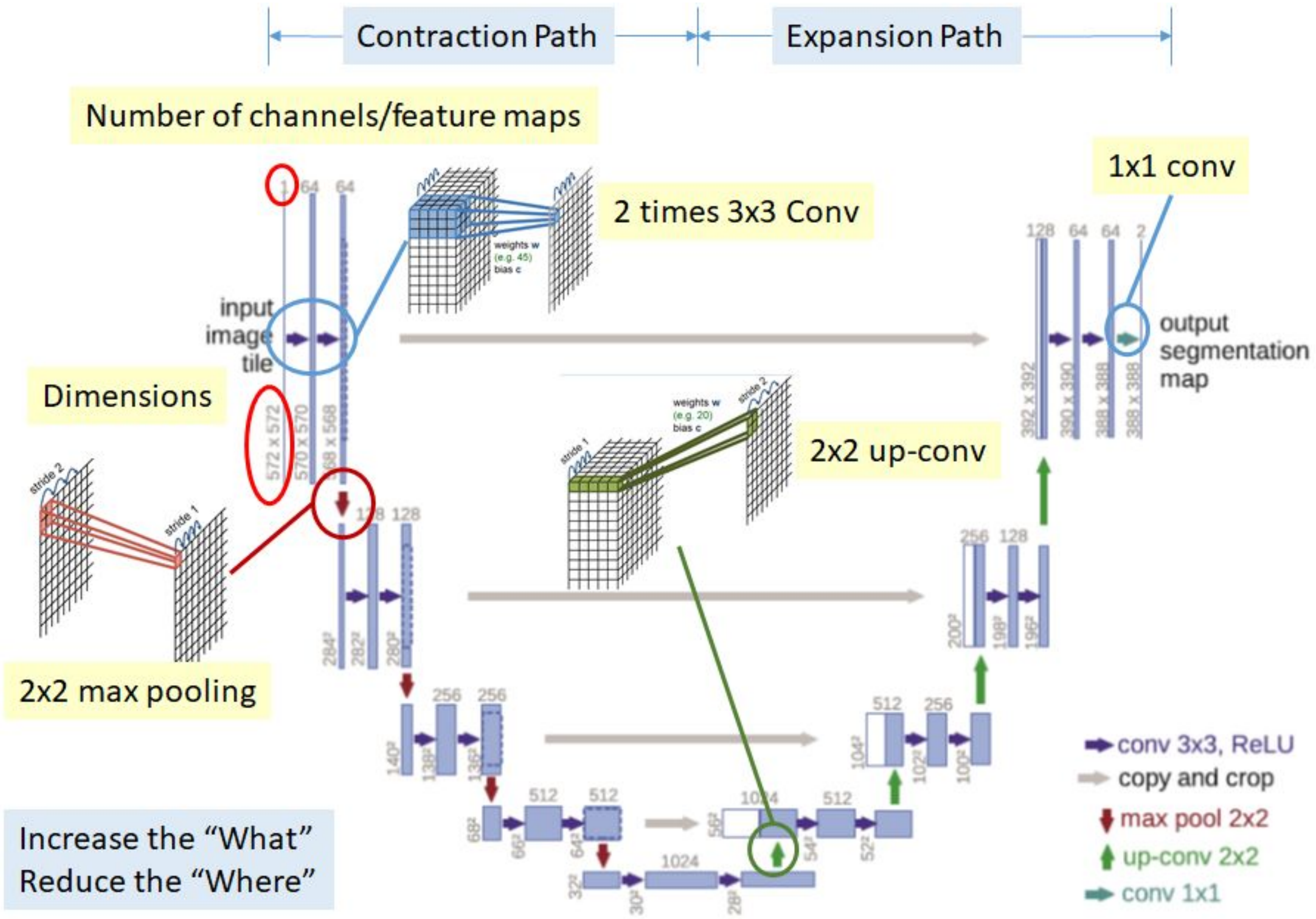
Segmentation net - UNet

Semantic segmentation





D·Labs



Increase the "What"
Reduce the "Where"

Face Recognition net

05

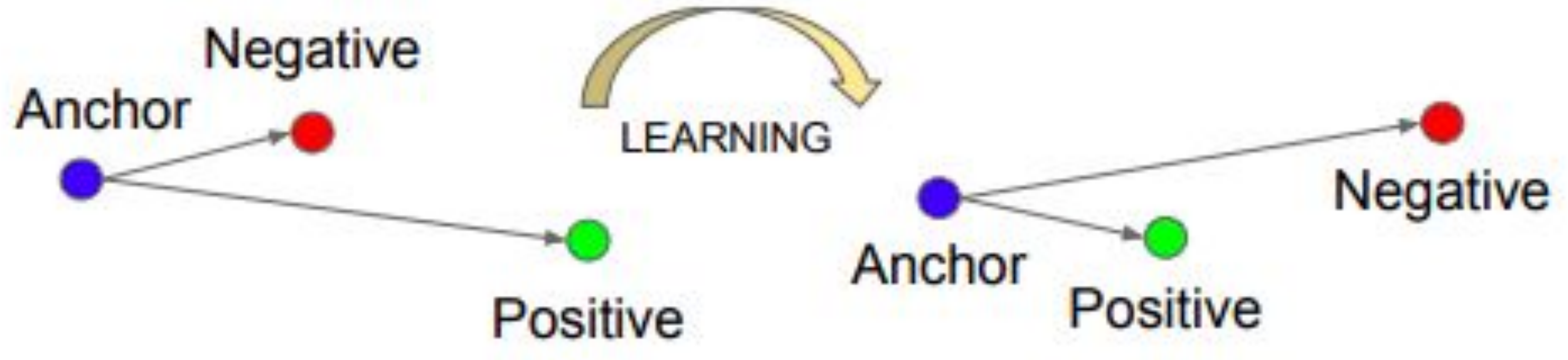
Face Recognition net

- What input data is required?
- What is expected output?
- How to handle new identities without re-training?
- What kind of net architecture is required?
- Which loss function should we apply?
- Which metric would help us to measure net effectiveness in interpretable way?

Face Recognition net

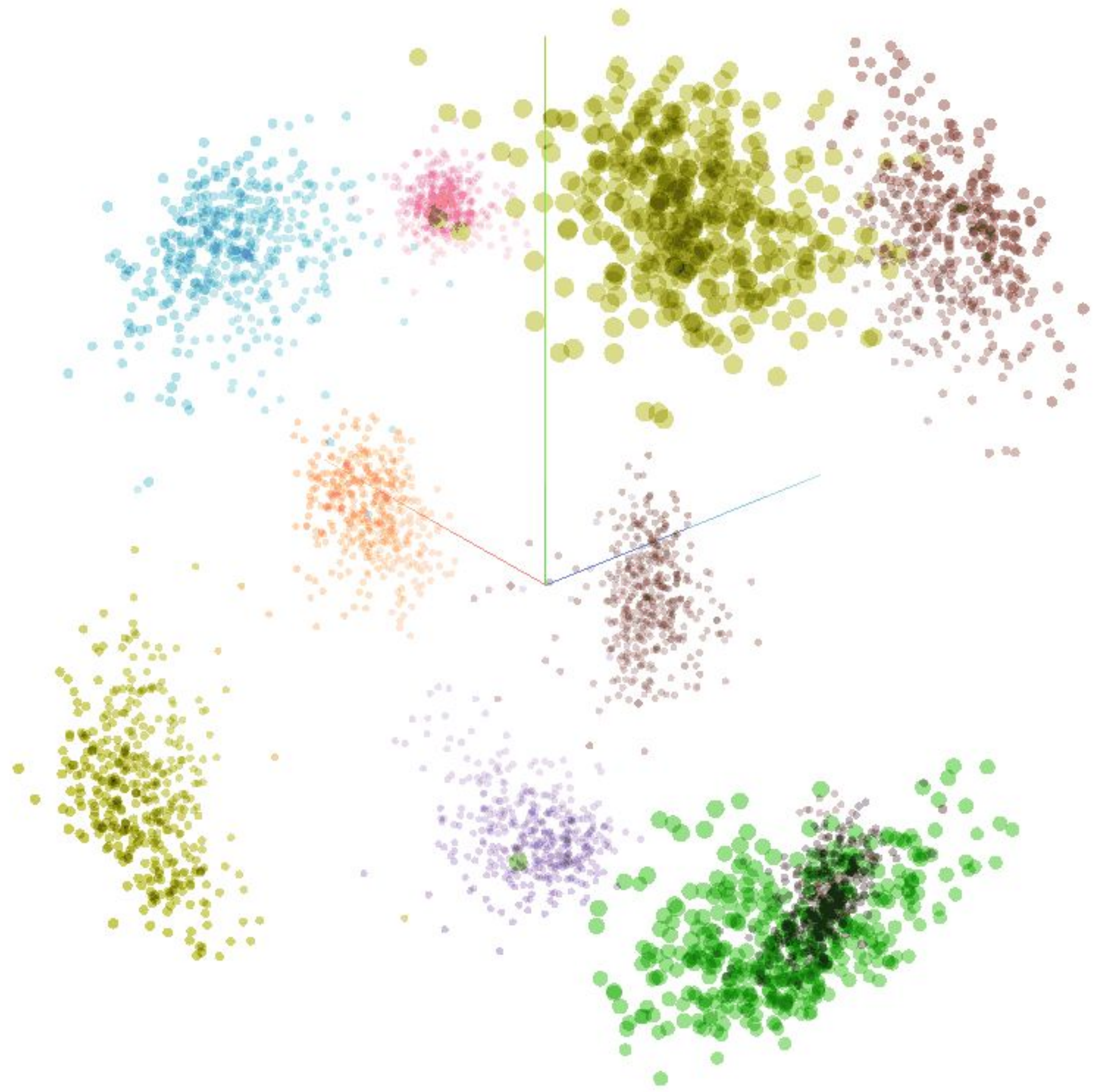
Triplet loss

$$L_{triplet}(x_a, x_p, x_n) = max(0, m + \|f(x_a) - f(x_p)\|_2^2 - \|f(x_a) - f(x_n)\|_2^2)$$



The objective is to learn embeddings such that the anchor is closer to the positive example than it is to the negative example by some margin value.

To enforce ranking of the distances (distance between anchor and positive sample is smaller than distance between anchor and negative sample) we introduce margin to separate those embeddings



D·Labs

Face Recognition net

Problems with triplet loss

- The number of possible triplets increases significantly with increasing number of classes and data samples
- It is computationally expensive to process all possible combinations - possible triplets
- Triplets provided to network might be generated randomly. Therefore while network training continues, we provide more frequent examples that are easy to deal with (triplet loss is equal to 0) which prevents network from further training - loss equal 0 cannot be back-propagated to modify weights

Face Recognition net

Types of triplets

- **Easy triplets** - triplets with loss equal 0
- **Hard triplets** - triplets with distance between anchor and negative sample being lower than distance between anchor and positive sample
- **Semi-hard triplets** - triplets with distance between anchor and positive sample being lower than distance between anchor and negative sample but loss is still greater than 0 (“margin requirement” is not met)

Face Recognition net

Strategies of generating triplets

- **Offline Triplet Mining** - triplets are generated manually and before training of network

- **Online Triplet Mining** - Batch Sampler is being used to generate triplets for further training. For each mini batch positive and negative pairs will be selected using provided labels.
  - **Strategies**
    - All possible triplets
    - Hardest negative for each positive pair (using the same negative for each anchor)
    - Random hard/semi-hard negative for each positive pair (consider only triplets with positive triplet loss value)

D·Labs

THANK YOU

Time for questions