Training Camp 2024

Second Day – 5th September 2024

Alessandro Nicolosi

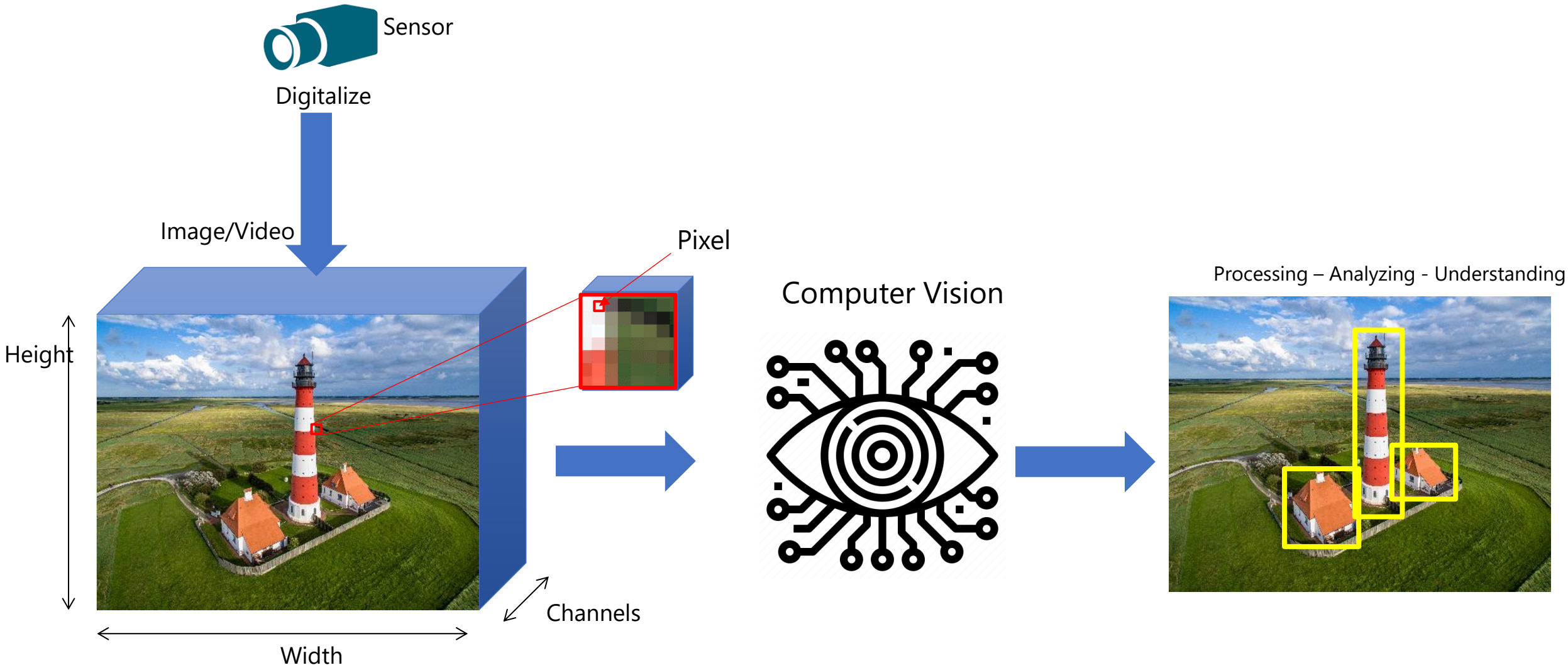
https://github.com/alenic

https://www.linkedin.com/in/alessandro-nicolosi/

LEONARDO

https://www.leonardo.com/it/

Computer Vision – An overview



Computer Vision – An overview

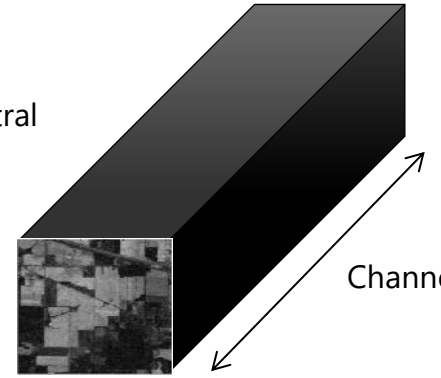
Electro-Optical

Channels = 3 (RGB)

Infrared

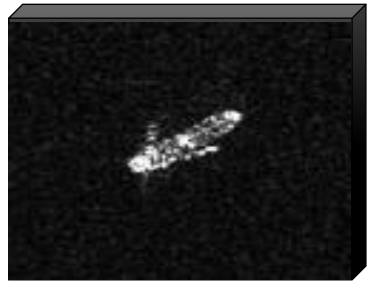
Channels = 1

Hyperspectral



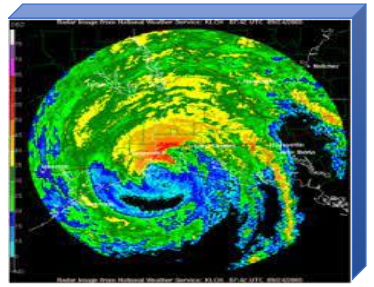
Channels = 288 (400-2500 nm)

SAR



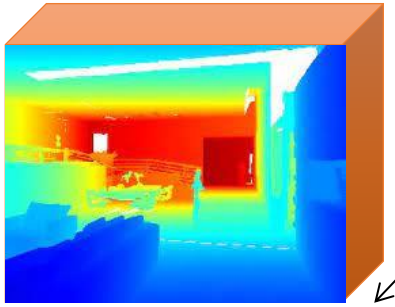
Channels = 1

Radar



Channels = 1

RGB-D

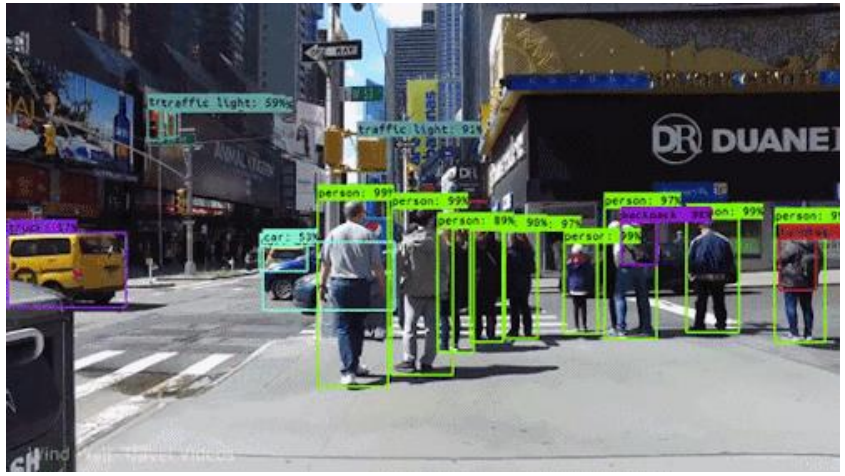


Channels = 4 (RGB - Depth)

And many others...

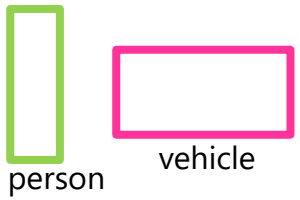
Computer Vision – An overview

Object Detection



Input
Image, Video

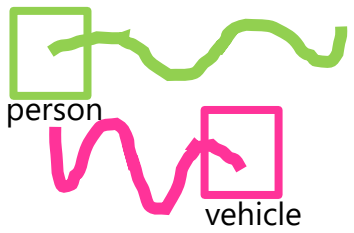
Output
Object positions + Object class



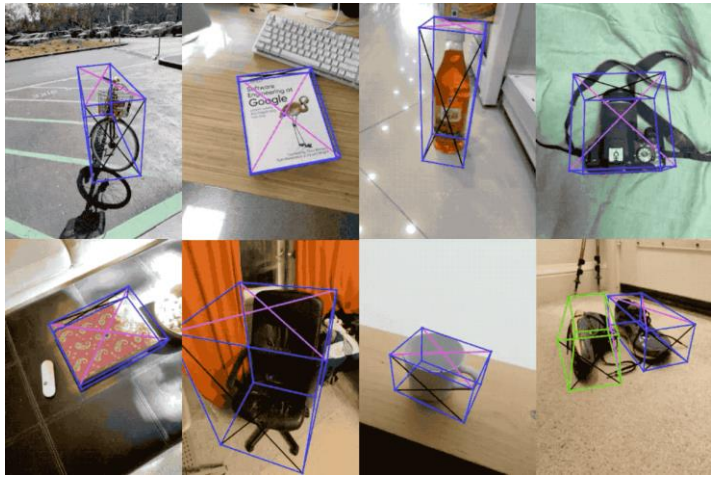
Object positions + Track

Input
Image, Video

Output
Object positions + Object class +
Track history

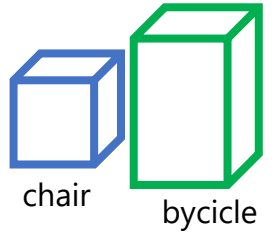


3D Object-Detection



Input
Image, Video

Output
Object positions + Object class +
Object pose



Computer Vision – An overview

Image Enhancement

Input
Image, Video

Output
Enhanced image/video: deblurred, denoised,...

Super Resolution

Input
Image, Video

Output
2x,4x,8x image/video

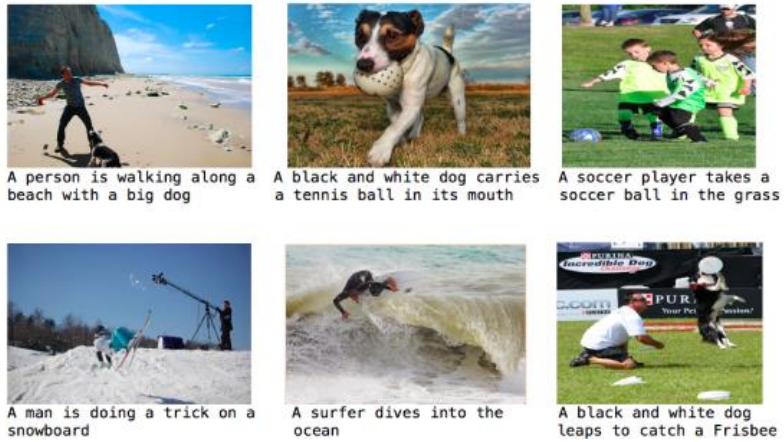
Object inpatining

Input
Image + pixels to remove

Output
Inpainted image

Computer Vision – An overview

Visual Captioning



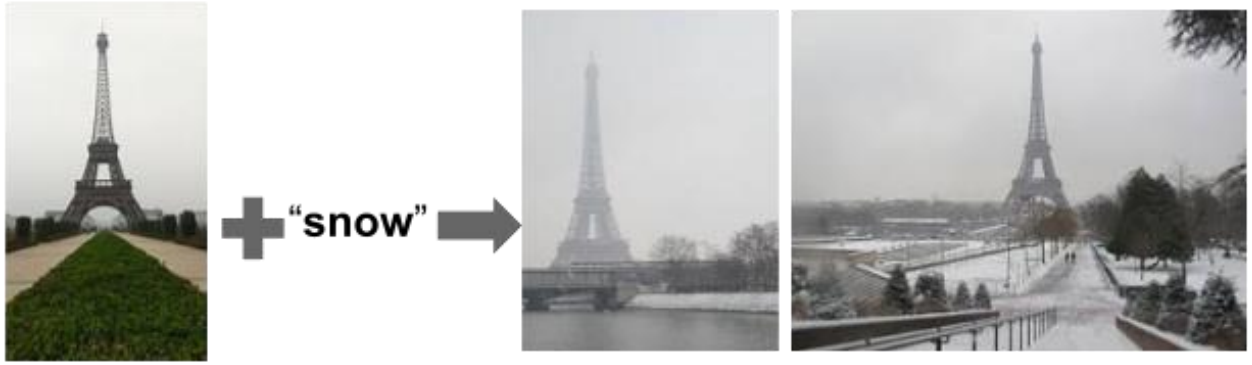
Input

Image, Video

Output

Image/video textual description

Image+Text Retrieval



Input

Image + Description

Output

Retrieved image

Text to Image

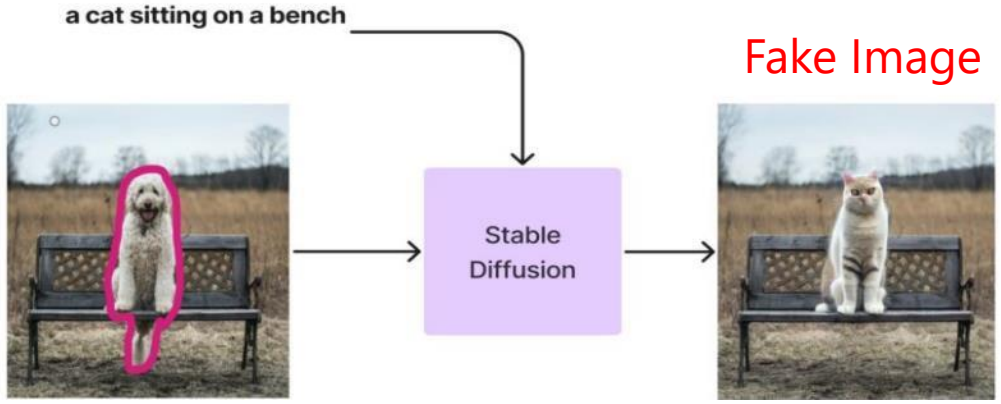
"President Donald Trump being arrested by New York City police"

Fake Image

Input
Textual description

Output
Image representing the input text

Generative inpatining

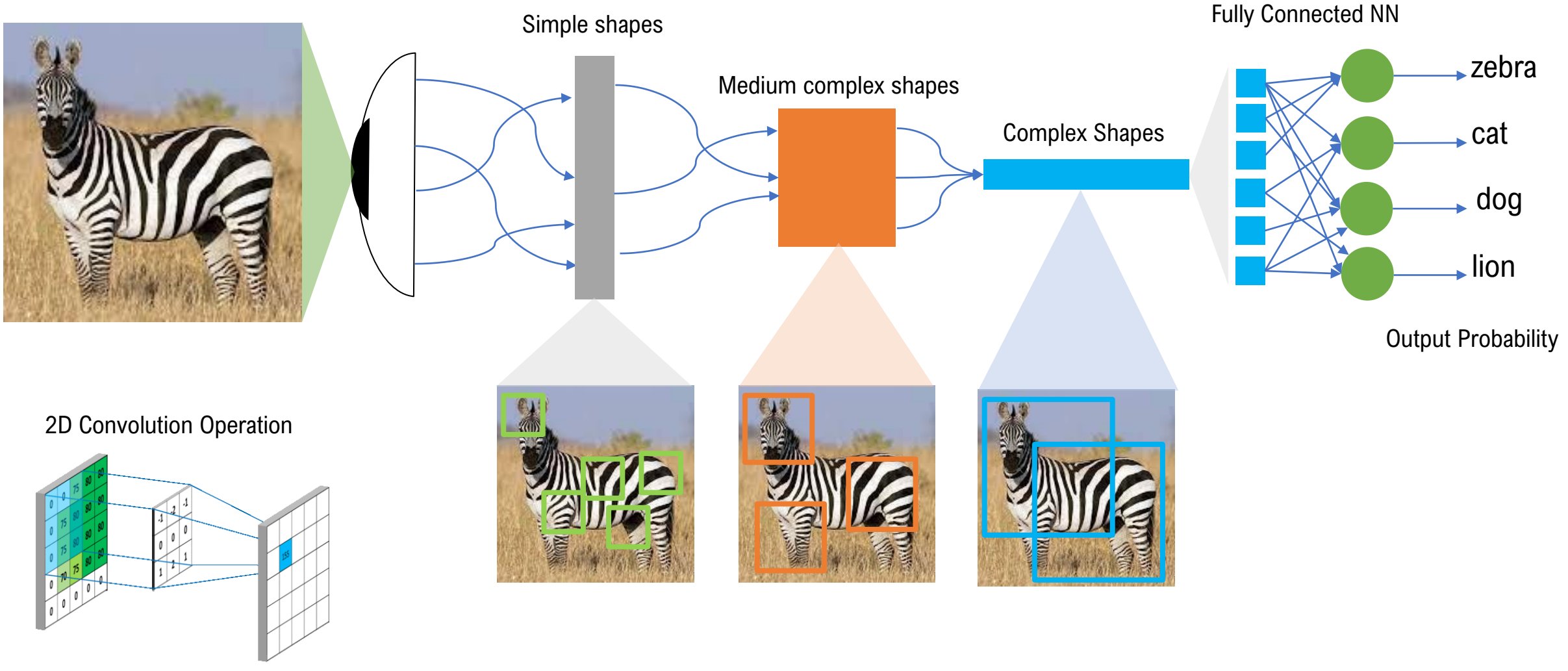


Fake Image

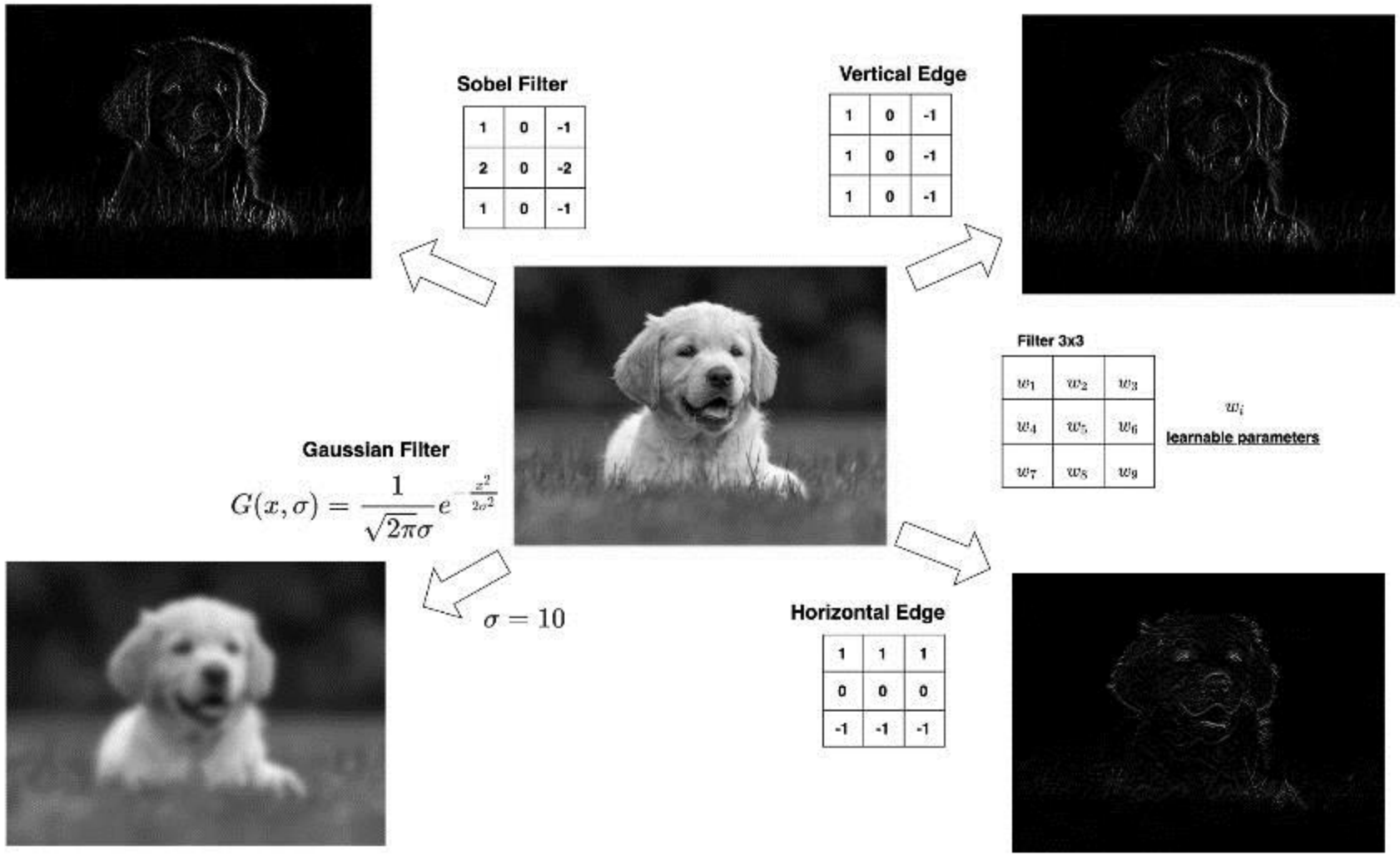
Input
Image + Pixels + Textual description

Output
Image with the replaced description

Convolutional Neural Network



Convolution in 2D

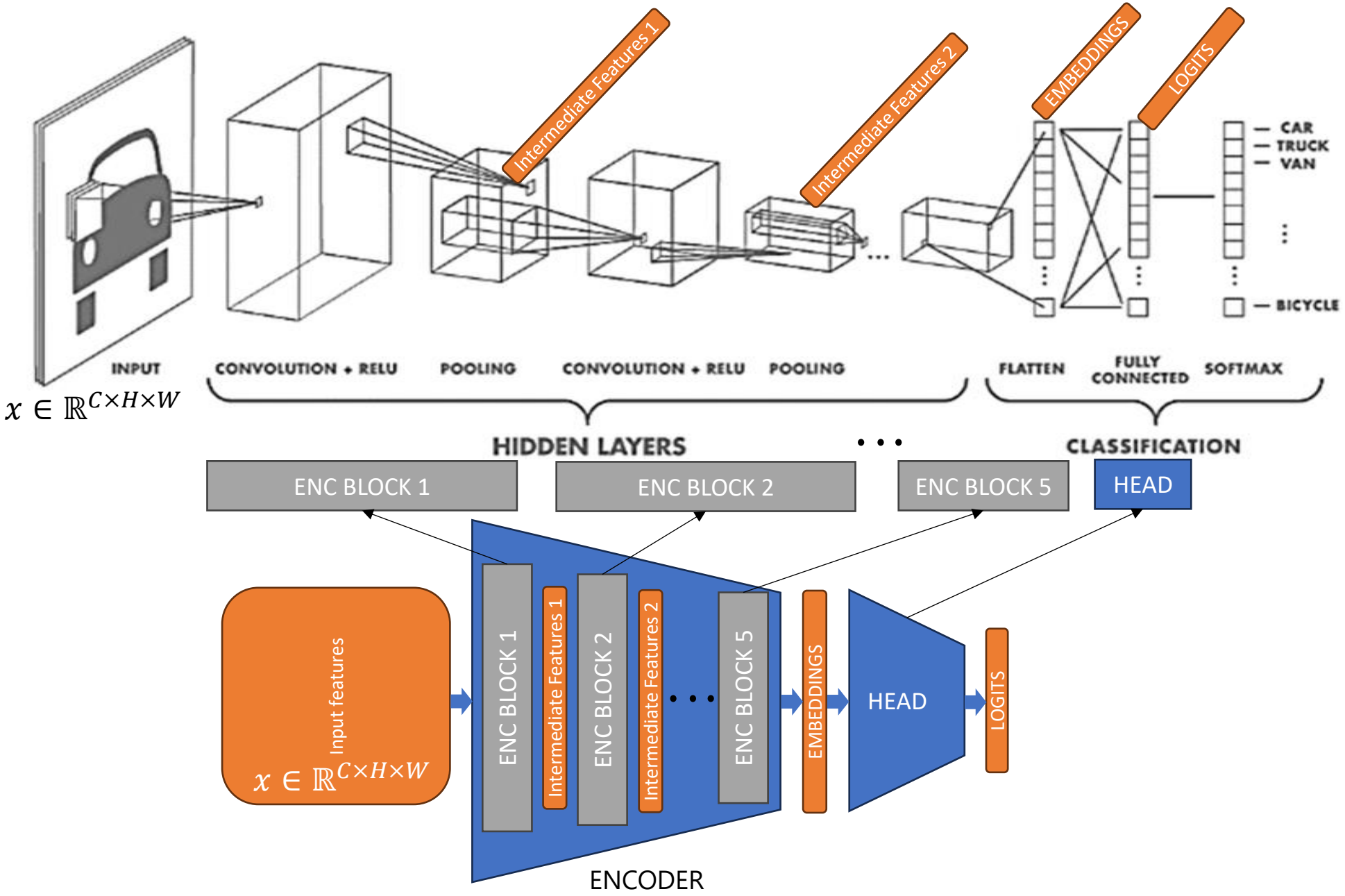


- A filter can detect a specific feature in the image
- A filter of size (f,f) has fxf trainable parameters

- Filter examples:
  - **Edge detection**
    - Vertical edges
    - Horizontal edges
    - Sobel filter

  - **Gaussian Filter**
    - Blur images

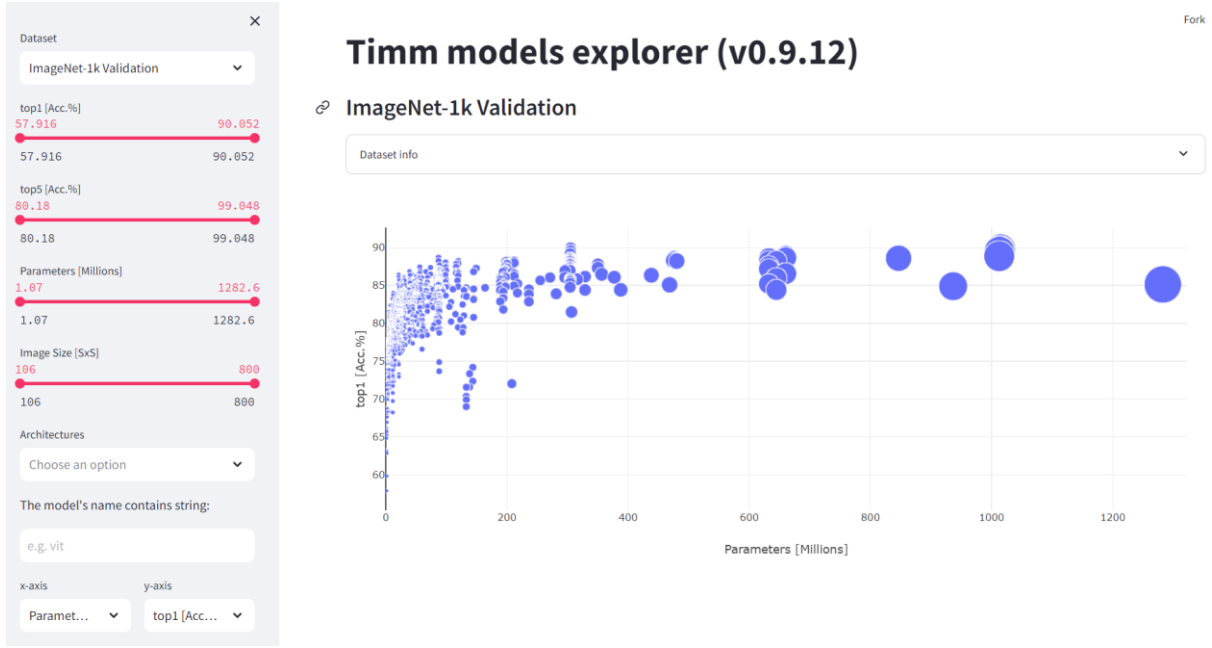
General CNN architecture



CV – Pytorch Models

- **timm:** https://github.com/huggingface/pytorch-image-models
  More than 700 pre-trained models!

- **Torchvision**: https://pytorch.org/vision/stable/models.html
  Best for start and for other tasks



https://timm-models-explorer.streamlit.app/

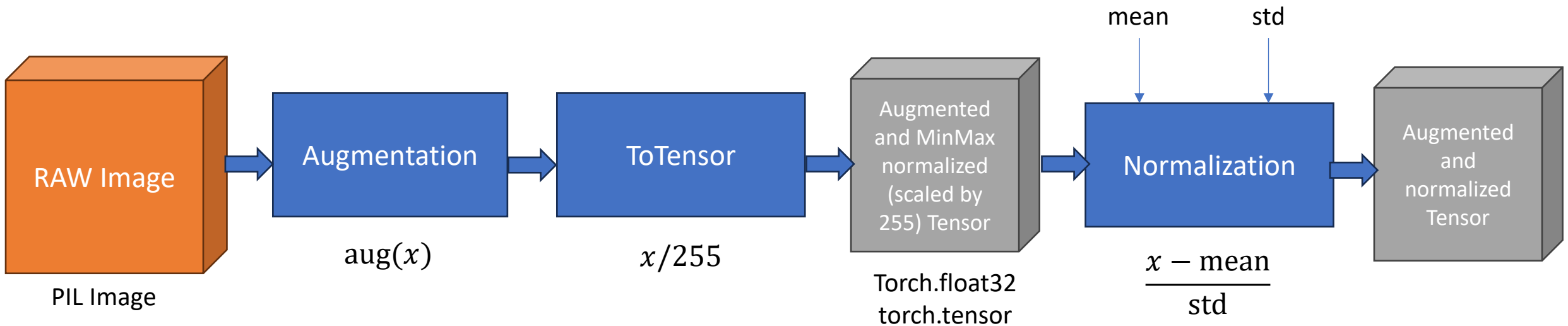
```
Some timm models

resnet18, resnet50
efficientnet_b0,efficientnet_b,
efficientnet_b2
convnext_tiny_in22k
swin_s3_tiny_224
vit_small_patch16_224
```

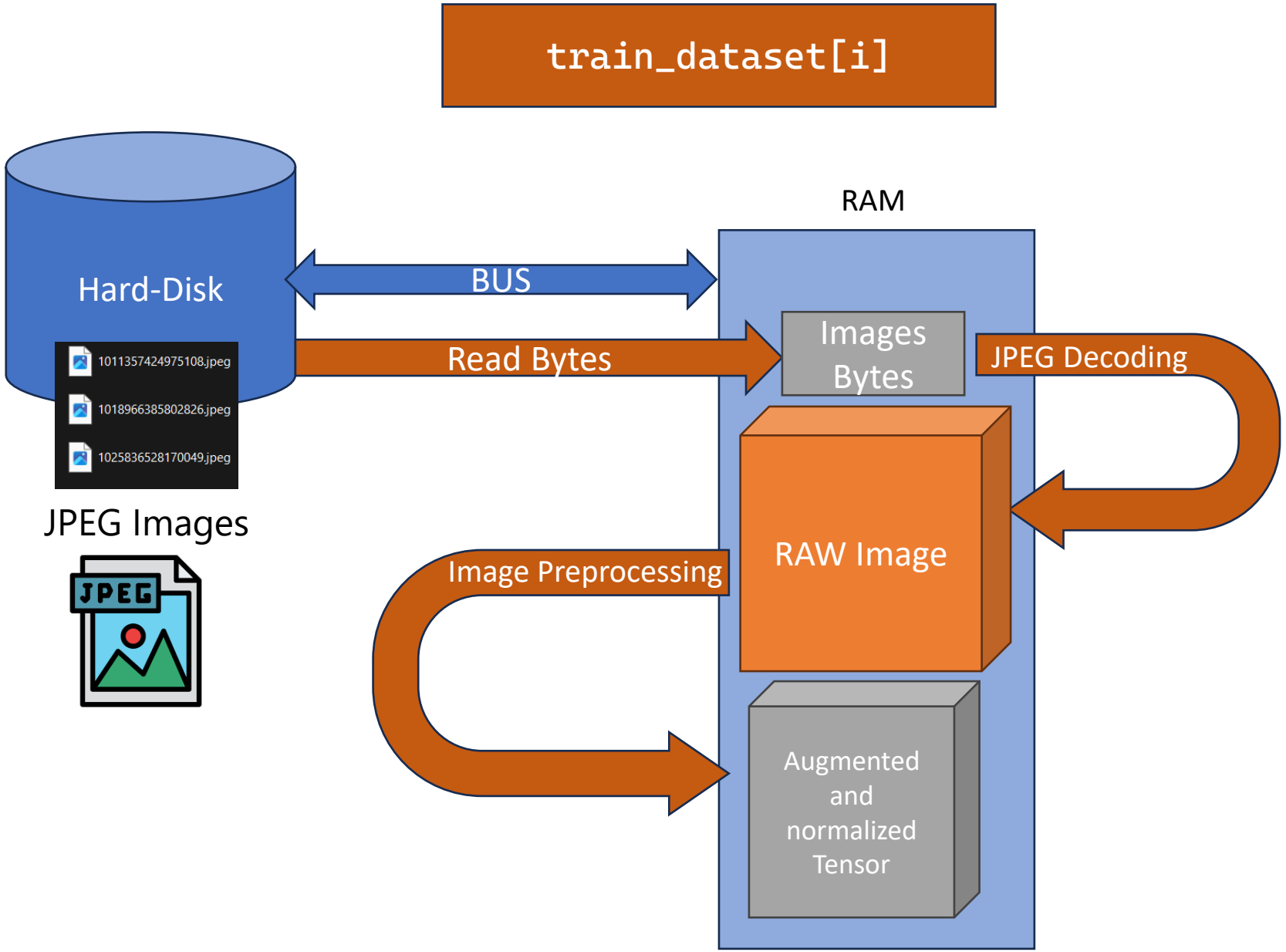
Image Preprocessing – Data Augmentation + Normalization

Albumentations: https://github.com/albumentations-team/albumentations

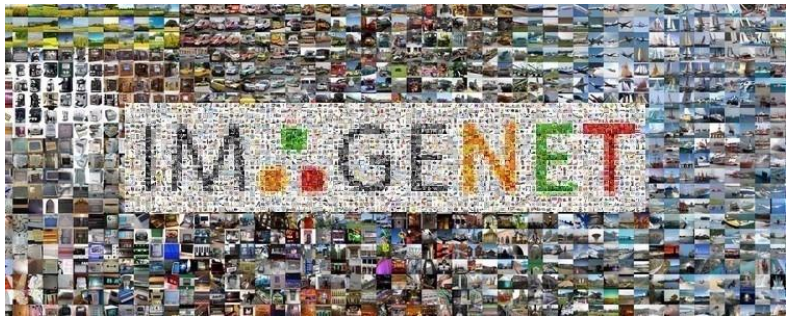
Torchvision: https://pytorch.org/vision/stable/transforms.html



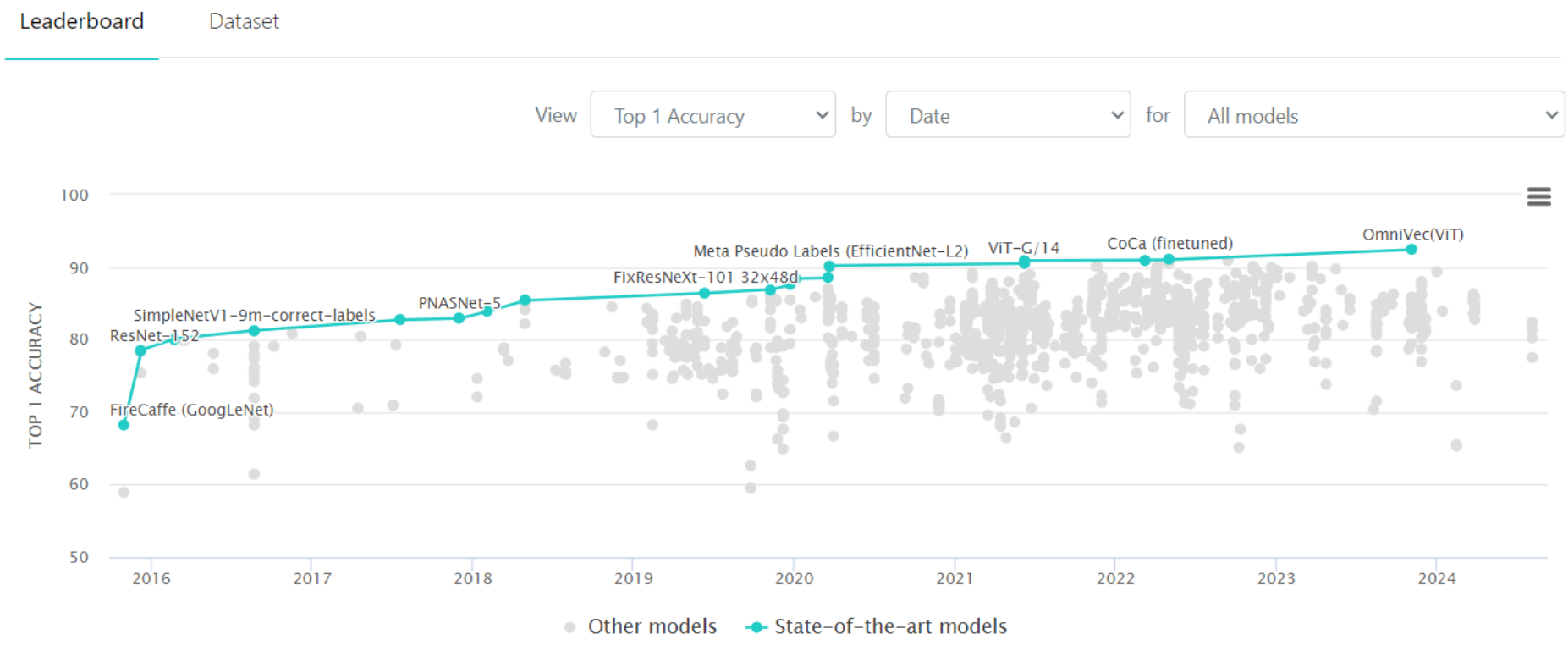
Computer Vision Pytorch Dataset



Transfer Learning

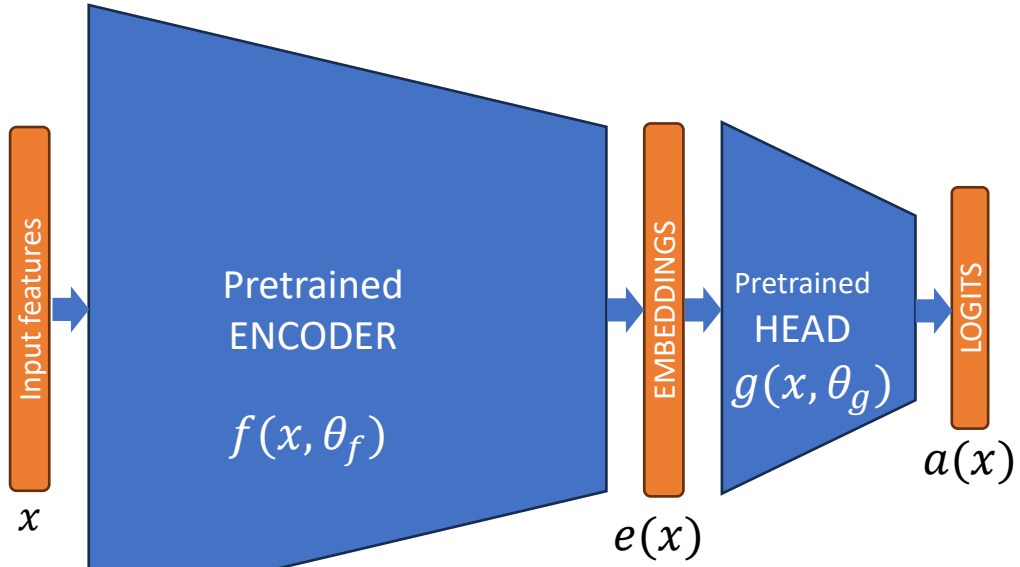


https://www.image-net.org/

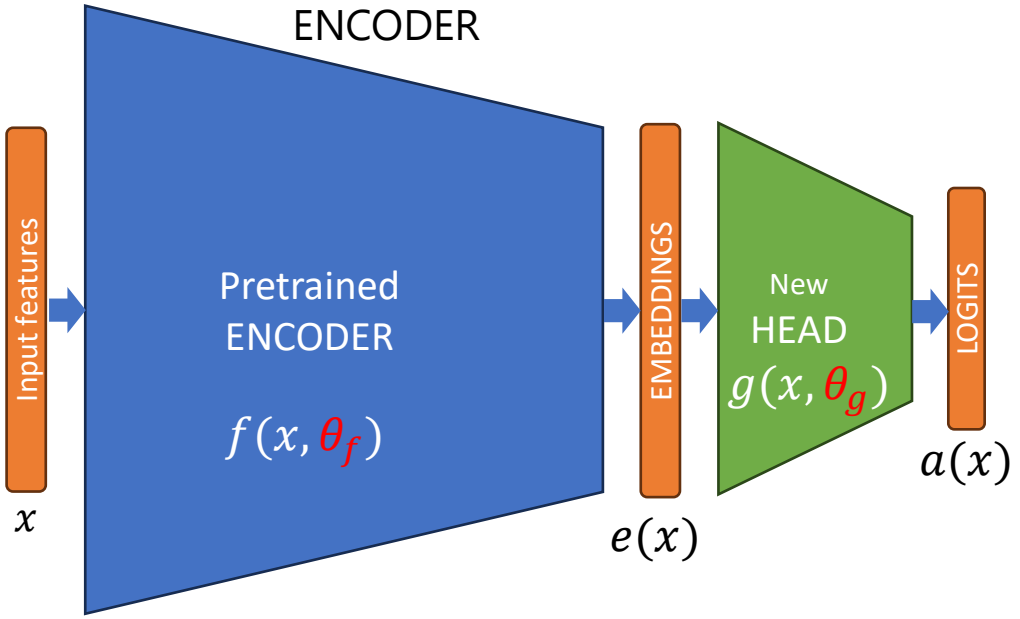


https://paperswithcode.com/sota/image-classification-on-imagenet

Transfer Learning / Finetuning



Pretrained on Imagenet (1000 classes)
Head = Linear(emb_size, 1000)



1. Change the head !

If you freeze some Layer -> transfer learning
If you unfreeze all -> finetuning

Thank you for your attention!