Machine Learning Laboratory

Session 4 – CNN for Classification

Theoretical Preparation

This session focuses on deep learning. We will steer away from implementing complex networks by hand and instead use the PyTorch framework to set up network architectures and automatically optimize the parameters without explicitly defining every gradient function ourselves. As established network structures, we will use VGG16—a 16-layer deep neural network with a simple structure that is still the basis for many other architectures—and ResNet-18, which introduces the super important concept of skip connections.

Deep Learning typically requires larger datasets than classical machine learning approaches. In this session, we will use the PascalVOC dataset, which provides different forms of annotations for various supervised learning setups. This session utilizes the class labels in the set for a multi-label classification, while in the next session you will use the semantic annotation maps for a different learning task.

- **General Deep Learning and VGG16 Architecture**

    - Read through the introduction of the stanford lecture CS231n `https://cs231n.github.io/convolutional-networks/`.
    - Zissermann et al. introduced VGG in `https://arxiv.org/pdf/1409.1556`.
    - Kaiming He et al. introduced Skip Connections and Resnet architectures: `https://arxiv.org/pdf/1512.03385`.
    - Pytorch documentation on BatchNorm2d in `https://docs.pytorch.org/docs/stable/generated/torch.nn.BatchNorm2d.html`
    - Pytorch documentation on Dropout in `https://docs.pytorch.org/docs/stable/generated/torch.nn.Dropout.html`

- **PascalVOC 2012 dataset**

    Browse `http://host.robots.ox.ac.uk/pascal/VOC/voc2012/` to obtain an overview of the dataset. Relevant are the Introduction and Data section.

- **PyTorch**

    To familiarize yourself with the framework, take a look at the basics of PyTorch: `https://docs.pytorch.org/tutorials/beginner/basics/intro.html`