In this course, we will consider the foundations of deep computer vision. The goal of the course is to develop practical skills for solving common computer vision tasks. In the last lecture, we will briefly discuss vision transformers and multimodal approaches. The basic knowledge of PyTorch or other deep learning libraries is assumed for all participants.

1. Introduction. Computer Vision development and trends. Practice: image classification with PyTorch.

2. Working with images: data preparation and augmentation. Practice 1: image processing with Python and TorchVision. Practice 2: saliency maps.

3. Segmentation: Architectures and losses.

4. Detection: Architectures and losses.

5. Foundation models and multimodal approaches. Practice 1: Transformer finetuning. Practice 2: Image quality assessment with CLIP.