

# Syllabus for Advanced Computer Vision Class

Duration: 10 weeks

Week	Content
1	<p>Object classification with SOTA models</p> <ul style="list-style-type: none"> <li>- CNN models: VGG -&gt; ResNet -&gt; MobileNet -&gt; EfficientNet</li> <li>- Attention, Transformer, VisionTransformer, and Swin Transformer</li> </ul>
2	<p>Ways to use pretrained models</p> <ul style="list-style-type: none"> <li>- Transfer learning: Feature Extraction + KNN/ DNN/ SVM</li> <li>- Transfer learning: Fine Tuning</li> </ul> <p><u>Assignment 1</u>: Face Recognition with KNN, SVM, and Fine Tuning</p> <p>Project Introduction</p>
3	<p><b>Object Detection (part 1)</b></p> <ul style="list-style-type: none"> <li>- Task definition and evaluation</li> <li>- Object detection before deep learning</li> <li>- Proposal-driven Detection: RCNN, Fast R-CNN, Faster R-CNN</li> </ul>
4	<p><b>Object Detection (part 2)</b></p> <ul style="list-style-type: none"> <li>- Single-state Object Detection: YOLO, SSD, Feature Pyramid Network, RetinaNet, CornerNet, CenterNet</li> <li>- Object Detection with Transformers: DETection TRansformer - DETR, Pix2Seq</li> </ul> <p><u>Assignment 2</u>: Two-stage object detection - Faster R-CNN</p> <p>Implement a simple two-stage object detection network. In the first part, you will implement a Region Proposal Network and use it to construct your first Faster R-CNN. We will train/evaluate your models on the PASCAL VOC 2007 dataset.</p>

5	<b>From Image classification to Image segmentation</b> <ul style="list-style-type: none"> <li>- Object/Semantic Segmentation</li> <li>- Classical Segmentation algorithms</li> <li>- Reuse Image Classification models in Image Segmentation</li> <li>- Popular architectures</li> <li>- Public datasets</li> <li>- Evaluation metrics</li> </ul>
6	<b>Other tasks in image segmentation</b> <ul style="list-style-type: none"> <li>- Instance/ Panoptic/ Entity Segmentation: architectures, datasets, metrics.</li> <li>- Video Object Segmentation: architectures, datasets, metrics</li> <li>- Interactive Segmentation: architectures, datasets, metrics.</li> </ul> <b>Assignment 3:</b> <ul style="list-style-type: none"> <li>- Facial segmentation: Build an end-to-end facial semantic segmentation on portrait images and manipulate the skin tone, background, eyes, hair color.</li> <li>- Instance segmentation in image fill-in/inpainting: Using instance segmentation to remove/replace objects in images with Stable Diffusion.</li> </ul>
7	<b>Image generation</b> <ul style="list-style-type: none"> <li>- AutoEncoder, Variational AutoEncoder</li> <li>- Generative Adversarial Networks (GANs), CycleGan</li> </ul> <b>Assignment 4:</b> <ul style="list-style-type: none"> <li>- Building AutoEncoder and VAE with pytorch</li> <li>- Building GANs to generate fashion images with pytorch</li> </ul>
8	<b>Visual - Text Multi-modal</b> <ul style="list-style-type: none"> <li>- Image Captioning with CNN - LSTM</li> <li>- Image captioning with Transformer</li> </ul>
9	Project discussion Guest lecture 1
10	Project discussion Guest lecture 2
11	Project representation Summary Certificate ceremony