Modules - Practical Computer Vision with PyTorch

This is a practical online course that will be offered through OpenHPI in May 2025. Learners will gain hands-on experience with modern deep learning techniques for solving computer vision problems, including image classification, object detection, and image generation.

Trailer Video

An overview of the course's content
Trailer video

Prerequisites

A basic knowledge of programming and high-school level math is sufficient to follow the course. The following resources serve as helpful references:

Programming Prerequisites

- Python Fundamentals
  - Basic knowledge of Python
  - Writing functions and classes

Mathematics Prerequisites

- Statistics and Probability
  - Mean, median, and mode
  - Probability
  - Variance and standard deviation
- Linear Algebra
  - Dot product
  - Matrix multiplication
- Calculus
  - Derivatives

Course Modules

Each module is paired with at least one practical programming notebook. Notebooks are currently in development and you can find them here.

Module 1: Foundations

1. Intro to Common Computer Vision Tasks
   - Slides
   - Video
2. Digital Image Representation in PIL and NumPy
   - Slides
   - Video
3. Image Tensors in PyTorch
   - Slides
   - Video

Module 2: Neural Networks Fundamentals

4. Introduction to Neural Networks
   - Slides
   - Video
5. Building a Multilayer Perceptron for Regression in PyTorch
   - Slides
   - Video
6. Matrix Multiplication and Network Shape
   - Slides
   - Video

Module 3: Training and Evaluation

7. Training a Feedforward Network for Classification in PyTorch
   - Slides
   - Video
8. Performance Metrics for Classification and Experiment tracking
   - Slides
   - Video
9. PyTorch Datasets and DataLoaders
   - Slides
   - Video

Module 4: Convolutional Neural Networks

10. Fundamentals of Convolutions

- Slides
- Video

11. Pooling in Neural Networks
    - Slides
    - Video
12. Upsampling and Channel Mixing with Convolutions
    - Slides
    - Video

Module 5: Training Techniques for Image Classification

13. Normalizing Input Values, and Inference with Pretrained Models
    - Slides
    - Video
14. Usage of Binary Cross Entropy Loss
    - Slides
    - Video
15. Skip Connections
    - Slides
    - Video

Module 6: Model Optimization and Interpretability

16. Data Augmentation
    - Slides
    - Video
17. Regularization with Dropout and Batch Normalization
    - Slides
    - Video
18. Transfer Learning and Fine Tuning Pre-trained Models
    - Slides
    - Video
19. Class Activation Mapping
    - Slides
    - Video

Module 7: Image Embeddings

20. Image Embeddings

- Slides
- Video

21. Vision Transformers
    - Slides
    - Video
22. CLIP (Contrastive Language-Image Pretraining)
    - Slides
    - Video

Module 8: Object Detection and Image Segmentation

23. Approaches to Object Detection
    - Slides
    - Video
24. Approaches to Image Segmentation
    - Slides
    - Video

Module 9: Image Generation

25. Image Generation with Diffusion Models
    - Slides
    - Video

GitHub Repository

Notebooks will be extended and published on Kaggle before May 2025
https://github.com/andandandand/practical-computer-vision