The Hong Kong University of Science and Technology

UG Course Syllabus Template

Course Title: Deep Learning in Computer Vision

Course Code: COMP4471

No. of Credits: 3-credit

Any pre-/co-requisites: (COMP 2011 OR COMP 2012 OR COMP 2012H) AND (MATH 2111 OR MATH 2121 OR MATH 2131 OR MATH 2350)

Name: CK Tang

Email: cktang@cse.ust.hk

Course Description

Deep learning has significantly advanced the performance of computer vision system from object recognition to image processing. This course covers the basics and various applications of deep learning in computer vision. Students will study the details of convolutional neural networks as well as recurrent neural networks and train deep networks with end-to-end optimization, and learn deep learning based approaches for both high-level and low-level computer vision tasks such as image recognition and image enhancement. Through programming projects, students will implement, train, and test deep neural networks on cutting-edge computer vision research. Students would be required to study or do research in a final course project related to deep learning and computer vision and present their work by the end of the course.

List of Topics

Topic	Title
1	**Course introduction** Computer vision overview Historical context Course logistics
2	**Image classification** Data-driven approach K-nearest neighbor Linear classification I

3	**Loss function and optimization** Linear classification II Higher-level representations, image features Optimization, stochastic gradient descent
4	**Introduction to neural networks** Backpropagation Multi-layer Perceptrons The neural viewpoint
5	**Convolutional neural networks** History Convolution and pooling ConvNets outside vision
6	**Training neural networks, part I** Activation functions, initialization, dropout, batch normalization
7	**Training neural networks, part II** Update rules, ensembles, data augmentation, transfer learning
8	**Deep learning hardware & software** Caffe, Torch, Theano, TensorFlow, Keras, PyTorch, etc
9	**CNN architectures** AlexNet, VGG, GoogLeNet, ResNet, etc
10	**Recurrent neural networks** RNN, LSTM, GRU Language modeling Image captioning, visual question answering Soft attention
11	**Detection and segmentation** Semantic segmentation Object detection Instance segmentation
12	**Visualizing and understanding** Feature visualization and inversion Adversarial examples Style transfer
13	**Generative models** PixelRNN/CNN Variational Autoencoders Generative Adversarial Networks Diffusion Models

14	**Attention and Transformers**
15	**Video Models**
16	**Self-supervised Learning**
17	**Generative AI for Visual Content Generation**
18	**Deep 3D Vision**
19	**Denosing Diffusion Model**
20	**Deep Reinforcement Learning for Robots**

Textbooks

Ian Goodfellow, Yoshua Bengio and Aaron Courville. *Deep Learning*, MIT Press, 2016.
Aston Zhang, Zachary C. Lipton, Mu Li, Alexander J. Smola. Dive into Deep Learning. 2021
https://d2l.ai/

Reference books

N/A

Grading Scheme

3 Programming Assignments	36%
Midterm exam	35%
Final Project	29%
Total	100%

Course Intended Learning Outcomes

1. Students understand the basics of deep neural networks
2. Students can train deep neural networks on several computer vision tasks
3. Students can use deep learning as a tool to solve a research problem of their interests

Assessment Rubrics

N/A