

(http://svl.stanford.edu) (http://stanford.edu/)



CS231n: Deep Learning for Computer Vision

Stanford - Spring 2022

Schedule

- **Lectures** will occur Tuesday/Thursday from 1:30-3:00pm Pacific Time at NVIDIA Auditorium (https://goo.gl/maps/hRjQYd6MqxB2).
- **Discussion** sections will (generally) occur on Fridays between 1:30-2:30pm Pacific Time on Zoom. Check Ed (https://edstem.org/us/courses/21177) for any exceptions.

Updated lecture slides will be posted here shortly before each lecture. For ease of reading, we have color-coded the lecture category titles in blue, discussion sections (and final project poster session) in yellow, and the midterm exam in red. Note that the schedule is subject to change as the quarter progresses.

Date		
Description		
Course Materials		
Events		
Deadlines		

03/29

Lecture 1: Introduction

Computer vision overview

Historical context Course logistics

[slides 1 (slides/2022/lecture_1_1_feifei.pdf)] [slides 2 (slides/2022/lecture_1_2_ruohan.pdf)]

Deep Learning Basic	ne e		
Deep Learning Dask	,3 		
N2/21			

Lecture 2: Image Classification with Linear Classifiers

The data-driven approach

K-nearest neighbor

Linear Classifiers

Algebraic / Visual / Geometric viewpoints

SVM and Softmax loss

[slides (slides/2022/lecture_2_ruohan.pdf)]

Image Classification Problem (https://cs231n.github.io/classification/)

Linear Classification (https://cs231n.github.io/linear-classify/)

04/01

Python / Numpy Review Session

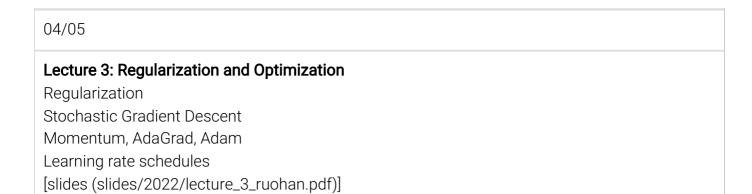
[Colab

(https://colab.research.google.com/github/cs231n/cs231n.github.io/blob/master/python-colab.ipynb)] [Tutorial (https://cs231n.github.io/python-numpy-tutorial/)]

② 1:30-2:30pm PT

Assignment 1 out

[handout (https://cs231n.github.io/assignments2022/assignment1/)] [colab (https://cs231n.github.io/assignments/2022/assignment1_colab.zip)]



Optimization (https://cs231n.github.io/optimization-1/)

04/07

Lecture 4: Neural Networks and Backpropagation

Multi-layer Perceptron

Backpropagation

[slides (slides/2022/lecture_4_ruohan.pdf)]

Backprop (http://cs231n.github.io/optimization-2)

Linear backprop example (handouts/linear-backprop.pdf)

Suggested Readings:

- 1. Why Momentum Really Works (https://distill.pub/2017/momentum/)
- 2. Derivatives notes (handouts/derivatives.pdf)
- 3. Efficient backprop (http://yann.lecun.com/exdb/publis/pdf/lecun-98b.pdf)
- 4. More backprop references: [1] (http://colah.github.io/posts/2015-08-Backprop/), [2] (http://neuralnetworksanddeeplearning.com/chap2.html), [3] (https://www.youtube.com/watch?v=q0pm3BrIUFo)

04/08

Backprop Review Session [slides (slides/2022/discussion_2_backprop.pdf)]

② 1:30-2:30pm PT

Perceiving and Understanding the Visual World
04/12
Lecture 5: Image Classification with CNNs History
Higher-level representations, image features Convolution and pooling [slides (slides/2022/lecture_5_ruohan.pdf)]
Convolutional Networks (http://cs231n.github.io/convolutional-networks)
04/13
Final Project Overview and Guidelines [slides (slides/2022/discussion_3_project.pdf)]
② 3:00-4:00pm PT
04/14
Lecture 6: CNN Architectures Batch Normalization Transfer learning AlexNet, VGG, GoogLeNet, ResNet [slides (slides/2022/lecture_6_jiajun.pdf)]
AlexNet (https://papers.nips.cc/paper/4824-imagenet-classification-with-deep-convolutional-neural-networks.pdf), VGGNet (https://arxiv.org/abs/1409.1556), GoogLeNet (https://arxiv.org/abs/1409.4842), ResNet (https://arxiv.org/abs/1512.03385)

04/15	
Assignment 2 out [handout (https://cs231n.github.io/assignments2022/assignment2/)] [colab (https://cs231n.github.io/assignments/2022/assignment2_colab.zip)]	
Assignment 1 due	

04/18		
Project proposal due		

Lecture 7: Training Neural Networks

Activation functions

Data processing

Weight initialization

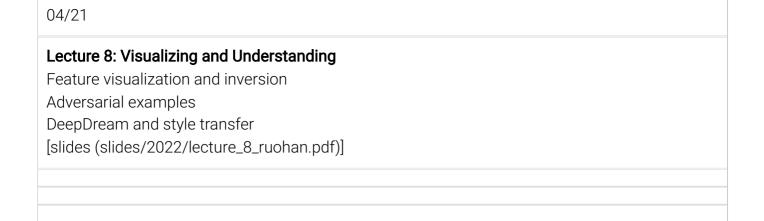
Hyperparameter tuning

Data augmentation

[slides (slides/2022/lecture_7_ruohan.pdf)]

Neural Networks, Parts 1 (http://cs231n.github.io/neural-networks-1), 2 (http://cs231n.github.io/neural-networks-2), 3 (http://cs231n.github.io/neural-networks-3) Suggested Readings:

- 1. Stochastic Gradient Descent Tricks (http://research.microsoft.com/pubs/192769/tricks-2012.pdf)
- 2. Efficient Backprop (http://yann.lecun.com/exdb/publis/pdf/lecun-98b.pdf)
- 3. Practical Recommendations for Gradient-based Training (https://arxiv.org/pdf/1206.5533v2.pdf)
- 4. Deep Learning, Nature 2015 (https://www.nature.com/articles/nature14539)
- 5. An Overview of Gradient Descent Algorithms (https://ruder.io/optimizing-gradient-descent/)
- 6. A Disciplined Approach to Neural Network Hyper-Parameters (https://arxiv.org/abs/1803.09820)



PyTorch Review Session [slides (slides/2022/discussion_4_pytorch.pdf)]

② 1:30-2:30pm PT

04/26

Lecture 9: Object Detection and Image Segmentation

Single-stage detectors

Two-stage detectors

Semantic/Instance/Panoptic segmentation

[slides (slides/2022/lecture_9_jiajun.pdf)]

FCN (https://arxiv.org/abs/1411.4038), R-CNN (https://arxiv.org/abs/1311.2524), Fast R-CNN (https://arxiv.org/abs/1504.08083), Faster R-CNN (https://arxiv.org/abs/1506.01497), YOLO (https://arxiv.org/abs/1506.02640)

04/28
Lecture 10: Recurrent Neural Networks RNN, LSTM, GRU Language modeling Image captioning Sequence-to-sequence [slides (slides/2022/lecture_10_ruohan.pdf)]
Suggested Readings: 1. DL book RNN chapter (http://www.deeplearningbook.org/contents/rnn.html) 2. Understanding LSTM Networks (https://colah.github.io/posts/2015-08-Understanding-LSTMs/)
04/29
Object Detection & RNNs Review Session [slides (slides/2022/discussion_5_detection.pdf)]
② 2:30-3:30pm PT
05/02
Assignment 2 due

05/03	
00,00	

Lecture 11: Attention and Transformers

Self-Attention

Transformers

[slides (slides/2022/lecture_11_ruohan.pdf)]

Suggested Readings:

- 1. Attention is All You Need [Original Transformers Paper (https://arxiv.org/abs/1706.03762)]
- 2. Attention? Attention [Blog by Lilian Weng (https://lilianweng.github.io/lillog/2018/06/24/attention-attention.html)]
- 3. The Illustrated Transformer [Blog by Jay Alammar (http://jalammar.github.io/illustrated-transformer/)]
- 4. ViT: Transformers for Image Recognition [Paper (https://arxiv.org/abs/2010.11929)] [Blog (https://ai.googleblog.com/2020/12/transformers-for-image-recognition-at.html?m=1)] [Video (https://www.youtube.com/watch?v=TrdevFK_am4)]
- 5. DETR: End-to-End Object Detection with Transformers [Paper (https://arxiv.org/abs/2005.12872)] [Blog (https://ai.facebook.com/blog/end-to-end-object-detection-with-transformers/)] [Video (https://www.youtube.com/watch?v=utxbUlo9CyY)]

05/5

Lecture 12: Video Understanding

Video classification

3D CNNs

Two-stream networks

Multimodal video understanding

[slides (slides/2022/lecture_12_ruohan.pdf)]

05/06

Midterm Review Session

2:30-3:30pm PT

05/07
Project milestone due
05/10
In-Class Midterm
② 1:30-3:00pm
Assignment 3 out [handout (https://cs231n.github.io/assignments2022/assignment3/)] [colab (https://cs231n.github.io/assignments/2022/assignment3_colab.zip)]
Reconstructing and Interacting with the Visual World
05/12
Lecture 13: Generative Models
Supervised vs. Unsupervised learning
Pixel RNN, Pixel CNN
Variational Autoencoders
Generative Adversarial Networks
[slides (slides/2022/lecture_13_jiajun.pdf)]
Suggested Readings:
1. Image GPT: Generative Pretraining From Pixels [Paper
(https://cdn.openai.com/papers/Generative_Pretraining_from_Pixels_V2.pdf)] [Blog
(https://openai.com/blog/image-gpt/)]

Lecture 14: Self-supervised Learning

Pretext tasks

Contrastive learning

Multisensory supervision

[slides (slides/2022/lecture_14_jiajun.pdf)]

Suggested Readings:

- 1. Lilian Weng Blog Post (https://lilianweng.github.io/lil-log/2019/11/10/self-supervised-learning.html)
- 2. DINO: Emerging Properties in Self-Supervised Vision Transformers [Paper (https://arxiv.org/abs/2104.14294)] [Blog (https://ai.facebook.com/blog/dino-paws-computer-vision-with-self-supervised-transformers-and-10x-more-efficient-training)] [Video (https://youtu.be/h3ij3F3cPlk)]

05/19

Lecture 15: Low-Level Vision

(Guest Lecture by Prof. Jia Deng (https://www.cs.princeton.edu/~jiadeng/) from Princeton University)

Optical flow

Depth estimation

Stereo vision

[slides (slides/2022/lecture_15_jia.pdf)]

05/24

Lecture 16: 3D Vision

3D shape representations

Shape reconstruction

Neural implicit representations

[slides (slides/2022/lecture_16_jiajun.pdf)]

Assignment 3 due

Human-Centered Applications and Implications
05/26
Lecture 17: Human-Centered Artificial Intelligence Al & healthcare
05/31
Lecture 18: Fairness in Visual Recognition
(Guest Lecture by Prof. Olga Russakovsky (https://www.cs.princeton.edu/~olgarus/) from
Princeton University)
06/02
Project final report due
06/04
Final Project Poster Session
Note: Only open to the Stanford community and invited guests.
② 3:30-6:30pm
Location: Alumni Center McCaw Hall/Ford Gardens
Click here (https://edstem.org/us/courses/21177/discussion/1553751) for the logistics and expectations.

06/05		
Project poster PDF due		