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

CS231n: Deep Learning for Computer Vision

Stanford - Spring 2023

Schedule

- **Lectures** will occur Tuesday/Thursday from 12:00-1:20pm Pacific Time at NVIDIA Auditorium (https://goo.gl/maps/hRjQYd6MqxB2).
- **Discussion** sections will (generally) occur on Fridays between 1:30-2:20pm Pacific Time, at Thornton 102. Check Ed (https://edstem.org/us/courses/38489) 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

04/04

Lecture 1: Introduction
Computer vision overview
Course overview
Course logistics
[slides 1 (slides/2023/lecture_1_part_1.pdf)] [slides 2 (slides/2023/lecture_1_part_2.pdf)]

———

Deep Learning Basics

04/06

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/2023/lecture_2.pdf)]

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

04/07
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/)]
1:30-2:20pm PT
Assignment 1 **out** [handout (https://cs231n.github.io/assignments2023/assignment1/)] [colab (https://cs231n.github.io/assignments/2023/assignment1_colab.zip)]

04/11
Lecture 3: Regularization and Optimization Regularization Stochastic Gradient Descent Momentum, AdaGrad, Adam Learning rate schedules [slides (slides/2023/lecture_3.pdf)]
Optimization (https://cs231n.github.io/optimization-1/)

04/13

Lecture 4: Neural Networks and Backpropagation
Multi-layer Perceptron
Backpropagation
[slides (slides/2023/lecture_4.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)

04/14

Backprop Review Session
[slides (slides/2023/section_2.pdf)] [Colab (https://colab.research.google.com/drive/1FSP70oJG9RQtAqq67UkQws7D62jvJoRo?usp=sharing)]

1:30-2:20pm PT

———

Perceiving and Understanding the Visual World

04/18

Lecture 5: Image Classification with CNNs
History
Higher-level representations, image features
Convolution and pooling
[slides (slides/2023/lecture_5.pdf)]

Convolutional Networks (http://cs231n.github.io/convolutional-networks)

04/20

Lecture 6: CNN Architectures
Batch Normalization
Transfer learning
AlexNet, VGG, GoogLeNet, ResNet
[slides (slides/2023/lecture_6.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)

04/21

Final Project Overview and Guidelines
[slides (slides/2023/section_3.pdf)]

1:30-2:20pm PT

Assignment 2 **out**
[handout (https://cs231n.github.io/assignments2023/assignment2/)] [colab (https://cs231n.github.io/assignments/2023/assignment2_colab.zip)]

Assignment 1 **due**

04/24

Project proposal **due**

04/25

Lecture 7: Training Neural Networks
Activation functions
Data processing
Weight initialization
Hyperparameter tuning
Data augmentation
[slides (slides/2023/lecture_7.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)

04/27

Lecture 8: Recurrent Neural Networks
RNN, LSTM, GRU
Language modeling
Image captioning
Sequence-to-sequence
[slides (slides/2023/lecture_8.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/)

04/28

PyTorch Review Session
[Colab (https://colab.research.google.com/drive/17i_lVAcObw_63ta_ao69EgS0AUQzMMmu?usp=sharing)]

1:30-2:20pm PT

05/02

Lecture 9: Attention and Transformers
Self-Attention
Transformers
[slides (slides/2023/lecture_9.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/lil-log/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=TrdevFK_am://www.youtube.com/watch?v=utxbUlo9CyY)]

05/04

Lecture 10: Video Understanding
Video classification
3D CNNs
Two-stream networks
Multimodal video understanding
[slides (slides/2023/lecture_10.pdf)]

05/05
RNNs & Transformers [Colab (https://colab.research.google.com/drive/1eFPBYzHy17svdTfL2FnPZOzBwOUQ-Aky?usp=sharing)] [slides (slides/2023/section_5.pdf)]
1:30-2:20pm PT

05/08
Assignment 2 **due**

05/09
Lecture 11: Object Detection and Image Segmentation Single-stage detectors Two-stage detectors Semantic/Instance/Panoptic segmentation [slides (slides/2023/lecture_11.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)

05/11
Lecture 12: Visualizing and Understanding Feature visualization and inversion Adversarial examples DeepDream and style transfer [slides (slides/2023/lecture_12.pdf)]

05/12

Midterm Review Session
[slides (https://docs.google.com/presentation/d/1w9kliBV7awiKNbAf97ci8MLRKjWF5Mw5pq_yfqhGauI/edit?usp=sharing)]

1:30-2:20pm PT

05/13

Project milestone **due**

———

Generative and Interactive Visual Intelligence

05/16

In-Class Midterm

12:00-1:20pm

Assignment 3 **out**

05/18

Lecture 13: Self-supervised Learning
Pretext tasks
Contrastive learning
Multisensory supervision
[slides (slides/2023/lecture_13.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/h3ij3F3cPIk)]

05/23

Lecture 14: Robot Learning
Deep Reinforcement Learning
Model Learning
Robotic Manipulation
[slides (slides/2023/lecture_14.pdf)]

05/25

Lecture 15: Generative Models (Guest Lecture by Dr. Ruiqi Gao (https://ruiqigao.github.io/) from Google DeepMind)
Generative Adversarial Network
Diffusion models
Autoregressive models
[slides (slides/2023/lecture_15.pdf)]

05/30

Lecture 16: 3D Vision
3D shape representations
Shape reconstruction
Neural implicit representations
[slides (slides/2023/lecture_16.pdf)]

Assignment 3 **due**

———

Human-Centered Applications and Implications

06/01

Lecture 17: Human-Centered Artificial Intelligence

06/06

Lecture 18: Guest Lecture by Prof. Sara Beery (https://beerys.github.io/) from MIT

06/08

Project final report **due**

06/14
Final Project Poster Session Time: 1:00 PM - 4:30 PM Location: Burnham Pavilion Session A Check-in: 12:30 PM - 1:00 PM (30 minutes) Session A: 1:00 PM - 2:30 PM (90 minutes) Session B Check-in: 2:30 PM - 3:00 PM (30 minutes) Session B: 3:00 PM - 4:30 PM (90 minutes)