上海科技大学 ShanghaiTech University

Project Practice for Deep Learning

Yujiao Shi

SIST, ShanghaiTech

Spring, 2024

Outline

- Course Logistics
  - Overall Objective
  - Syllabus/ Projects
- Introduction to Deep Learning

Course Objectives

- Understanding deep networks
  - Key concepts and principles
  - Selection for hyper-parameters, loss function, etc.
  - Deep network design and fine-tuning
- Applying standard networks to solve real problems
  - Data collection and processing
  - Model evaluation
  - Application to downstream tasks

Syllabus & Schedule

上海科技大学 ShanghaiTech University

- Part I: Basics
  - Basic neural networks (MLP, CNN, RNN and Transformer, Week 1-3)
- Part II: Projects
  1) Image super-resolution (2D, Week 4)
  2) Camera localization (2D &/ 3D, Week 5)
  3) Self-selected Project (as a group, with 2~3 members, Week 6)
  4) Hand pose estimation (2D to 3D, Week 7)
  5) 3D reconstruction and novel view synthesis (2D to/from 3D, Week 8)
  6) Generative models (Multi-modality, Week 9)
- Part III: Practice Weeks (Week 10 - 11)
- Final presentation of self-selected project (Week 12)

Course Overview

上海科技大学
ShanghaiTech University

Class Participation: 10%

Projects: 15% x 6

Projects

(1) (2) (3) (4) (5) (6)

	八月		九月				十月					十一月				十二月				一月		
星期一	19	26	2	9	16	23	30	7	14	21	28	4	11	18	25	2	9	16	23	30	6	13
星期二	20	27	3	10	17 中秋节	24	1 国庆节	8	15	22	29	5	12	19	26	3	10	17	24	31	7	14
星期三	21	28	4	11	18	25	2	9	16	23	30	6	13	20	27	4	11	18	25	1 元旦	8	15
星期四	22	29	5	12	19	26	3	10	17	24	31	7	14	21	28	5	12	19	26	2	9	16
星期五	23	30	6	13	20	27	4	11	18	25	1	8	15	22	29	6	13	20	27	3	10	17
星期六	24	31	7	14	21	28	5	12	19	26	2	9	16	23	30	7	14	21	28	4	11	18
星期日	25	1	8	15	22	29	6	13	20	27	3	10	17	24	1	8	15	22	29	5	12	19
周数	4	5	6	7	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
学期	暑假											秋学期										

Deep Learning Introductions

Projects practice

Final Presentation

Course Overview

上海科技大学
ShanghaiTech University

Class Participation: 10%

Projects: 15% x 6

Projects

(1) (2) (3) (4) (5) (6)

	八月		九月				十月					十一月				十二月				一月		
星期一	19	26	2	9		23	30	7	14	21	28	4	11	18	25	2	9	16	23	30	6	13
星期二	20	27	3	10	17 中秋节	24	1 国庆节	8	15	22	29	5	12	19	26	3	10	17	24	31	7	14
星期三	21	28	4	11	18	25	2	9	16	23	30	6	13	20	27	4	11	18	25	1 元旦	8	15
星期四	22	29	5	12	19	26	3	10	17	24	31	7	14	21	28	5	12	19	26	2	9	16
星期五	23	30	6	13	20	27		11	18	25	1	8	15	22	29	6	13	20	27	3	10	17
星期六	24	31	7	14	21	28	5	12	19	26	2	9	16	23	30	7	14	21	28	4	11	18
星期日	25	1	8	15	22	29	6	13	20	27	3	10	17	24	1	8	15	22	29	5	12	19
周数	4	5	6	7	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
学期	暑假				秋学期																	

Deep Learning Introductions

Projects practice

Final Presentation

Course Overview

Class Participation: 10%

Projects: 15% x 6

Projects

(2) (3) (4) (5) (6)

	八月		九月				十月					十一月				十二月				一月		
星期一	19	26	2	9	16	23	30	7	14	21	28	4	11	18	25	2	9	16	23	30	6	13
星期二	20	27	3	10	17 中秋节	24	1 国庆节	8	15	22	29	5	12	19	26	3	10	17	24	31	7	14
星期三	21	28	4	11	18	25	2	9	16	23	30	6	13	20	27	4	11	18	25	1 元旦	8	15
星期四	22	29	5	12	19	26	3	10	17	24	31	7	14	21	28	5	12	19	26	2	9	16
星期五	23	30	6	13	20	27	4	11	18	25	1	8	15	22	29	6	13	20	27	3	10	17
星期六	24	31	7	14	21	(1) 28	5	12	19	26	2	9	16	23	30	7	14	21	28	4	11	18
星期日	25	1	8	15	22	29	6	13	20	27	3	10	17	24	1	8	15	22	29	5	12	19
周数	4	5	6	7	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
学期	暑假				秋学期																	

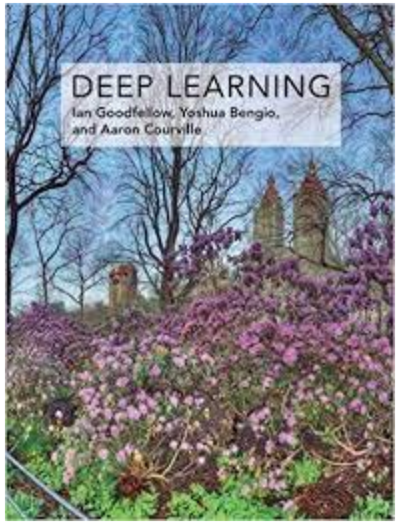
Deep Learning Introductions

Projects practice

Final Presentation

Reference books and materials

- Deep learning:
  - http://www.deeplearningbook.org/
  - https://d2l.ai/
- Online deep learning courses:
  - Stanford: CS230, CS231n
  - CMU: 11-785
  - MIT: 6.S191
- Additional reading materials on Piazza
  - Survey papers, tutorials, etc.



Instructor and TAs

上海科技大学 ShanghaiTech University

- Instructor: Prof Yujiao Shi
  - shiyj2@shanghaitech.edu.cn
  - SIST 1C-303C

- TAs:
  - Shaoxun Wu; Jiawei Yang； Jiaqi Yang

- Office hours: To be announced on Piazza

- We will use Piazza as the main communication platform

Grading policy

- Class Participation: 10%

- Five Projects with Pre-determined Topics: 15% x 5
  - Work independently
  - 12% basic + 3% advanced

- One Self-determined Project: 15%
  - May work as a group, 2~3 students
  - Project proposal
  - Final presentation + demo

Administrative Stuff

- Plagiarism
  - All assignments must be done individually
    - You may not look at solutions from any other source
    - You may not share solutions with any other students
    - Plagiarism detection software will be used on all the programming assignments
    - You may discuss together or help another student but you cannot give the exact solution
- Plagiarism punishment
  - When one student copies from another student, both students are responsible
  - Zero point on the assignment
  - Repeated violation will result in an F grade for this course as well as further discipline at the school/university level

Pre-requisite

上海科技大学 ShanghaiTech University

- Proficiency in Python
  - All class assignments will be in Python (and use numpy)

- Calculus, Linear Algebra, Probability and Statistics
  - Undergrad course level

- Equivalent knowledge of Andrew Ng’s CS229 (Machine Learning)
  - Formulating cost functions
  - Taking derivatives
  - Performing optimization with gradient descent

Outline

上海科技大学
ShanghaiTech University

- Course Logistics
  - Overall Objective
  - Syllabus/ Projects
- Introduction to Deep Learning

Introduction

- Our goal: Build intelligent algorithms to make sense of data
  - Example: Recognizing objects in images

red panda (*Ailurus fulgens*)

  - Example: Predicting what would happen next

Vondrick et al. CVPR2016

Introduction

- Our goal: Build intelligent algorithms to make sense of data
  - Example: Recognizing objects in images
  - Example: Predicting what would happen next

Given an initial still frame,

Introduction

上海科技大学 ShanghaiTech University

- A broad range of real-world applications
  - Speech recognition
    - Input: sound wave → Output: transcript
  - Language translation
    - Input: text in language A (Eng) → Output: text in language B (Chs)
  - Image classification
    - Input: images → Output: image category (cat, dog, car, house, etc.)
  - Autonomous driving
    - Input: sensory inputs → Output: actions (straight, left, right, stop, etc.)
- Main challenges: difficult to manually design the algorithms

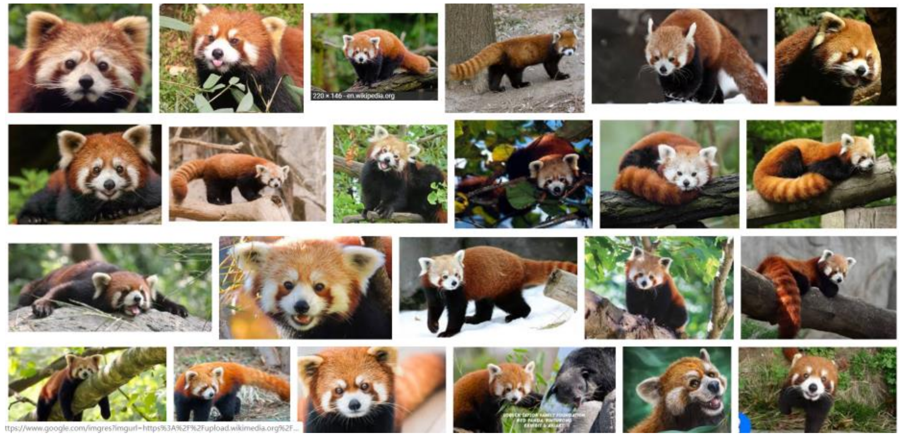
A data-driven approach

- Each task as a mapping function (or a model)

Input data → Mapping function → Expected output

  - input data: images
  - expected output: object or action names

- Building such mapping functions from data



red panda (*Ailurus fulgens*)

→ Mapping function

A data-driven approach

上海科技大学
ShanghaiTech University

- Building a mapping function (model)

$$y = f(x; \theta)$$

  - x: input data
  - y: expected output
  - θ: parameters to be estimated

- Learning the model from data
  - Given a dataset $\mathcal{D} = \{(x_n, y_n)\}_{n=1}^N$
  - Find the 'best' parameter $\hat{\theta}$, such that

$$y_n \simeq f(x_n; \hat{\theta}) \quad \forall n$$

  - And it can be generalized to unseen input data

What is deep learning?

- Using deep neural networks as the mapping function

- Model: Deep neural networks
  - A family of parametric models
  - Consisting of many ‘simple’ computational units
  - Constructing a multi-layer representation of input

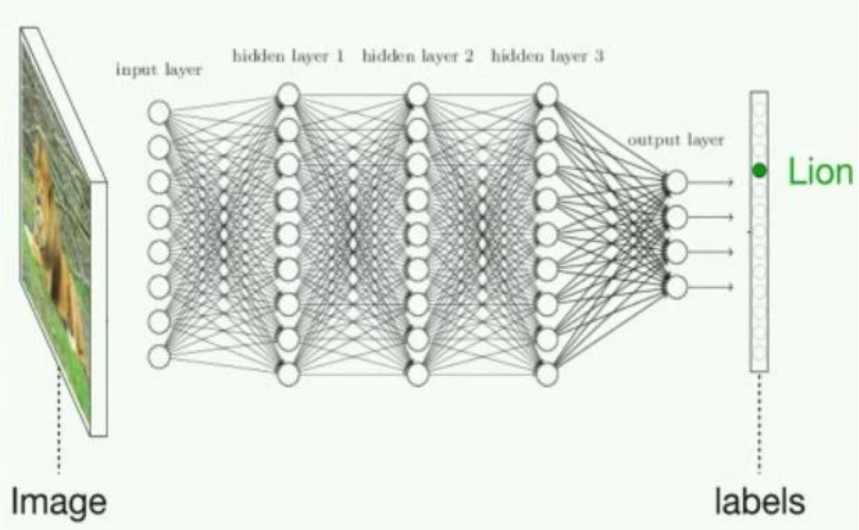
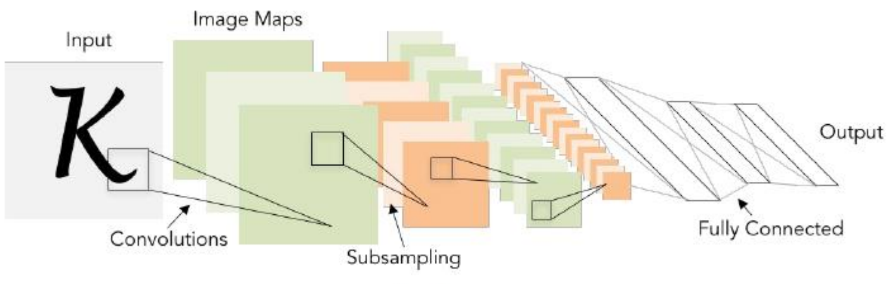


Image from Jeff Clune’s Deep Learning Overview

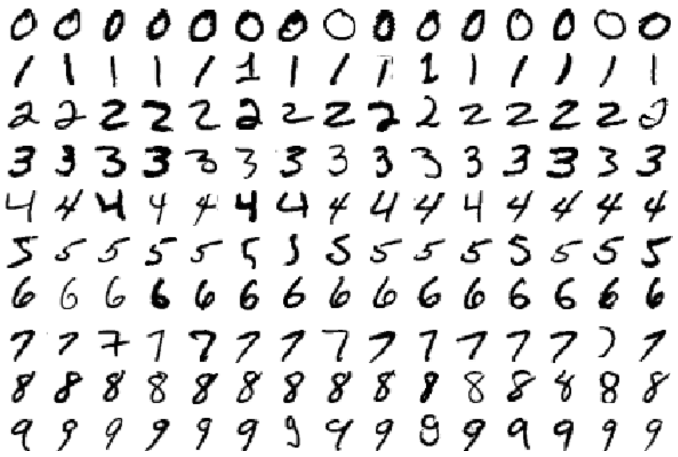
What is deep learning?

- Using deep neural networks as the mapping function

- Learning: Parameter estimation from data
  - Parameters: connection weights between units
  - Formulated as an optimization problem
  - Efficient algorithms for handling large-scale models & datasets



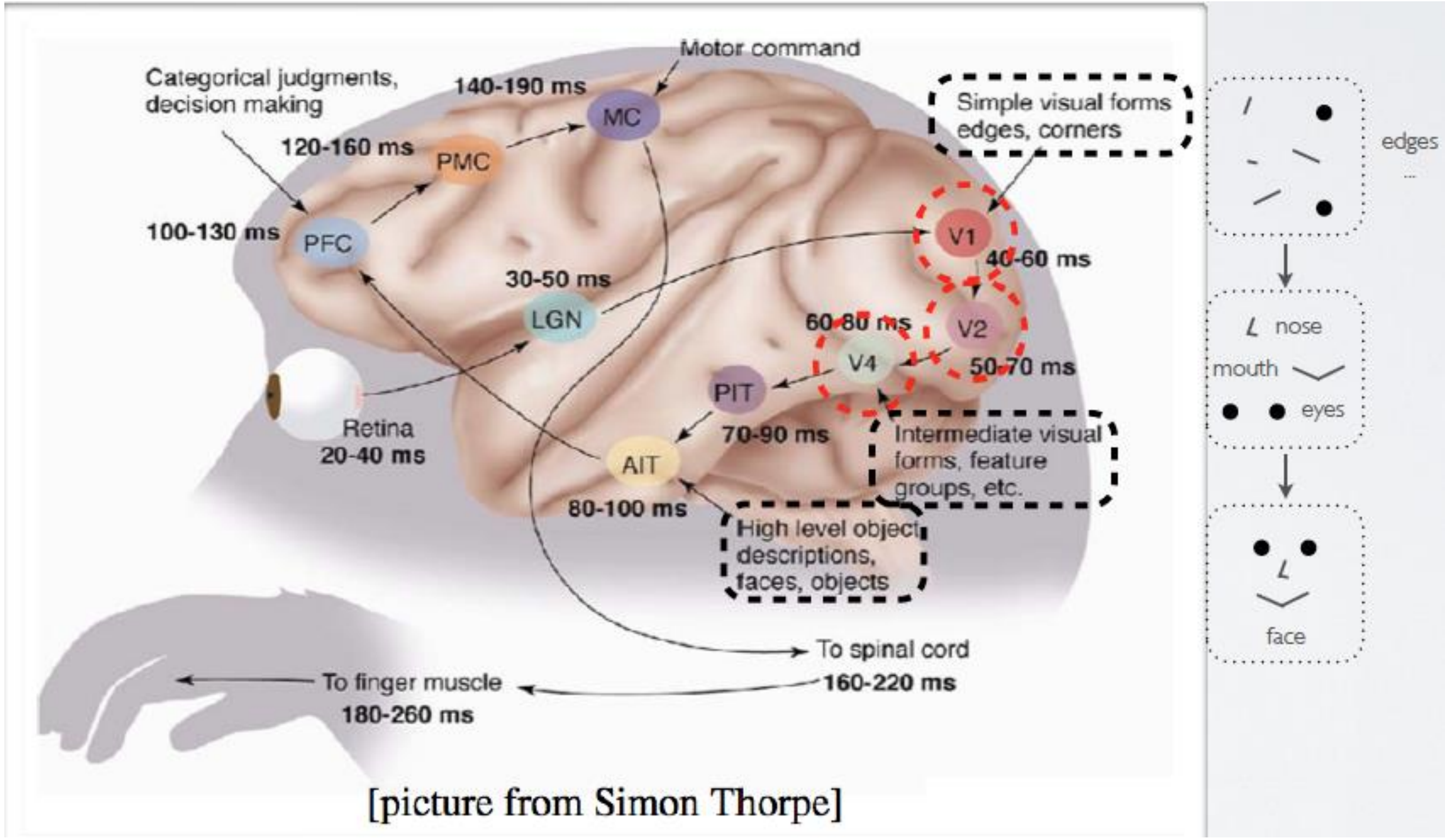
LeNet-5



Why deep networks?

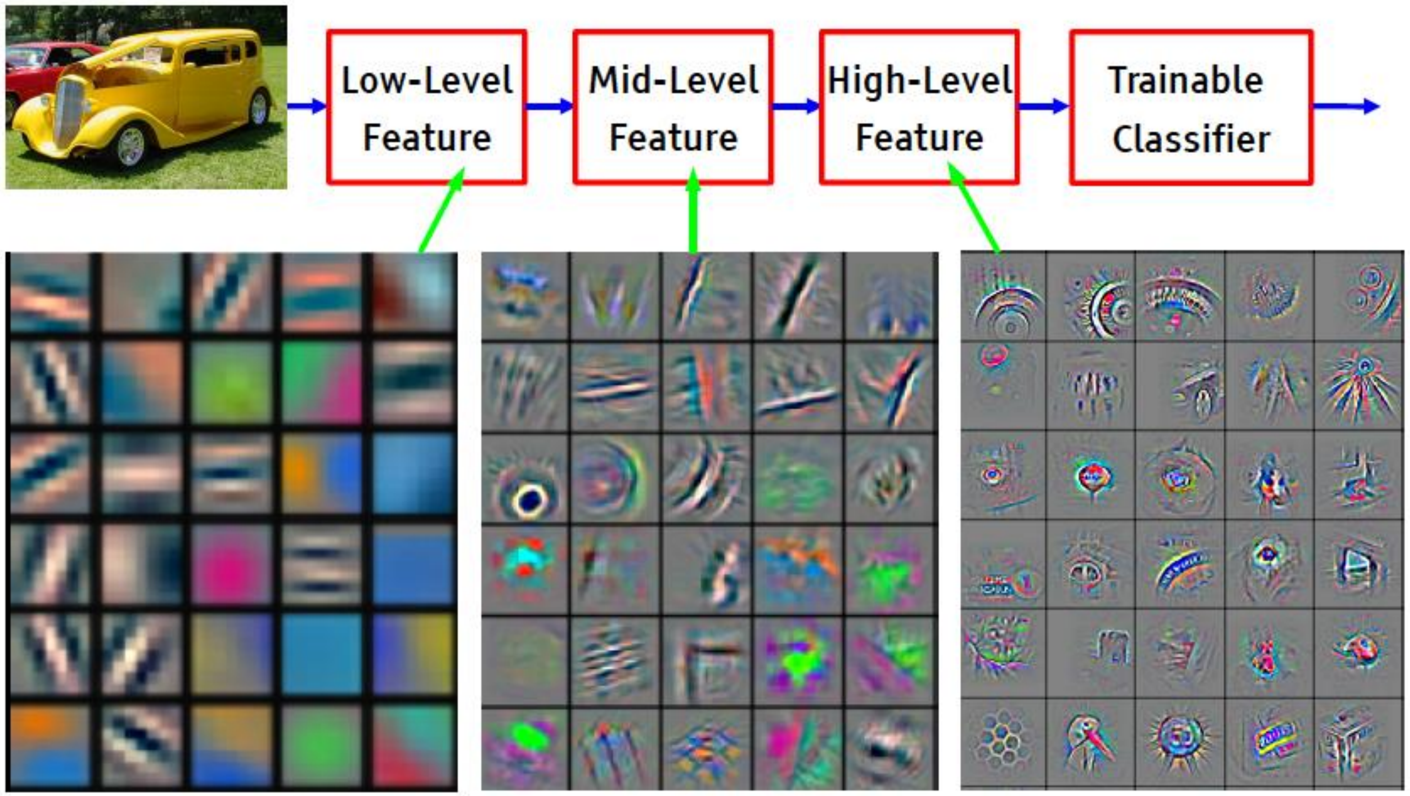
上海科技大学
ShanghaiTech University

- Inspiration from visual cortex



Why deep networks?

- A deep architecture can represent certain functions (exponentially) more compactly
- Learning a rich representation of input data



Feature visualization of convolutional net trained on ImageNet from [Zeiler & Fergus 2013]