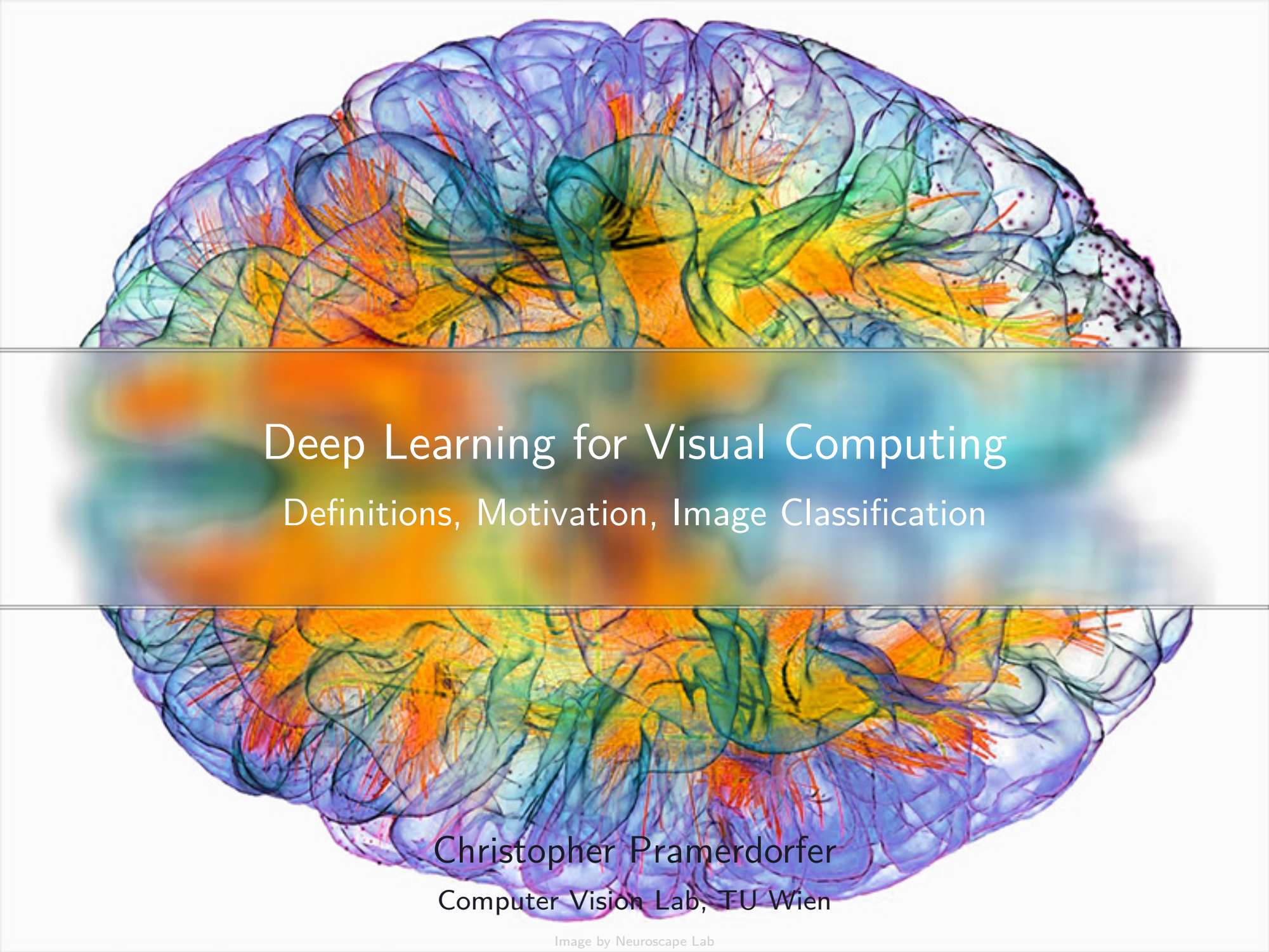


Deep Learning for Visual Computing

Definitions, Motivation, Image Classification

Christopher Pramerdorfer

Computer Vision Lab, TU Wien

Image by Neuroscape Lab

Topics

Definitions

- Deep Learning
- Convolutional Neural Networks

Motivation: Why care about Deep Learning?

Image classification

- Definition
- Challenges
- Datasets

Learn to solve problems in hierarchical fashion [1]

- Learn hierarchy of concepts
- Later concepts build upon earlier (simpler) ones

Graph of concepts has many layers

- Hence Deep Learning (DL)

Definitions
Convolutional Neural Networks (CNNs)

DL models for data with grid-like structure (e.g. images)

- Deep feedforward neural networks
- Include layers that perform convolutions

Most important models for image analysis

- Focus of this course

Definitions
Convolutional Neural Networks (CNNs)

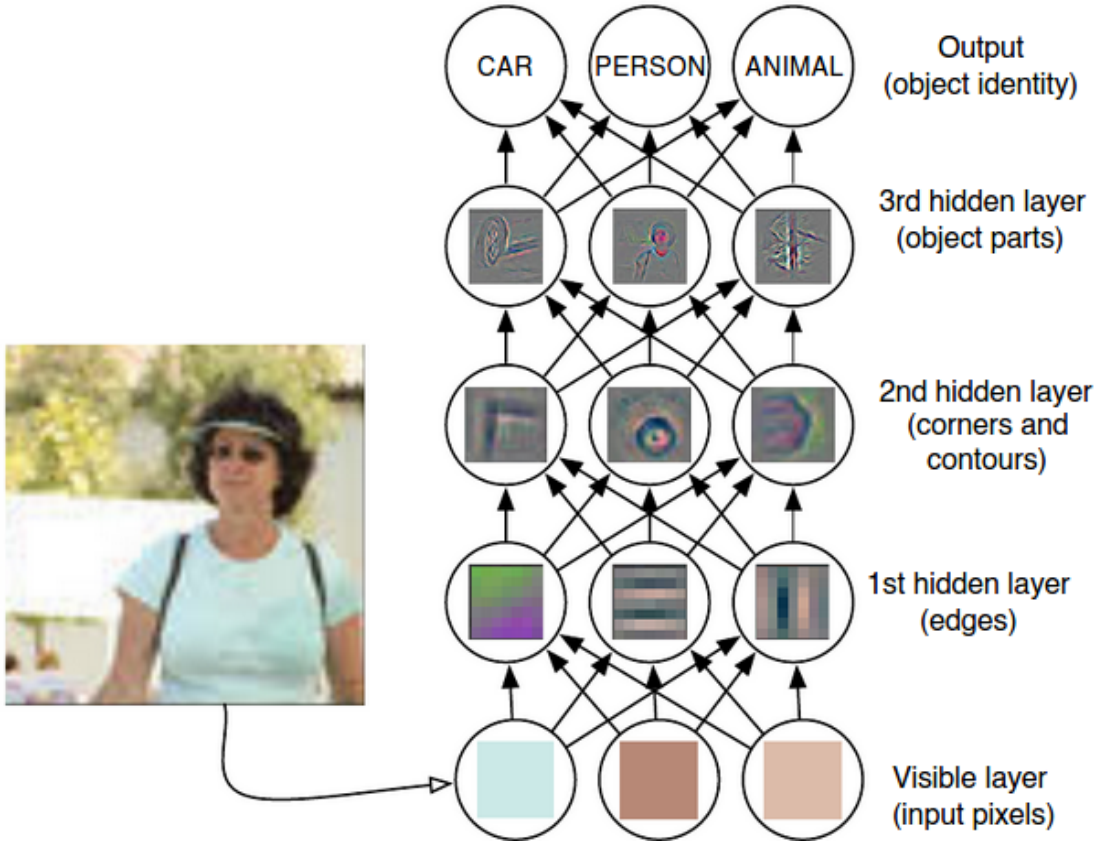


Image from [1]

blank page

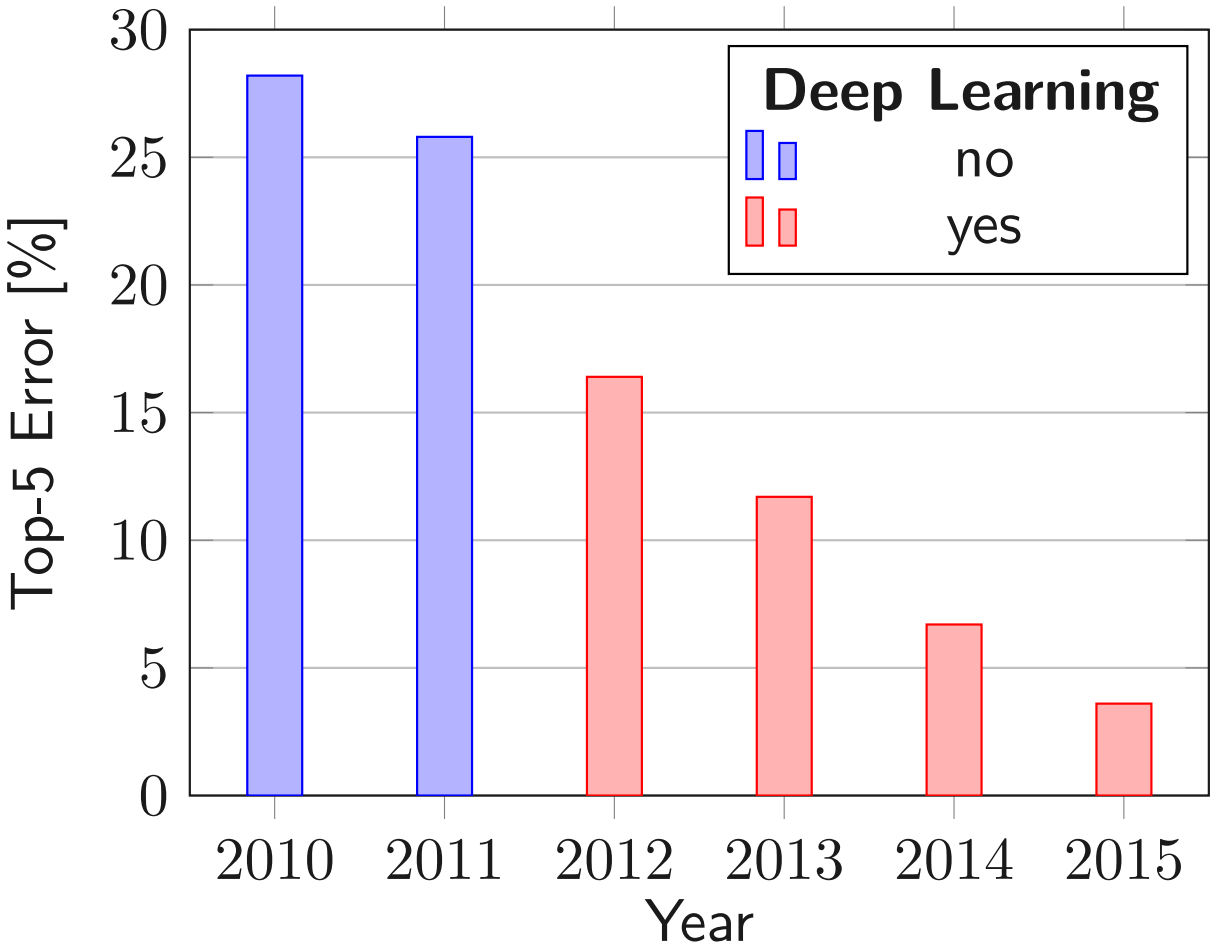
Motivation

Why care about Deep Learning?

- Significantly better performance on many tasks
- Flexible (can do alot more than classification)
- Companies want people with experience

Let's see some examples of what Deep Learning can do

Motivation
Image Classification (LSVRC)



Motivation
Image Classification

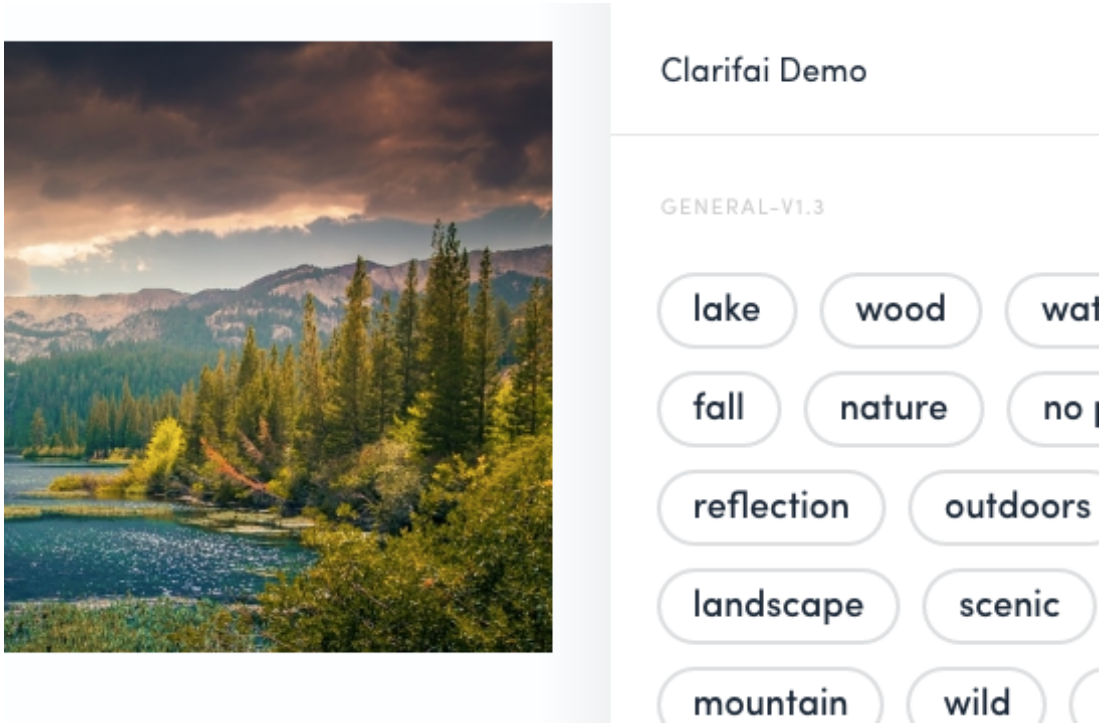


Image from clarifai.com

Motivation
Object Detection

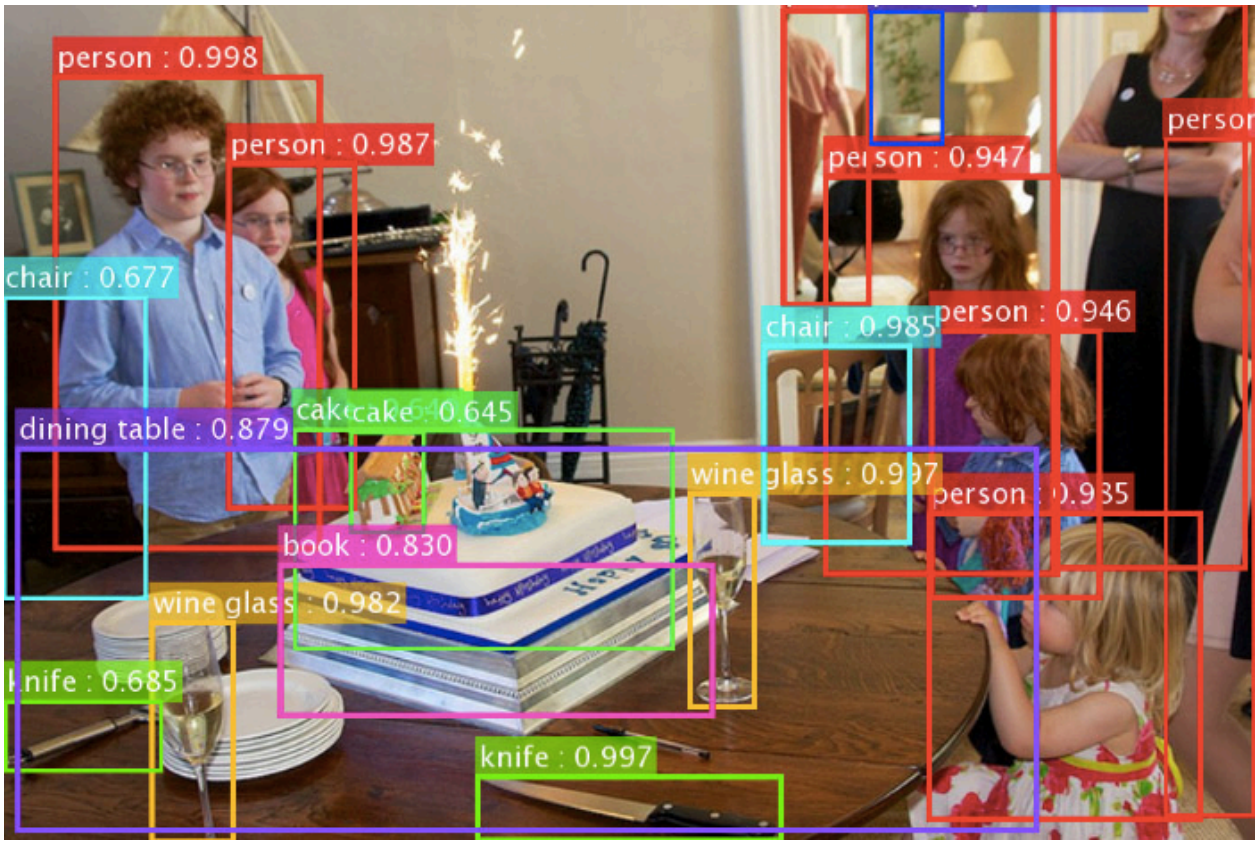


Image from kaiminghe.com

Motivation
Object Detection

Image from youtube

Motivation
Image Colorization

Image from richzhang.github.io

Motivation
Style Transfer

Image from [2]

Motivation
Facial Landmark Detection

Image from [3]

Motivation
3D Models from Single Images

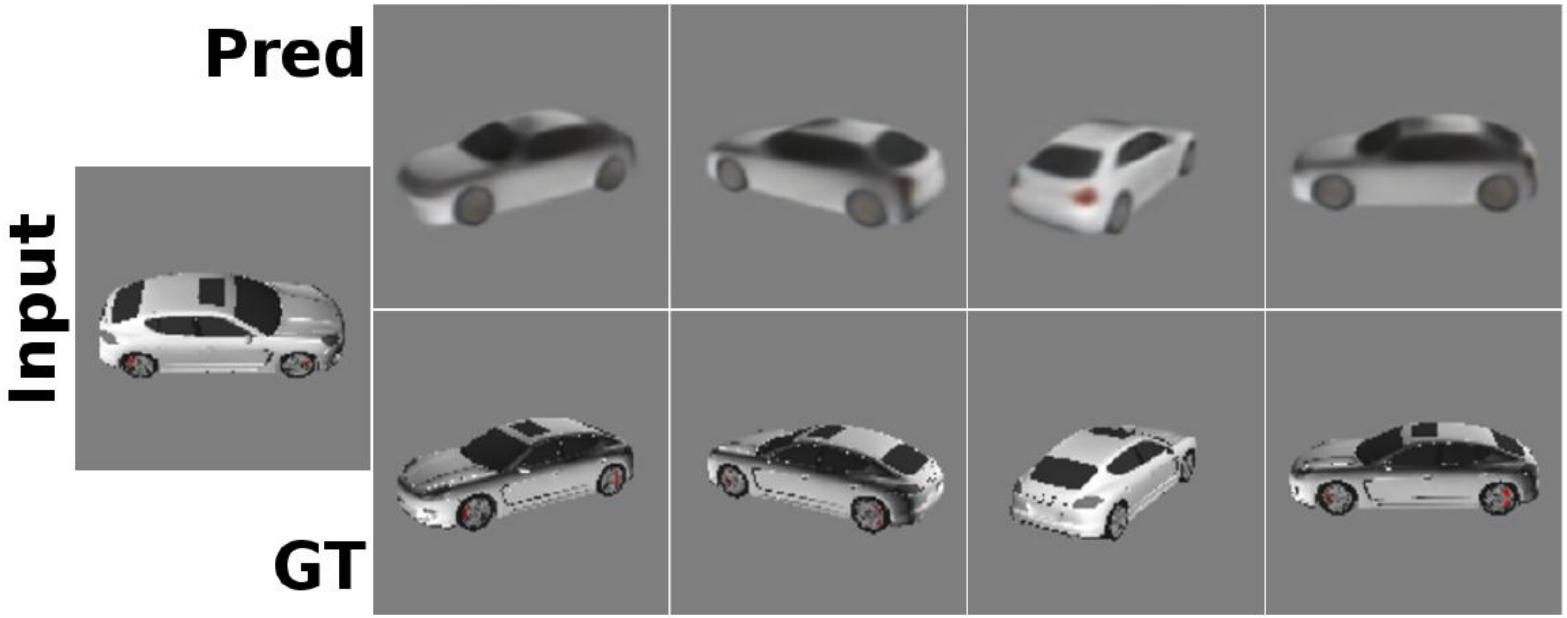


Image from [4]

Motivation
Scene Understanding

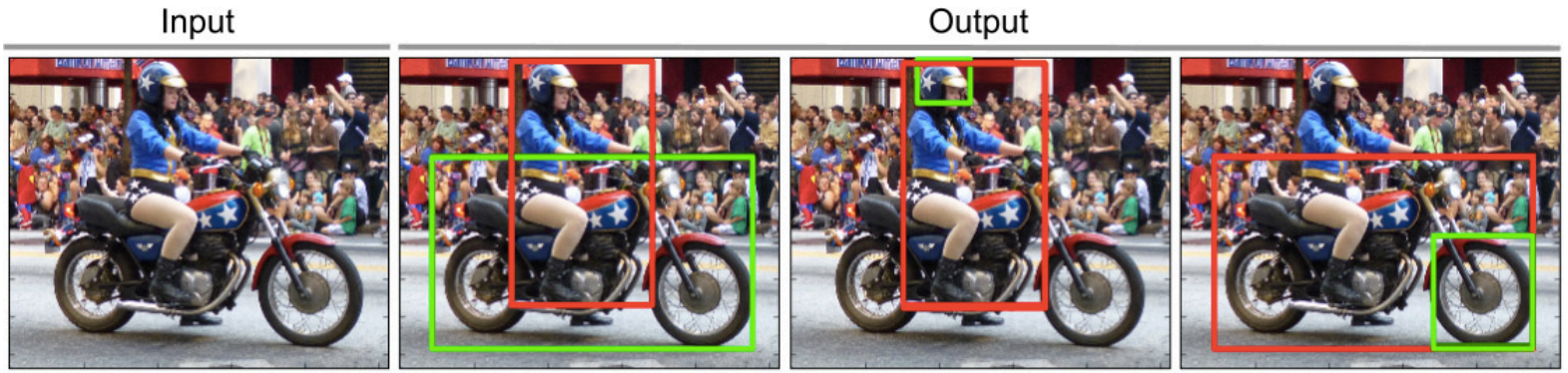
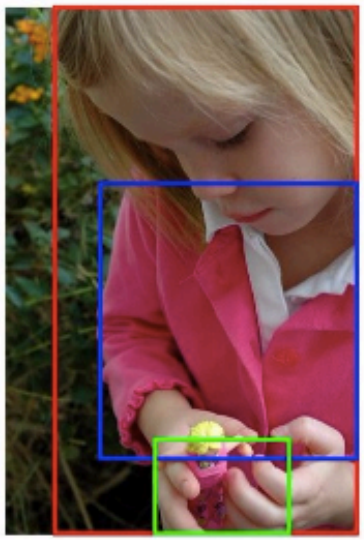


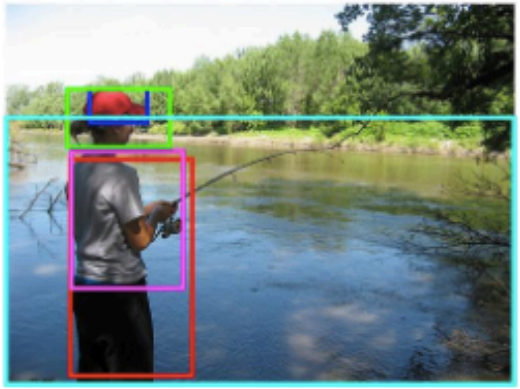
Image from stanford.edu

Motivation
Scene Understanding



A little girl in a pink shirt is looking at a toy doll.

A woman is riding a bicycle on the pavement.



A girl with a red cap, hair tied up and a gray shirt is fishing in a calm lake.

Image from [5]

blank page

Image Classification

One of main image analysis tasks

Task definition

- Given a set of class labels (e.g. {bird, cat, dog})
- Which class does the given image belong to?

$\Rightarrow$ cat

Image from youtube.com

Image Classification

Image belongs to exactly one class in the set

- Comparatively easy task
- On some datasets, machines now outperform humans!

But still very challenging

Image Classification
Challenges – Pose and Viewpoint

Image adapted from warrenphotographic.co.uk

Image Classification
Challenges – Illumination

Image from studioddt.com

Image Classification
Challenges – Deformation

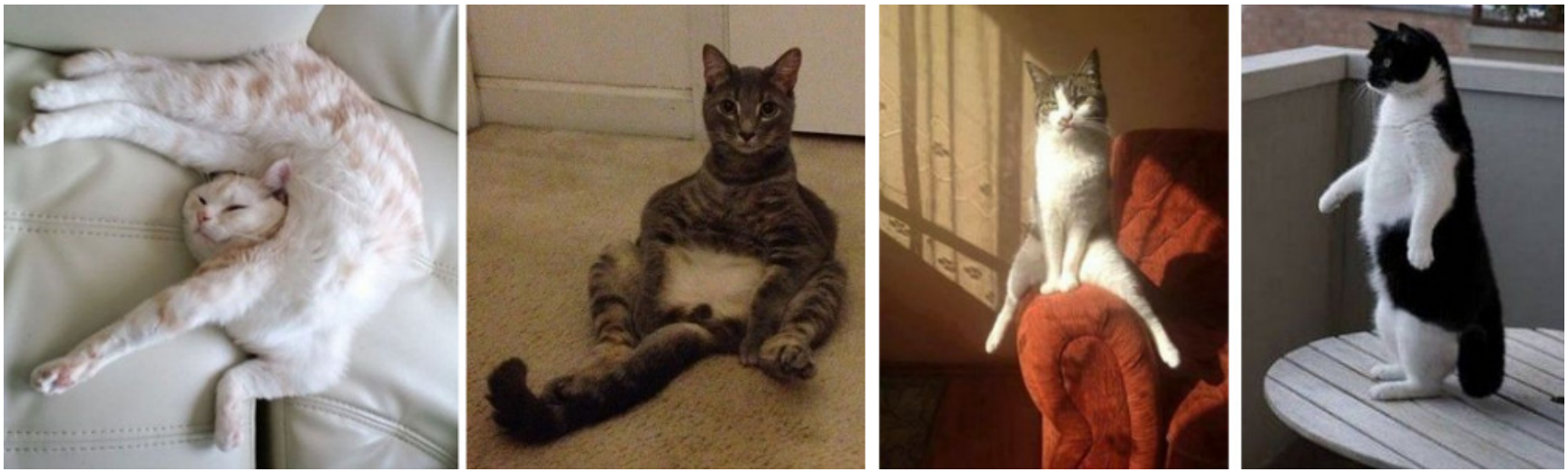


Image from cs231n.github.io

Image Classification
Challenges – Occlusion

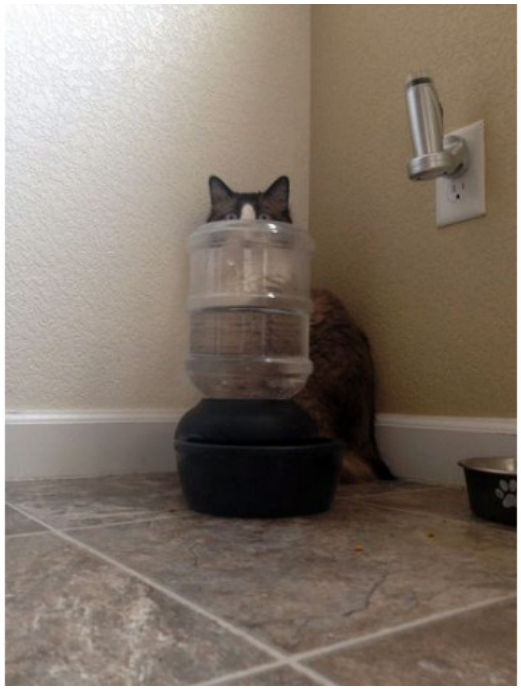




Image from cs231n.github.io

Image Classification
Challenges – Background

Image from cs231n.github.io

Image Classification
Challenges – Intraclass Variation

Image from cs231n.github.io

Image Classification
Datasets

Computer vision research is dataset-driven

- Data required for developing and testing
- Collecting and annotating takes lots of effort

Public image classification datasets available

- Frees us from having to collect data
- Facilitates method comparison

Image Classification
Datasets – CIFAR10

10 classes, 60k images

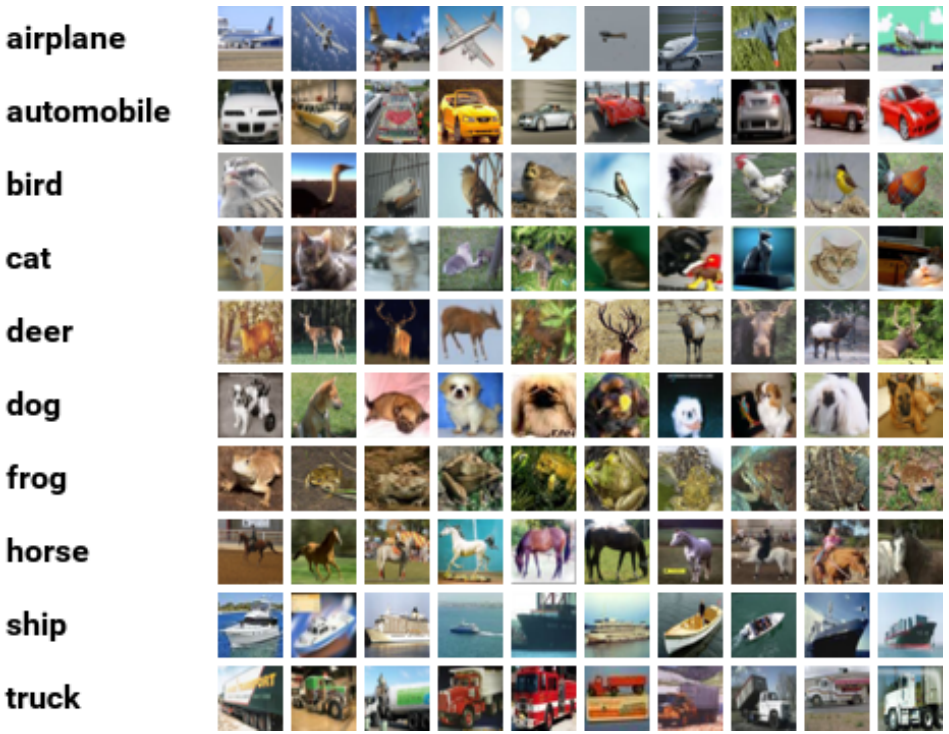


Image from cs.toronto.edu

Image Classification
Datasets – Pascal VOC

20 classes, 29k images

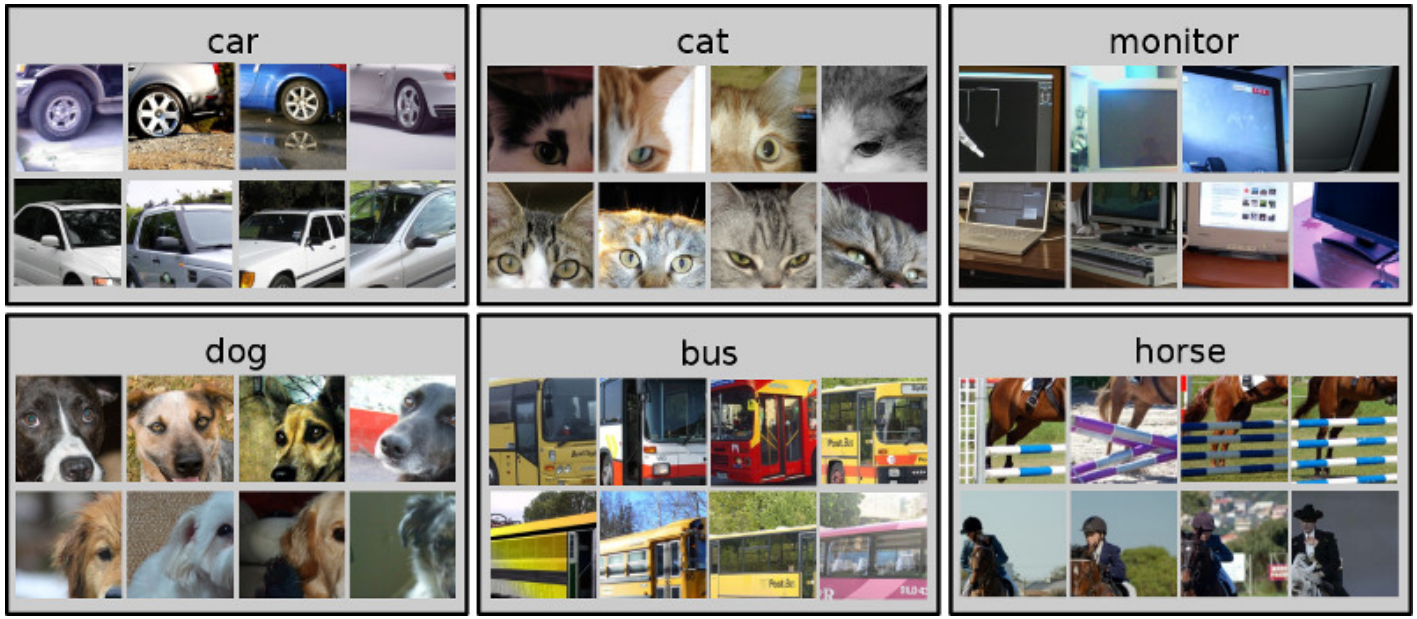
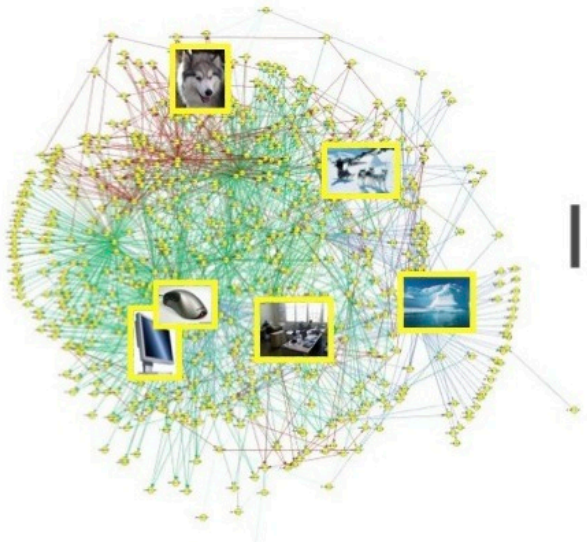


Image from cs.adelaide.edu.au

Image Classification
Datasets – ImageNet (LSVRC)

1000 classes, 1.4m images

- Subset for annual image classification challenge



IMAGENET

Image from umich.edu

Image Classification
Datasets

We always require three disjoint subsets

- Training set: for training (duh)
- Validation set: for tuning hyperparameters
- Test set: for a final performance analysis

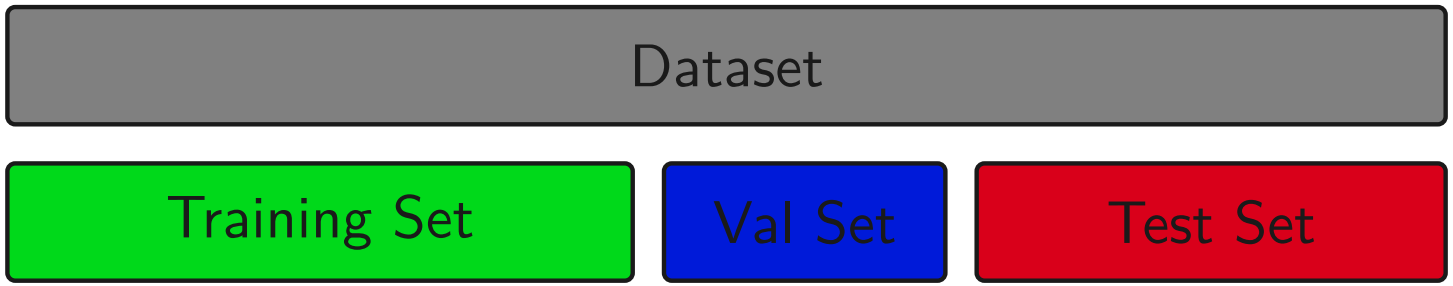


Image Classification

We know the problem and have data

How can we “solve” the image classification problem?

- Next lecture

Bibliography I

[1] *Deep learning*, 2016, [Online]. Available: http://www.deeplearningbook.org.

[2] *Image Style Transfer Using Convolutional Neural Networks*, CVPR, 2016.

[3] *A Recurrent Encoder-Decoder Network for Sequential Face Alignment*, ECCV, 2016.

[4] *Multi-View 3D Models from Single Images with a Convolutional Network*, ECCV, 2016.

[5] *Grounding of Textual Phrases in Images by Reconstruction*, ECCV, 2016.