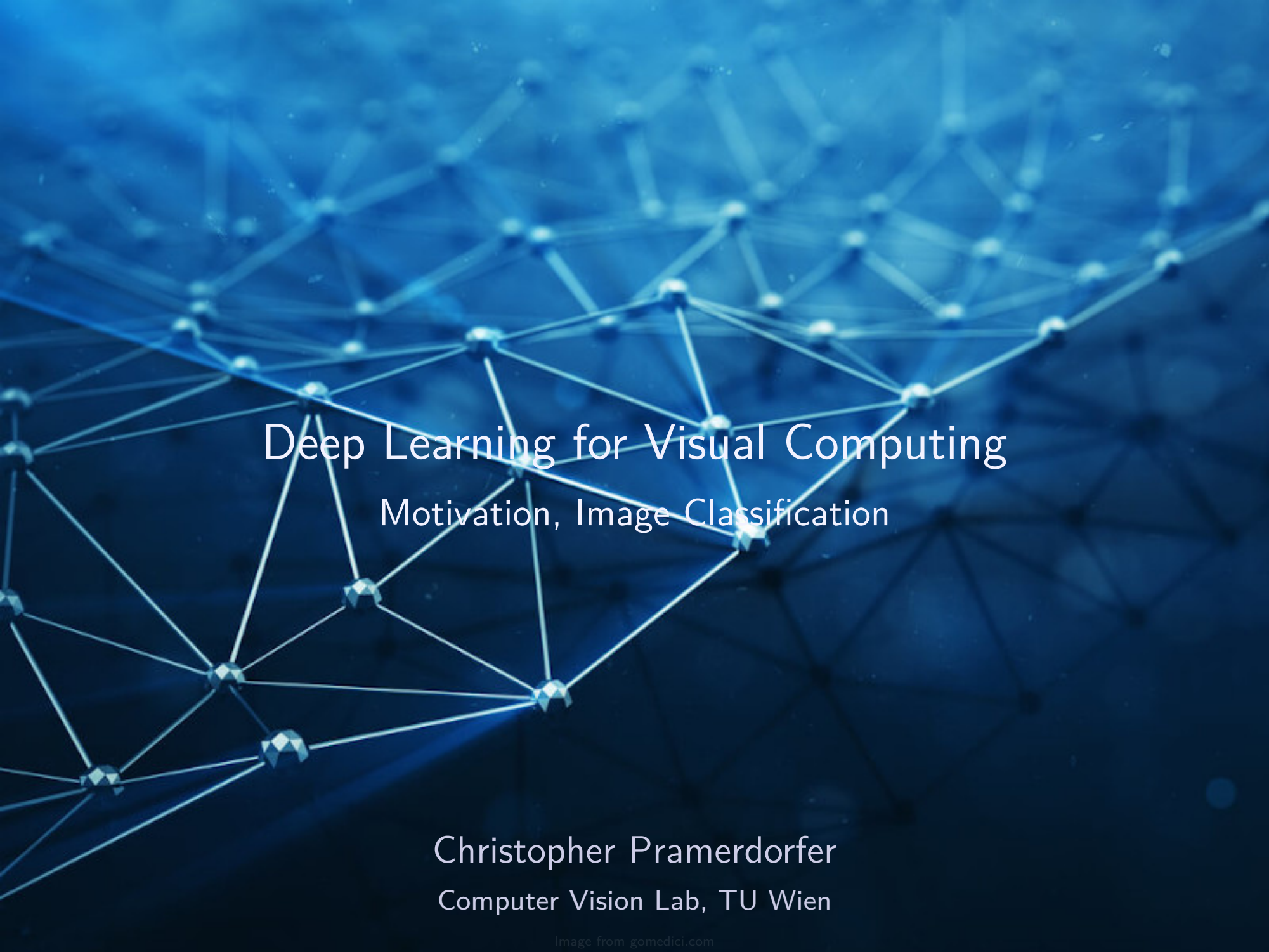


Deep Learning for Visual Computing

Motivation, Image Classification

Christopher Pramerdorfer

Computer Vision Lab, TU Wien

Image from gomedici.com

Topics

Deep Learning

- Motivation
- Introduction

Image classification

- Challenges
- Datasets
- Manual approach

Motivation for Deep Learning

Course is called Deep Learning for Visual Computing

- Very generic term (includes computer graphics etc.)

We'll focus on Computer Vision

- Make computers gain high-level understanding of images
- Goal is human-like understanding

Motivation for Deep Learning

Image Classification is a common example

- "What thing is shown in this image?"

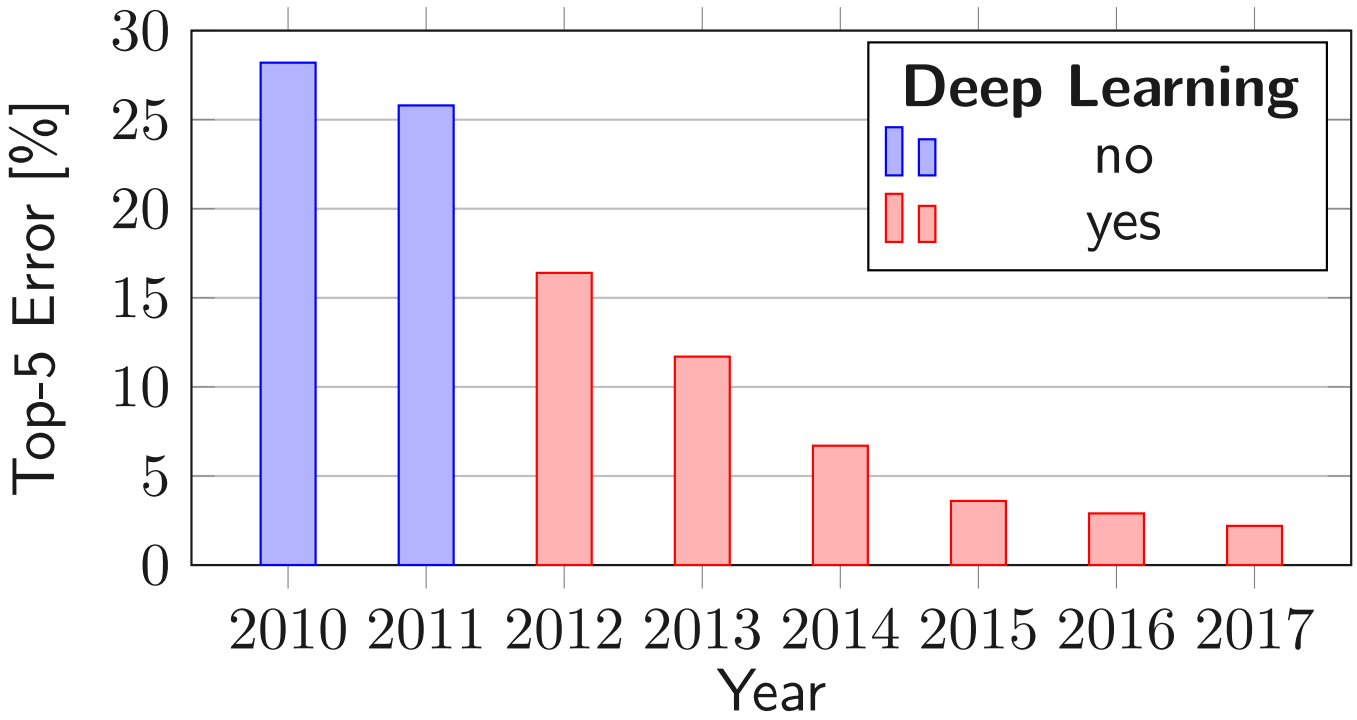
Image from youtube.com

$\Rightarrow$ cat

Motivation for Deep Learning

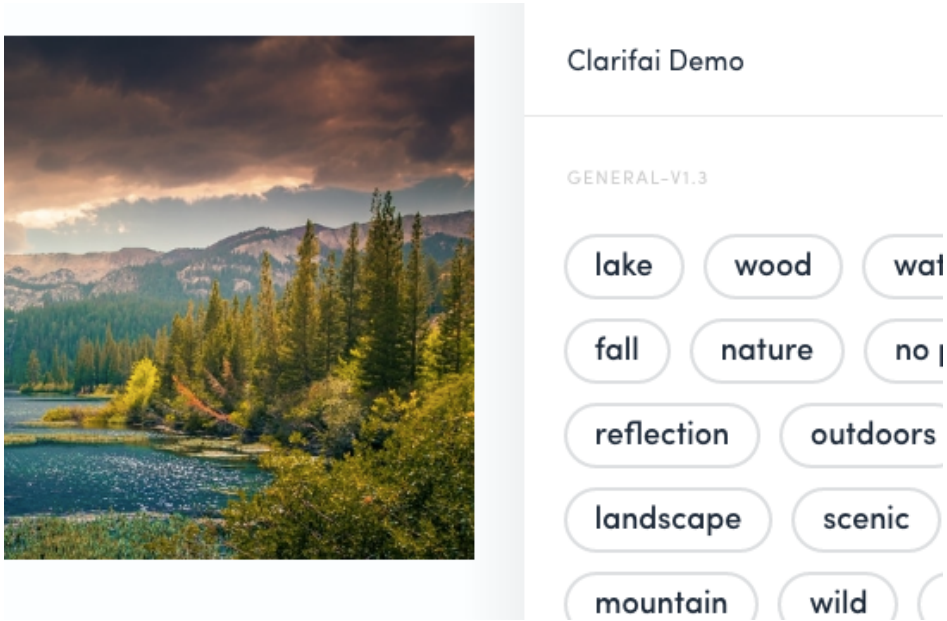
ImageNet benchmark performance over time

- 100k test images of 1000 classes



Motivation for Deep Learning
Other Forms of Image Understanding

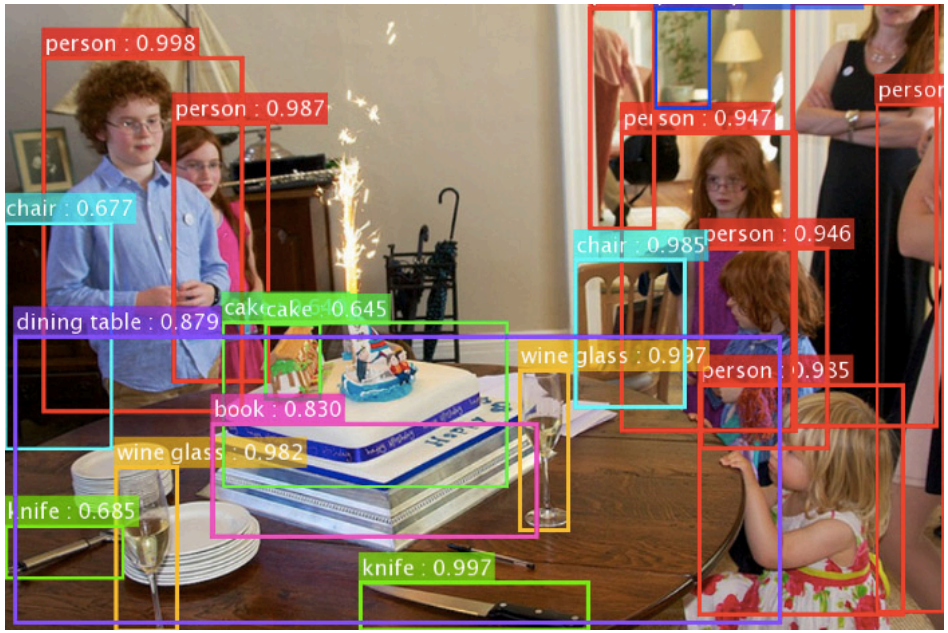
"Describe the image with keywords"



link

Motivation for Deep Learning
Other Forms of Image Understanding

"Detect notable objects"



Motivation for Deep Learning
Other Forms of Image Understanding

"Detect traffic participants"

link

Motivation for Deep Learning
Other Forms of Image Understanding

"Be an artist"

link

Motivation for Deep Learning
Other Forms of Image Understanding

"Estimate people's poses"

link

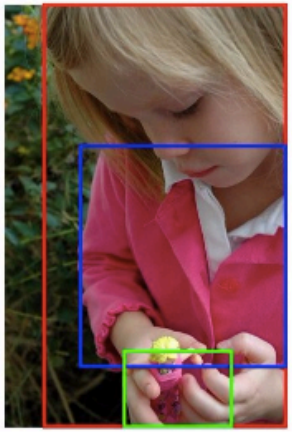
Motivation for Deep Learning
Other Forms of Image Understanding

"Generate videos that look real"

link

Motivation for Deep Learning
Other Forms of Image Understanding

"Describe the image with a sentence"



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.

paper

Motivation for Deep Learning

All these examples are based on Deep Learning

- Would be impossible otherwise at this quality
- We will take a closer look throughout the lecture

Deep Learning is state of the art

- In virtually any Computer Vision task
- In other fields as well (e.g. speech recognition)

Deep Learning Introduction

Sounds like magic? Not quite!

- Implemented using Neural Networks
- With neurons adapted for image data (convolutional)
- And arranged in many layers (hence “deep”)

Key ingredients

- Large datasets
- Lots of processing power (GPUs)

Deep Learning Introduction

Networks learn high-level concepts for lower-level ones

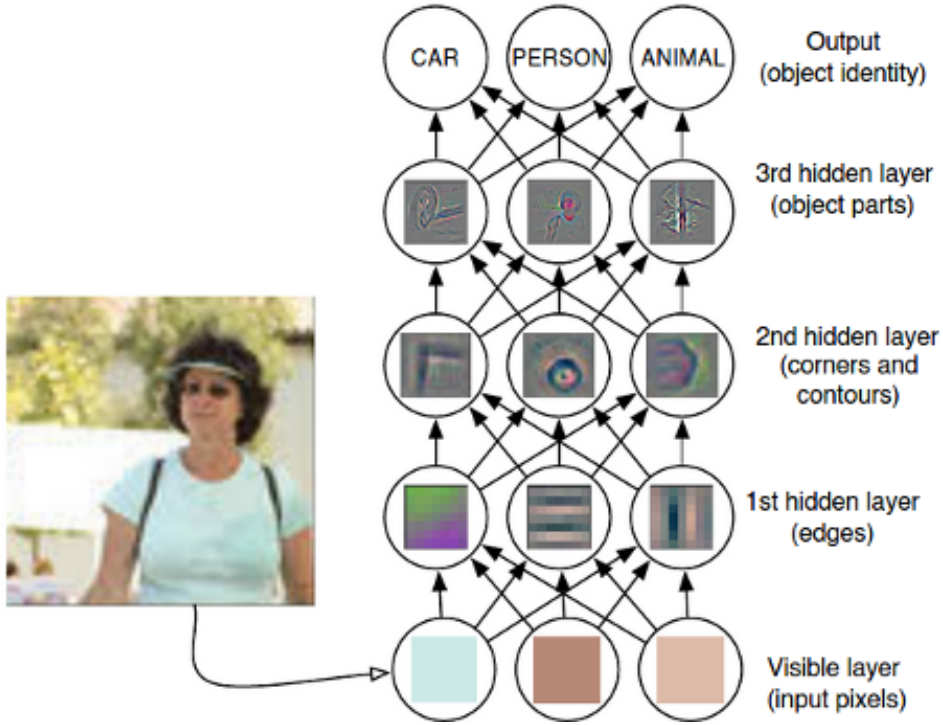


Image from deeplearningbook.org

Deep Learning Introduction

The approach is $30+$ years old

- But data and processing power were limited until recently

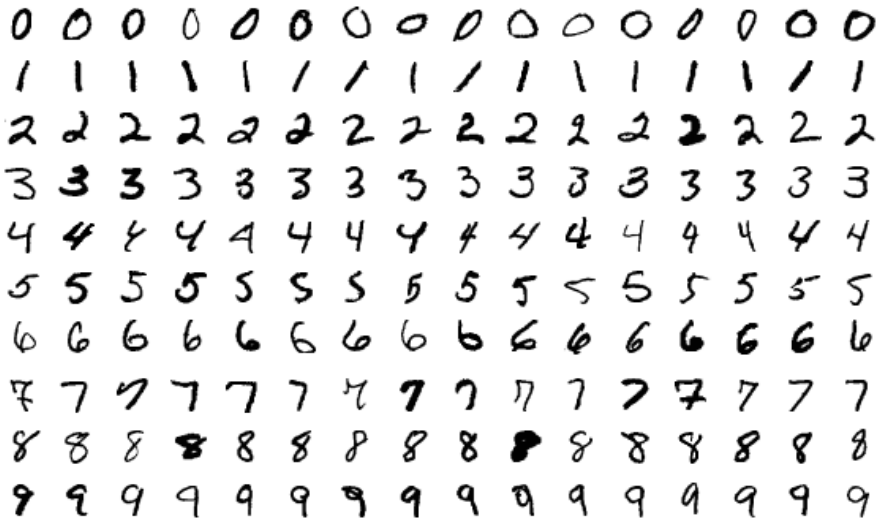


Image from Wikipedia

blank page

Image Classification

Fundamental Computer Vision 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 (but still challenging)
- On some datasets Deep Learning outperforms humans

Simple problem formulation

- Great for learning Deep Learning basics
- Why we will stick to classification for now

Image Classification
Challenges – Pose and Viewpoint

Image adapted from warrenphotographic.co.uk

Image Classification
Challenges – Illumination

Image from studioddt.com

Image Classification
Challenges – Deformation

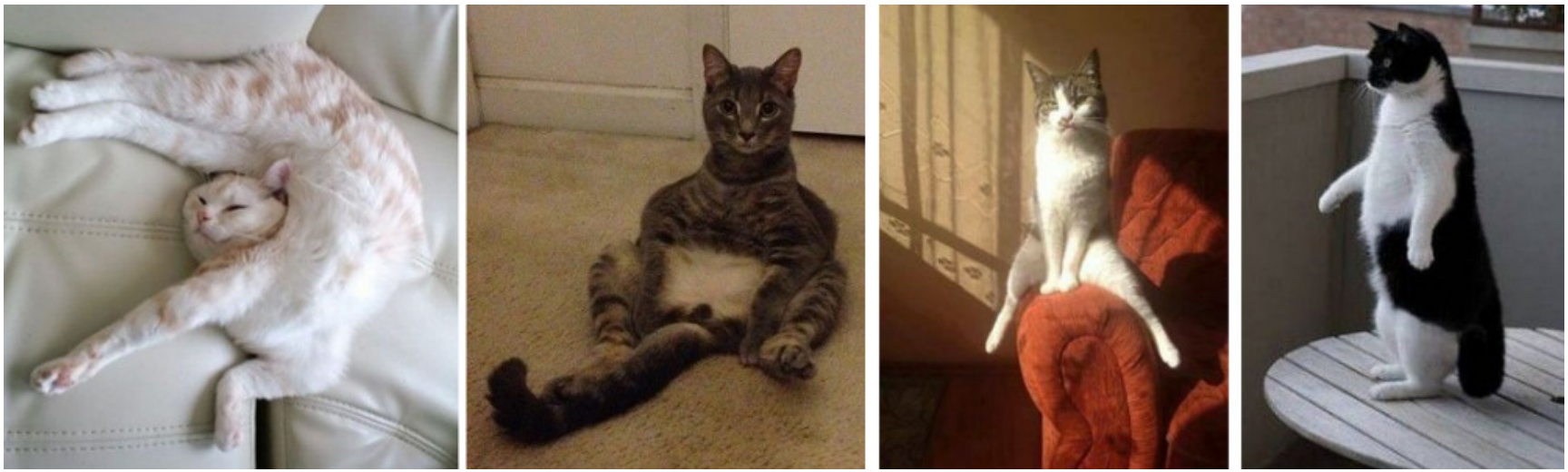


Image from cs231n.github.io

Image Classification
Challenges – Occlusion

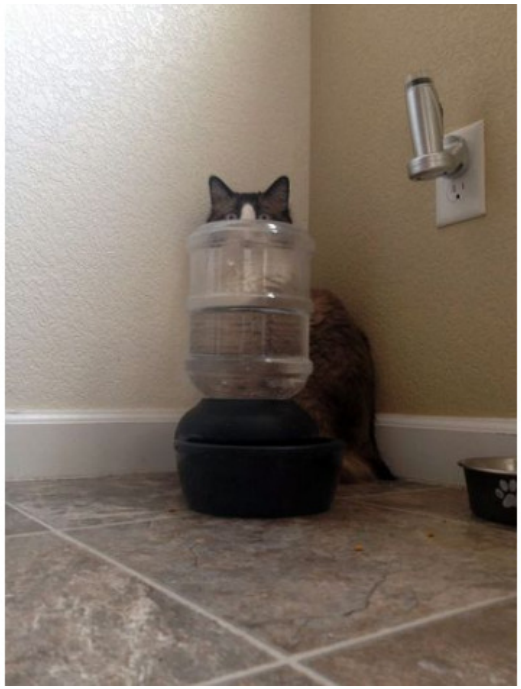




Image from cs231n.github.io

Image Classification
Challenges – Background

Image from cs231n.github.io

Image Classification
Challenges – Intraclass Variation

Image from cs231n.github.io

Image Classification
Datasets

A good classifier must cope with these challenges

- To verify this we need a representative dataset
- Such datasets are usually large

If we employ Machine Learning we also need training data

- Datasets must be disjoint (so need even more data)
- Deep Learning requires lots of data

Image Classification
Datasets

Dataset acquisition takes lots of effort

- Collect many (thousands or more) of images
- Assign class labels to enable automatic training and testing

Data acquisition and processing is central in Deep Learning

- Often the most time-consuming task
- Usually main bottleneck for performance

Thankfully many public datasets are available

Image Classification
Datasets – CIFAR-10

10 classes, 60k images

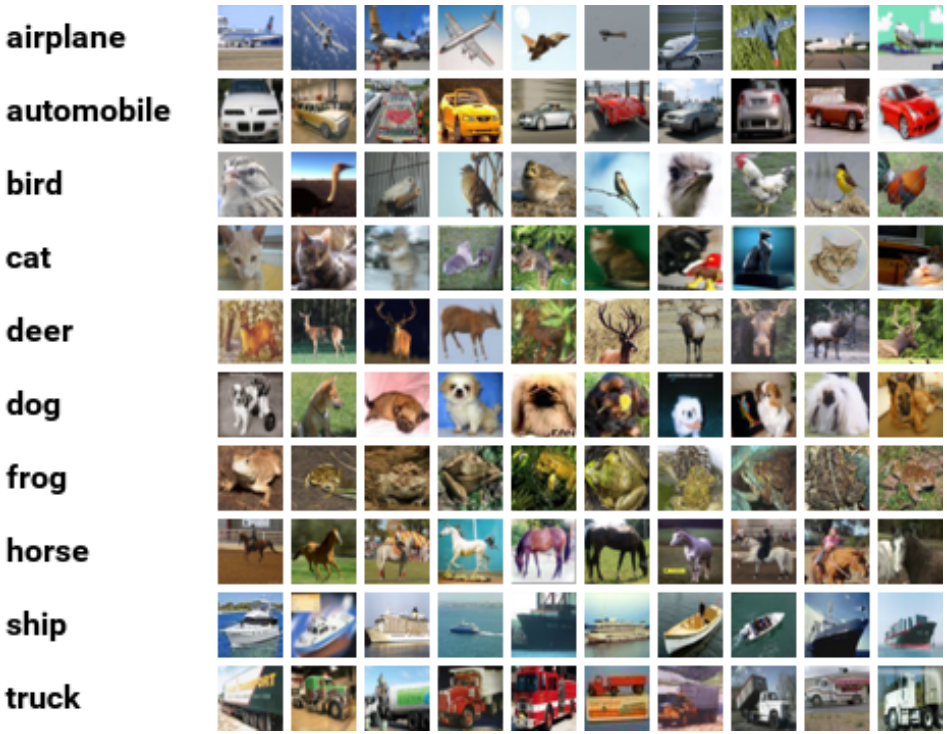
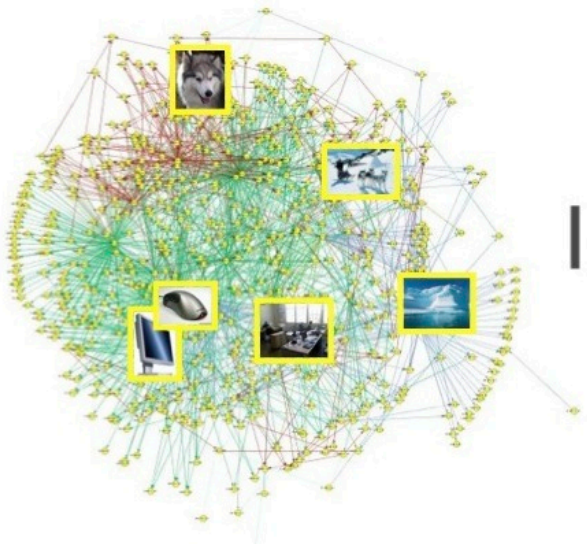


Image from cs.toronto.edu

Image Classification
Datasets – ImageNet

20k classes, 14m images

- Main driver for Deep Learning performance



IMAGENET

Image from umich.edu

Image Classification
Datasets – COCO

300k images, labels for classification, detection, segmentation, ...

Image from cocodataset.org

Image Classification

Let's build an image classifier

- Should support the classes {dog, cat}
- Using the CIFAR-10 dataset

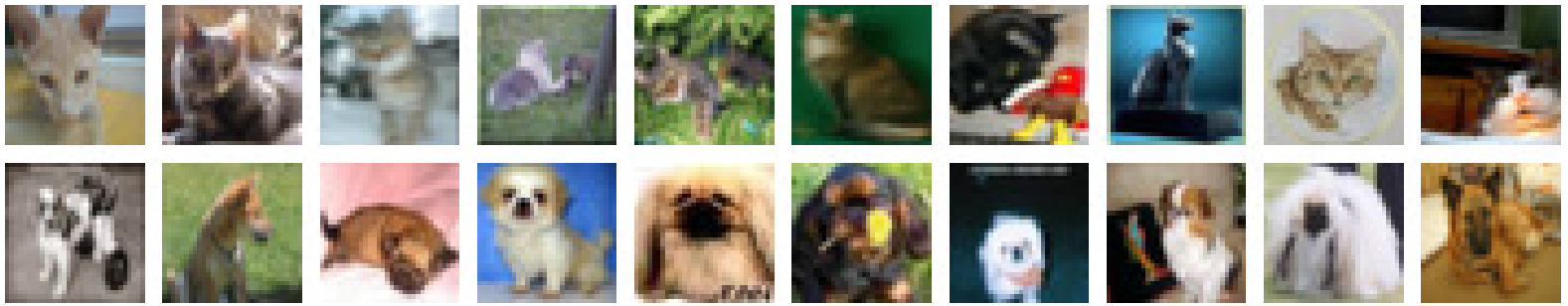


Image from cs.toronto.edu

Image Classification

How can we write an algorithm for this purpose?

Image from cs.toronto.edu

Image Classification

We cannot!

- No obvious unique and reliable features
- Not clear how to represent and use them

Image from cs.toronto.edu

Image Classification

We humans are incredible image classifiers

But we cannot describe formally how we do so

- Thus the standard `if {} else {}` approach fails

This applies to most vision problems

- Reason we need Machine and Deep Learning