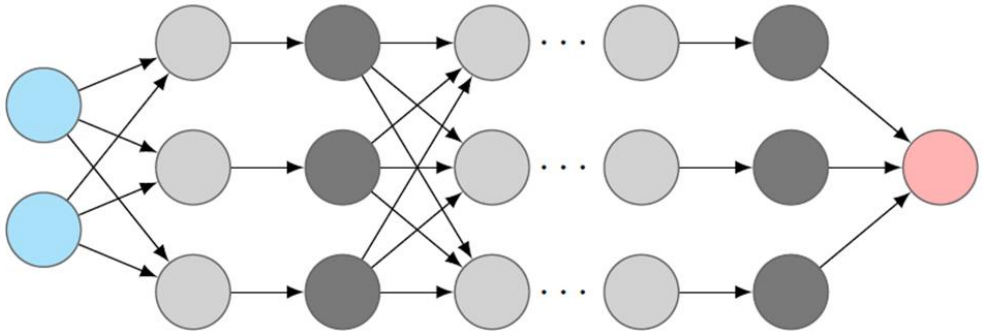
Deep Learning

Introductory course for Master students in computer science and mechatronics

Lecturer: Michael Möller – michael.moeller@uni-siegen.de
Exercises: Hartmut Bauermeister – hartmut.bauermeister@uni-siegen.de



Organizational Things

Necessary prior knowledge

- Linear Algebra
- Calculus (ideally with multiple variables, but we'll repeat this)
- Programming (we will introduce Python, NumPy, and PyTorch, but will not be able to repeat loops, conditions, data-types, classes…)

Nice to know but not necessary

- Image processing
- Optimization
- Basic machine learning, e.g. statistical learning theory, pattern recognition

Organizational Things

Lecture

- Starts at 12:15
- Will be in **room H-D 3206** most of the time, BUT

  On Oct. 17th the lecture will be in room H-C 3310! This will be crowded.

  On Dec. 5th the lecture will be in room H-B 4419/20!
- I'll try to organize a Christmas lecture with speakers from industry, which may lead to another change in the room.

Organizational Things

Exercises

- Will start next Monday 15th of October
- Will be in **room H-B 4419/20**
- Hartmut Bauermeister, hartmut.bauermeister@uni-siegen.de, will lead the exercises.
- We will hand out homeworks (in the form of Jupyter Notebooks), on which you may work in teams of 2 students
- The solution to the homework will be discussed in the exercises
- The exercises are also a good plattform to ask detailed questions about the course material
- In December, the exercises might turn into a challenge (tbd)

Organizational Things

Exercises

- The computations for the first ~5 exercises can carried out on a CPU
- After this, a CPU computation is possible, but a little lengthy. If you like you can get access to room H-A 7118 (outside of lectures).
  It has 12 Computers with NVIDIA graphics cards.
- Hartmut organizes this access to the computer room.
- Please send you final solution to Hartmut (hartmut.bauermeister@uni-siegen.de) in teams of 2 students.

Getting 50% of the assignments to work is a prerequisite for being admitted to the final exam!

Furthermore, programming questions are part of the exam!

Organizational Things

- My office: H-A 7106
- Hartmut's office: H-A 7116
- For appointments, please email us or contact us during the lecture/exercises
- The lecture and exercise start at quarter past.
- Course website: http://www.vsa.informatik.uni-siegen.de/en/deep-learning
- Username: student Password: 100%brain

This lecture is worth 5 credits. The final exam will be written (90 min.)!

Please do not be shy to say something and ask questions during the lecture!

The more we discuss, the more interesting the lecture is!

Deep Learning

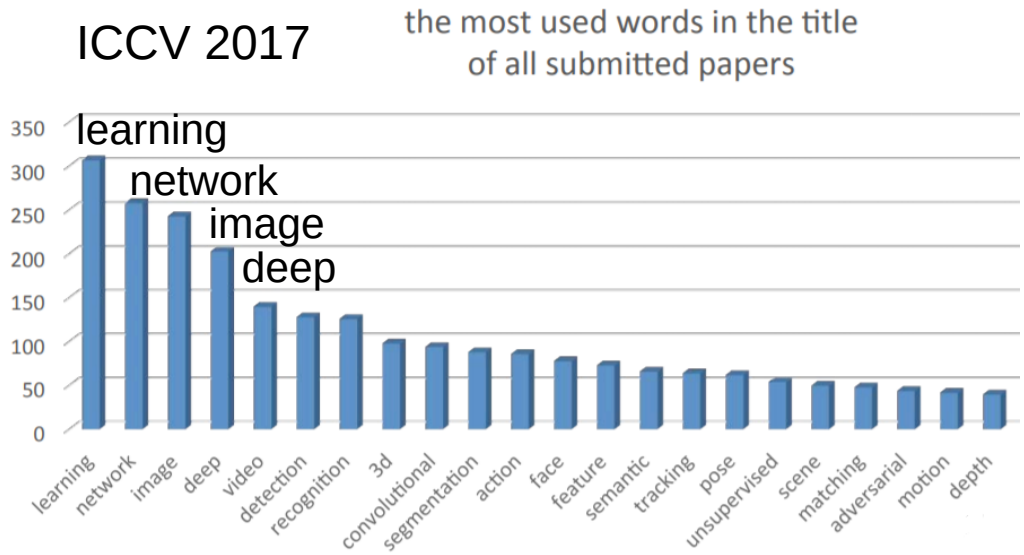
"So one way to think about all three of these ideas is that Machine Learning is the cutting edge of Artificial Intelligence. And Deep Learning is the cutting edge of the cutting edge."

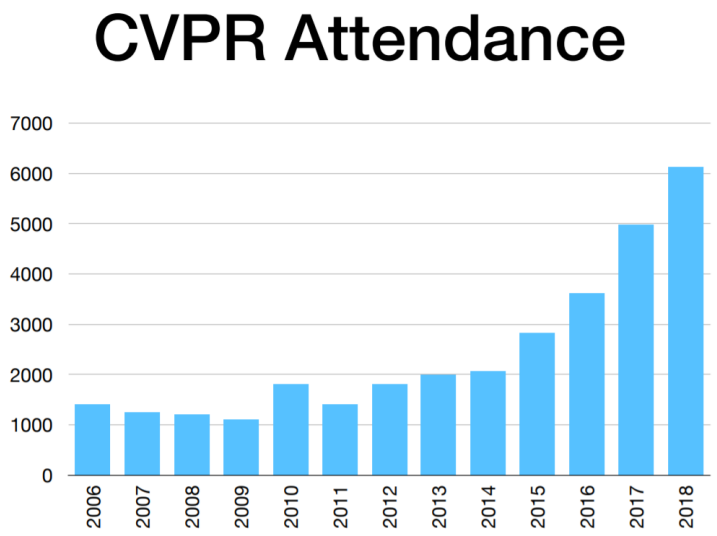
https://medium.com/machinevision/overview-of-artificial-intelligence-buzz-adb7a5487ac8, Oct. 4th

"Deep learning may be one of the most overhyped of modern technologies, but there is a good chance that it will one day become the secret sauce in many different business processes."

https://www.ft.com/content/0a879bec-48bd-11e8-8c77-ff51caedcde6, financial times, Oct. 4th

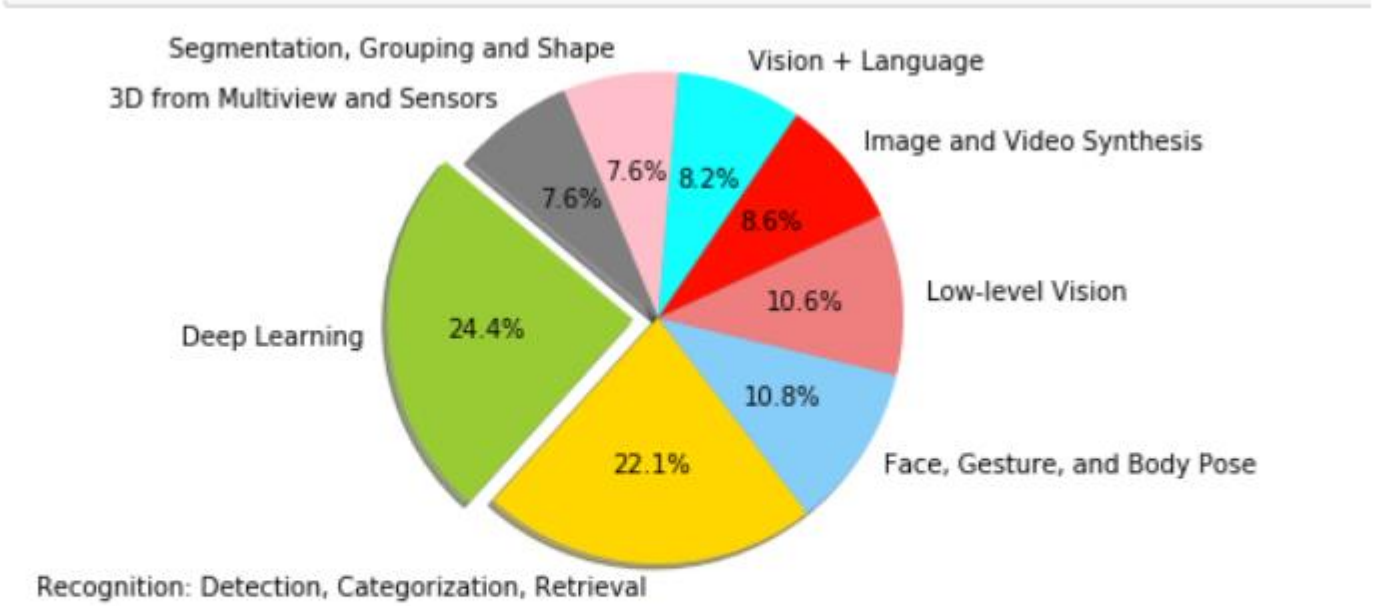
ICCV 2017





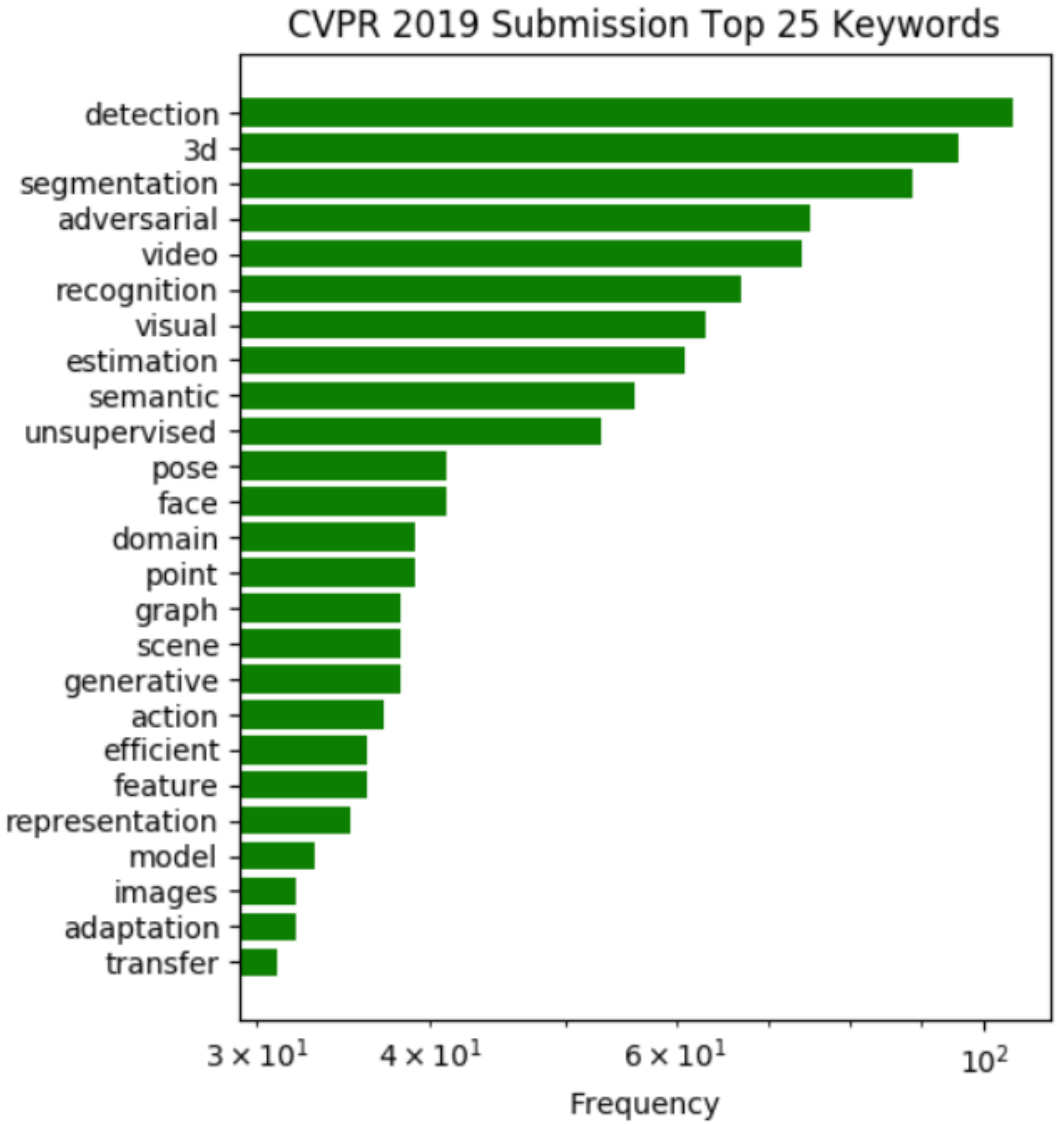
Deep Learning

About CVPR 2019 from https://towardsdatascience.com/latest-computer-vision-trends-from-cvpr-2019-c07806dd570b



Not surprisingly, most of the research is focused on Deep Learning (isn't everything deep learning now!)

Deep Learning



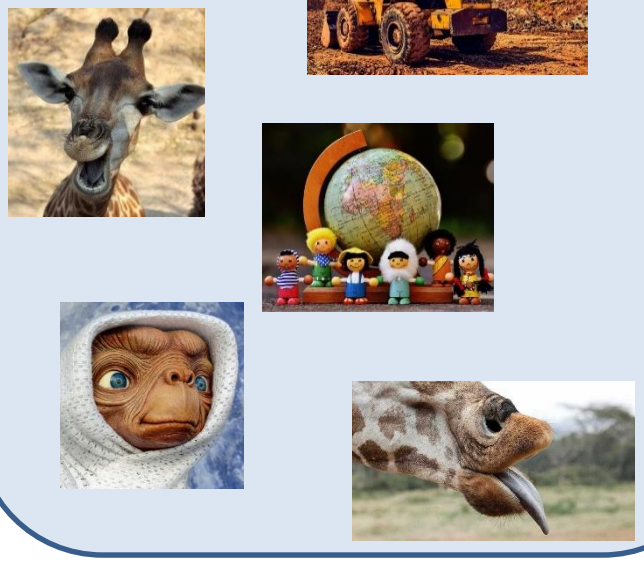
About CVPR 2019 from https://towardsdatascience.com/latest-computer-vision-trends-from-cvpr-2019-c07806dd570b

Deep Learning

My perspective on what (supervised) “Deep Learning” is:
A fancy word for function approximation

Assume there is an unknown function G that maps some kind of input data x to some kind of desired output y.

Space of all images



G

Answer to the question if the image shows a giraffe

NO!

YES!

Why are these problems difficult?

Yes, here it is!

Is there a dog in this image?

Unfortunately, this is how the dog looks like for the computer (red channel only)

178	189	190	187	182	174	120	139	172	175	141	124	149	168	183	176	184	182	184	192	192	195	185
170	174	181	162	144	178	174	183	178	180	183	181	180	186	183	181	185	189	190	190	170	172	208
186	187	185	185	188	187	186	185	181	182	170	176	188	189	189	185	187	194	192	182	182	180	193
182	188	185	178	164	150	159	170	180	184	173	153	128	114	119	153	182	137	145	195	188	191	191
186	137	75	45	33	20	25	32	50	63	41	22	2	0	0	18	46	29	38	154	186	191	175
119	26	2	2	2	0	0	0	0	0	0	0	0	3	9	2	0	0	3	17	52	91	120
19	9	7	5	5	3	6	8	8	4	3	2	2	3	6	7	4	6	4	0	0	0	1
3	0	4	6	7	4	6	9	8	1	4	5	6	4	5	6	6	8	11	10	10	6	0
118	60	1	4	6	7	6	6	6	5	8	6	9	10	10	7	3	8	27	29	23	11	5
204	98	2	7	8	13	13	12	12	14	11	4	4	6	8	9	6	8	13	30	38	18	16
113	104	26	9	11	11	13	13	13	15	12	7	6	5	7	7	9	21	44	68	50	23	13
38	102	56	14	16	13	26	31	20	22	11	7	9	8	8	10	42	101	179	165	113	29	4
52	23	40	29	14	4	78	160	82	46	25	6	7	7	9	40	121	166	126	92	127	85	26
37	41	51	25	16	9	77	204	180	134	107	60	14	16	36	114	182	161	131	120	168	160	134
76	174	144	31	12	9	87	187	194	198	198	186	59	36	39	122	197	189	204	207	200	203	212
156	201	202	98	4	7	124	197	191	191	194	206	79	92	65	67	185	194	202	212	200	145	125
209	201	206	136	6	4	143	196	182	193	195	186	69	169	156	39	137	219	202	201	190	137	137
140	113	93	93	22	44	174	179	188	199	206	186	79	172	206	97	70	159	122	106	169	190	194
107	134	155	148	31	116	203	199	194	192	210	173	77	135	160	175	60	80	141	130	167	190	194
132	157	195	134	20	156	197	172	151	176	181	94	96	168	163	177	114	49	172	161	132	139	137
137	172	212	105	25	164	188	171	153	88	74	53	117	164	143	130	147	101	93	134	155	143	111
183	186	177	106	31	157	196	188	95	63	139	144	152	166	158	175	175	188	123	69	67	118	132

Why are these problems difficult?

Image Source: Karen Zack, twitter.com/teenybiscuit

It is amazing how small visual differences can change our interpretation of images entirely!

Deep Learning

My perspective on what (supervised) “Deep Learning” is:
A fancy word for function approximation

Assume there is an unknown function G that maps some kind of input data x to some kind of desired output y.

Space of all images





G

Answer to the question if the image shows a giraffe

NO!

YES!

My perspective on what (supervised) “Deep Learning” is:
A fancy word for function approximation

Assume there is an unknown function G that maps some kind of input data x to some kind of desired output y.

Assume we are given some evaluations of this (unknown) function G. This is what we ill call ***training data***!

Giraffe

No giraffe

No giraffe

No giraffe

Giraffe

$$G\left(\quad\right) = 0, \quad G\left(\quad\right) = 1$$

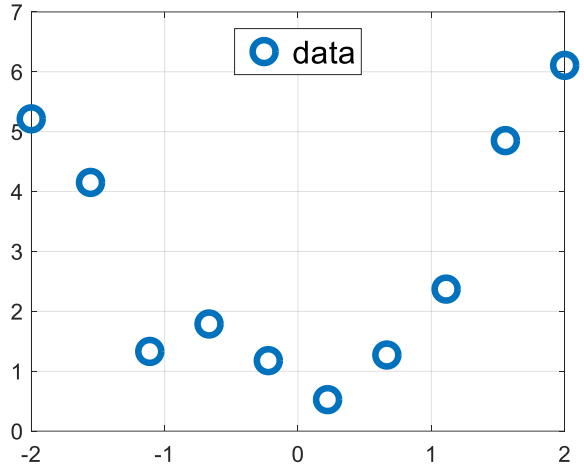
My perspective on what (supervised) "Deep Learning" is:
A fancy word for function approximation

Assume there is an unknown function G that maps some kind of input data x to some kind of desired output y.

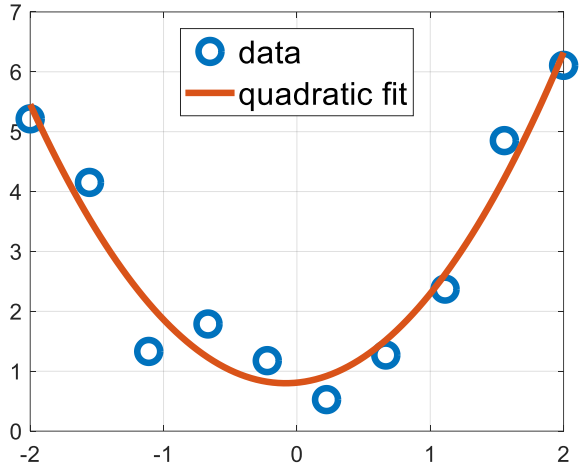
Assume we are given some evaluations of this (unknown) function G. This is what we will call ***training data***!

1. Choose a parameterized function $\mathcal{N}(x;\theta)$ in the hope that for the right choice of parameters θ it approximates the unknown function G well. We call $\mathcal{N}$ the ***network***, and sometimes refer to θ as the ***weights***.
2. Try to determine suitable weights θ in such a way that $\mathcal{N}(x_i;\theta) \approx y_i$ holds for all examples (x_i, y_i) from your training data set. This is referred to as ***training the network***.
3. Make try to ensure that both, the architecture as well as the training are chosen in such as way that the network makes good predictions during inference, i.e. on previously unseen data x: $\mathcal{N}(x;\theta) \approx G(x)$. We refer to this property as ***generalization***.

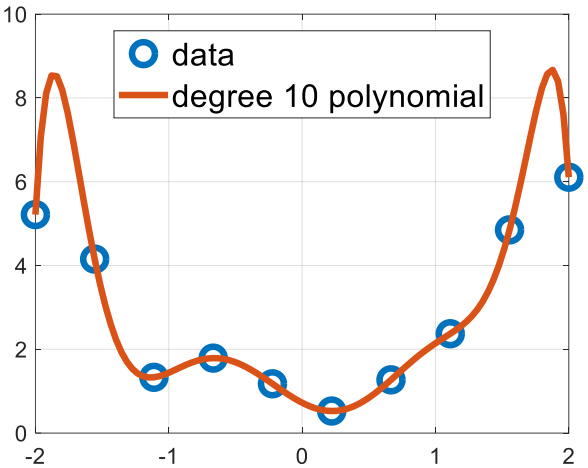
Shallow Example



Training data (x,y)



$$\mathcal{N}(x;\theta) = \theta_1 x^2 + \theta_2 x + \theta_3$$



$$\mathcal{N}(x;\theta) = \sum_{i=0}^{10} \theta_i x^i$$

Depending on the underlying function G, one or the other choice might be better!

This is a very simple 1d example! The power of deep learning, and the reason it receives a lot of attention are that similar concepts seem to work extremely well for incredibly complex functions G!!

Impressive Deep Examples

Predicting the sound objects make when you hit them:
https://www.youtube.com/watch?v=0FW99AQmMc8

Lip-synchronization from audio:
https://www.youtube.com/watch?v=9Yq67CjDqvw&t=268s

Lip-reading:

https://www.youtube.com/watch?v=5aogzAUPilE
https://www.youtube.com/watch?v=fa5QGremQf8&t=4s

Video reenactment:

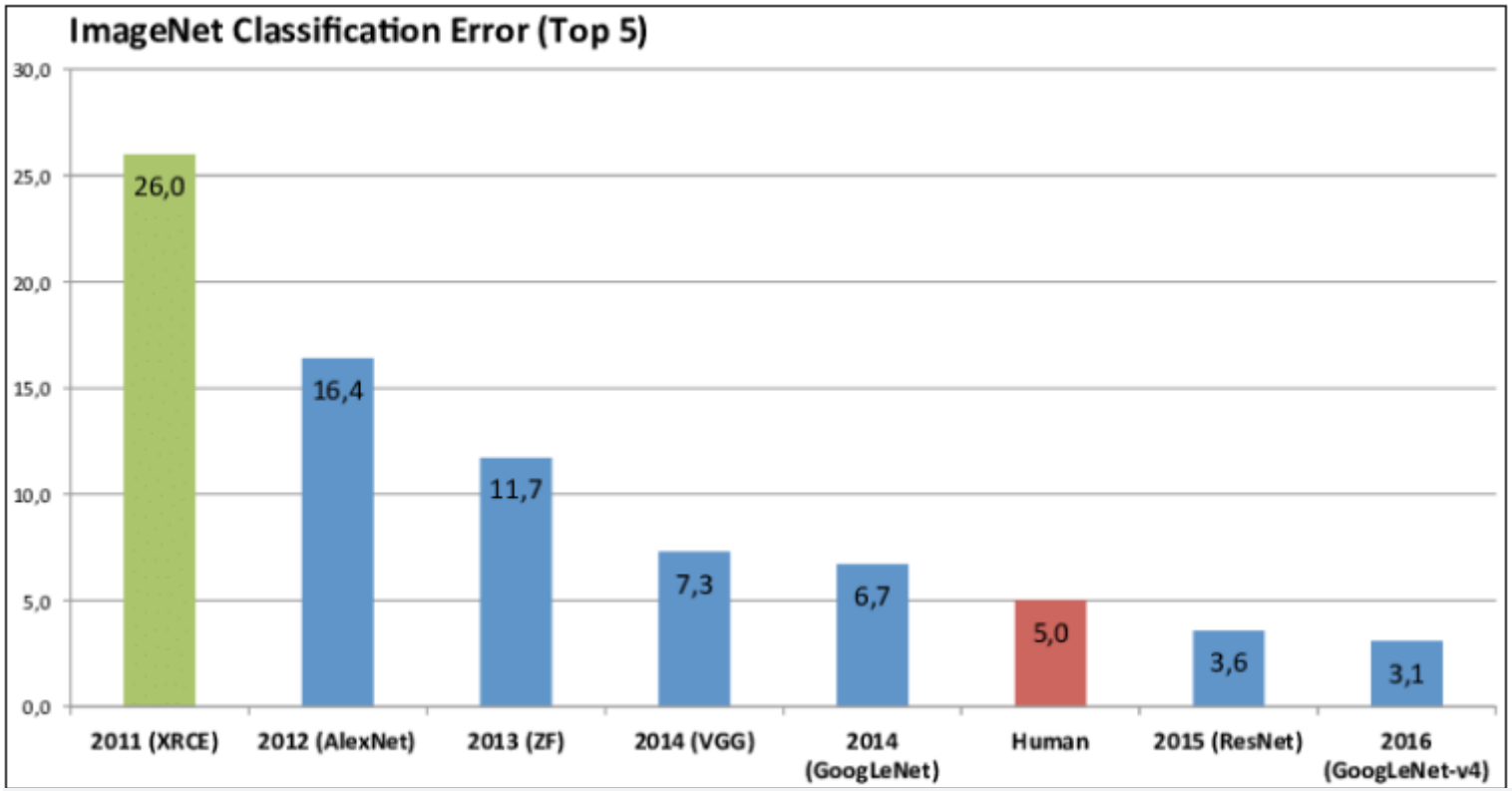
https://www.youtube.com/watch?v=qc5P2bvfl44

Image inpainting:
https://www.youtube.com/watch?v=gg0F5JjKmhA

How did we get there?

Milestones in the Development of Neural Networks

Today, deep networks are (partially) beating humans at image recognition!



https://www.embedded-vision.com/industry-analysis/blog/deep-learning-five-and-half-minutes

This course

Regression and classification using *fully connected* networks

- Learn main principles of deeply nested network architectures
- Implement fully connected networks yourself using NumPy
- Write your own optimization algorithm for training such networks
- Learn how to validate and test your performance

Advanced network architectures using PyTorch

- Learn how to work with images using *convolutional neural networks (CNNs)*
- Weight initialization, self-normalization, and skip-connections for improved training
- Regularization, early-stopping, dropout, and data augmentation for improved generalization

Enter a challenge

- Can you put your new skills to use to get maximum accuracy?