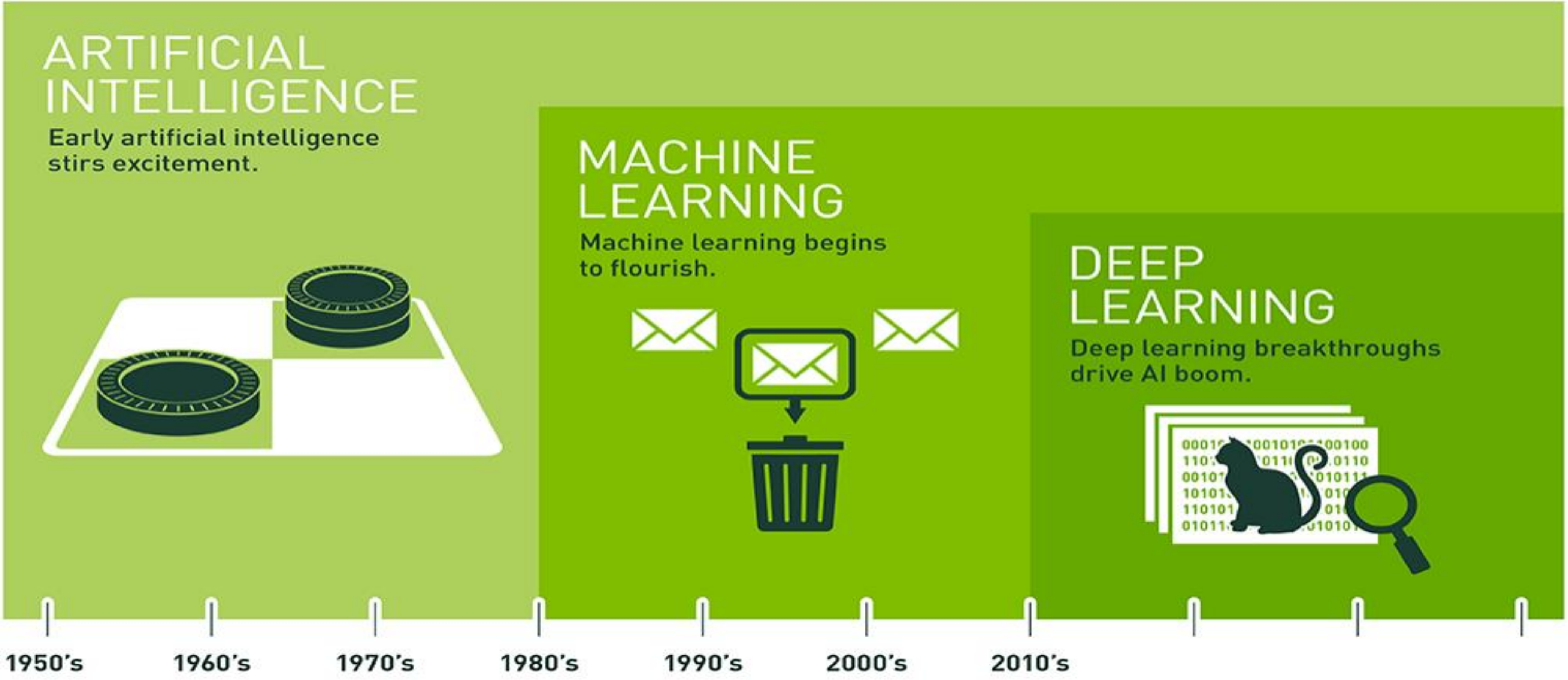
Introduction to Deep Learning

Dr. Alptekin Temizel
Professor, Graduate School of Informatics, METU
DLI Certified Instructor, DLI University Ambassador

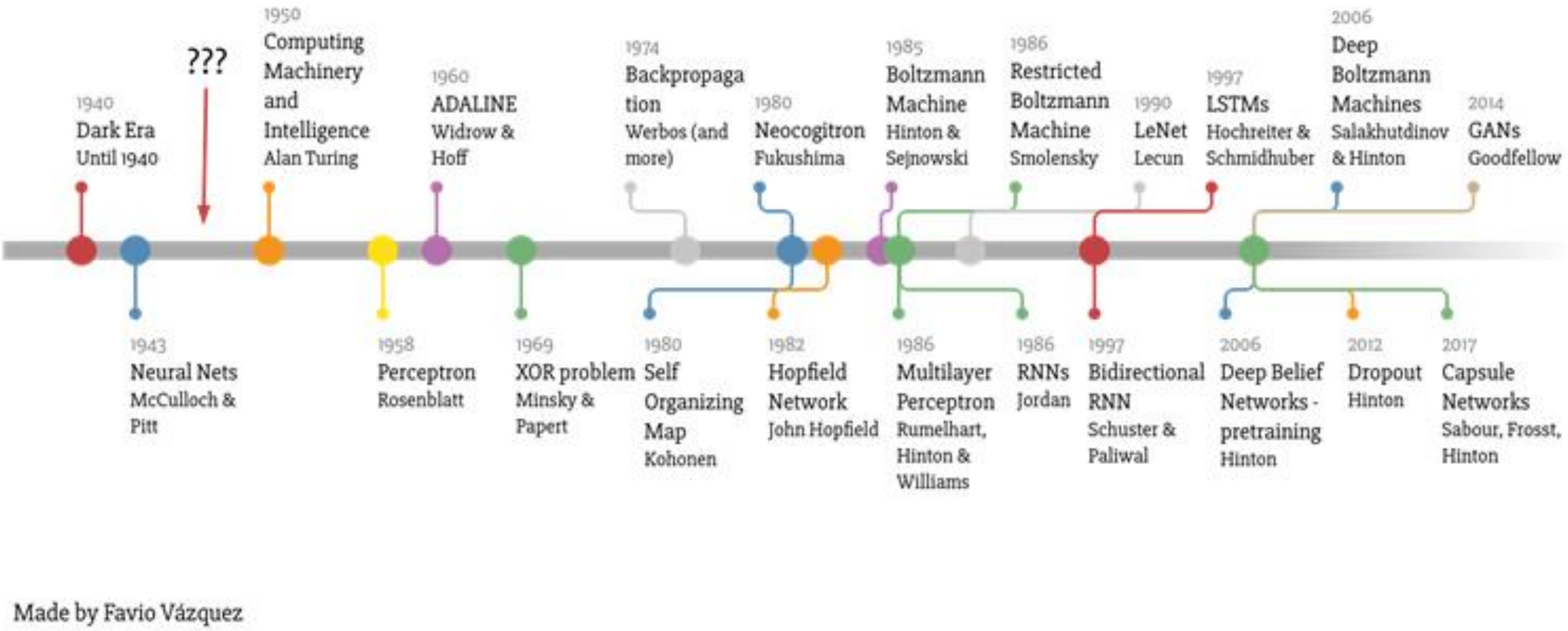
GPU/Deep Learning Related Activities

- NVIDIA Professor Partnership Award – July 2009
- GPU Teaching Center – December 2010
- GPU Research Center – February 2012
- Deep Learning Institute (DLI) Certification and DLI University Ambassadorship– Aug. 2017

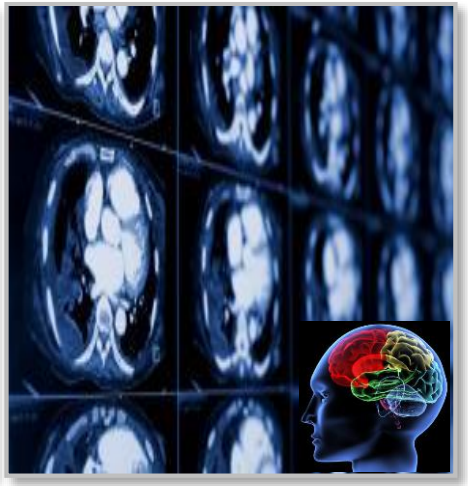
Definitions



Deep Learning Timeline



Deep Learning Across Industries

Internet Services	Medicine	Media & Entertainment	Security & Defense	Autonomous Machines
				
➢ Image/Video classification ➢ Speech recognition ➢ Natural language processing	➢ Cancer cell detection ➢ Diabetic grading ➢ Drug discovery	➢ Video captioning ➢ Content based search ➢ Real time translation	➢ Face recognition ➢ Video surveillance ➢ Cyber security	➢ Pedestrian detection ➢ Lane tracking ➢ Recognize traffic signs

ODTÜ METU

NVIDIA DEEP LEARNING INSTITUTE

THE EXPANDING UNIVERSE OF MODERN AI

"THE BIG BANG"

Big Data
GPU
Algorithms

RESEARCH

Berkeley
Carnegie Mellon University
DEEPMIND
MIT
NYU
OpenAI
Université de Montréal
OXFORD
TORONTO

CORE TECHNOLOGY / FRAMEWORKS

facebook. — torch
Google — TensorFlow
Microsoft — CNTK
Preferred Networks — Chainer
Université de Montréal — theano
Berkeley — Caffe
OXFORD
NVIDIA. cuDNN

AI-as-a-PLATFORM

amazon web services
IBM Watson
Google
Microsoft Azure

START-UPS

api.ai
Personal Assistants
conversational interface

BLUE RIVER TECHNOLOGY
Agriculture
crop-yield optimization

clarifai
Tech
visual recognition platform

deep genomics
Genomics
genetic interpretation

drive.ai
Automotive
computer vision

MetaMind
eCommerce & Medical
recommendation engines

Morpho
Tech
computer vision

Orbital Insight
Geospatial
predictions from images

nervana
Tech
AI-as-a-service

SADAKO
Waste Management
sorting robots

SocialEyes
Medical
diabetic retinopathy

Education
teaching robots

1,000+ AI START-UPS
$5B IN FUNDING
Source: Venture Scanner

INDUSTRY LEADERS

Alibaba.com
AstraZeneca
Audi
Baidu 百度
Bloomberg
BMW
charles SCHWAB
CISCO
ebay
FANUC ROBOTICS
Ford
GE
gsk
The Home Depot
Massachusetts General Hospital
Mercedes-Benz
MERCK
Pinterest
Schlumberger
Shell
SIEME
TARG
TESL
TOYO
Twitter
UBE
VOLV
Walma
YAHO
Yande
yelp

Popularity of Deep Learning

Albert Einstein

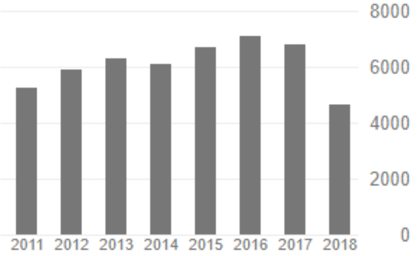
Institute of Advanced Studies, Princeton
No verified email
Physics

FOLLOW

TITLE	CITED BY	YEAR
Can quantum-mechanical description of physical reality be considered complete? A Einstein, B Podolsky, N Rosen Physical review 47 (10), 777	17154	1935
Uber einen die Erzeugung und Verwandlung des Lichtes betreffenden heurischen Gesichtpunkt A Einstein Ann. Phys. 17, 132-148	11080 *	1905

Cited by VIEW ALL

	All	Since 2013
Citations	121368	37858
h-index	112	67
i10-index	369	219



Alex Krizhevsky

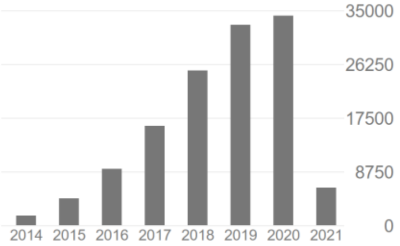
Dessa
Verified email at dessa.com
Machine Learning

FOLLOW

TITLE	CITED BY	YEAR
Imagenet classification with deep convolutional neural networks A Krizhevsky, I Sutskever, GE Hinton Advances in neural information processing systems 25, 1097-1105	82404	2012
Dropout: a simple way to prevent neural networks from overfitting N Srivastava, G Hinton, A Krizhevsky, I Sutskever, R Salakhutdinov The journal of machine learning research 15 (1), 1929-1958	26688	2014

Cited by

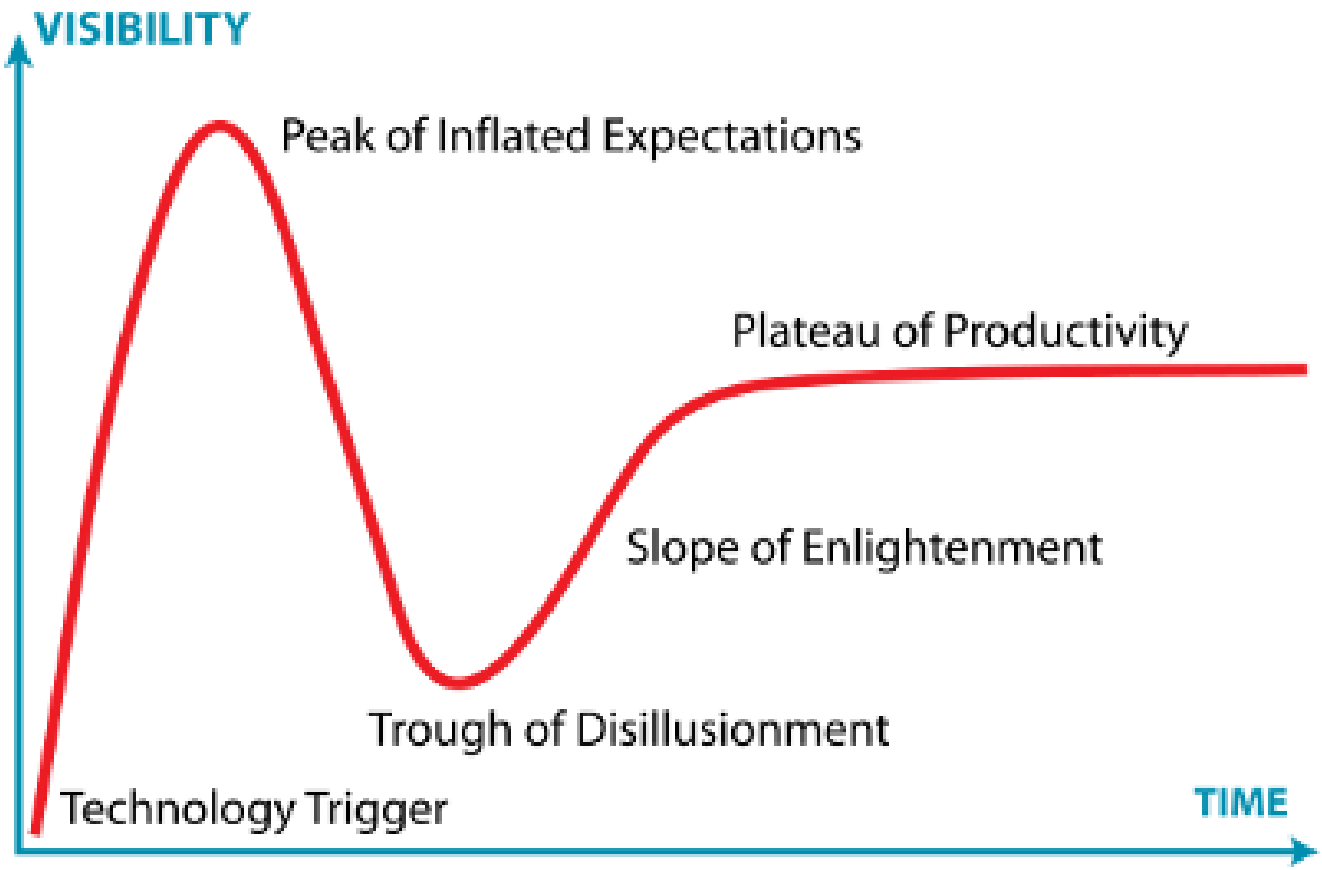
	All	Since 2016
Citations	132918	123929
h-index	24	24
i10-index	25	25



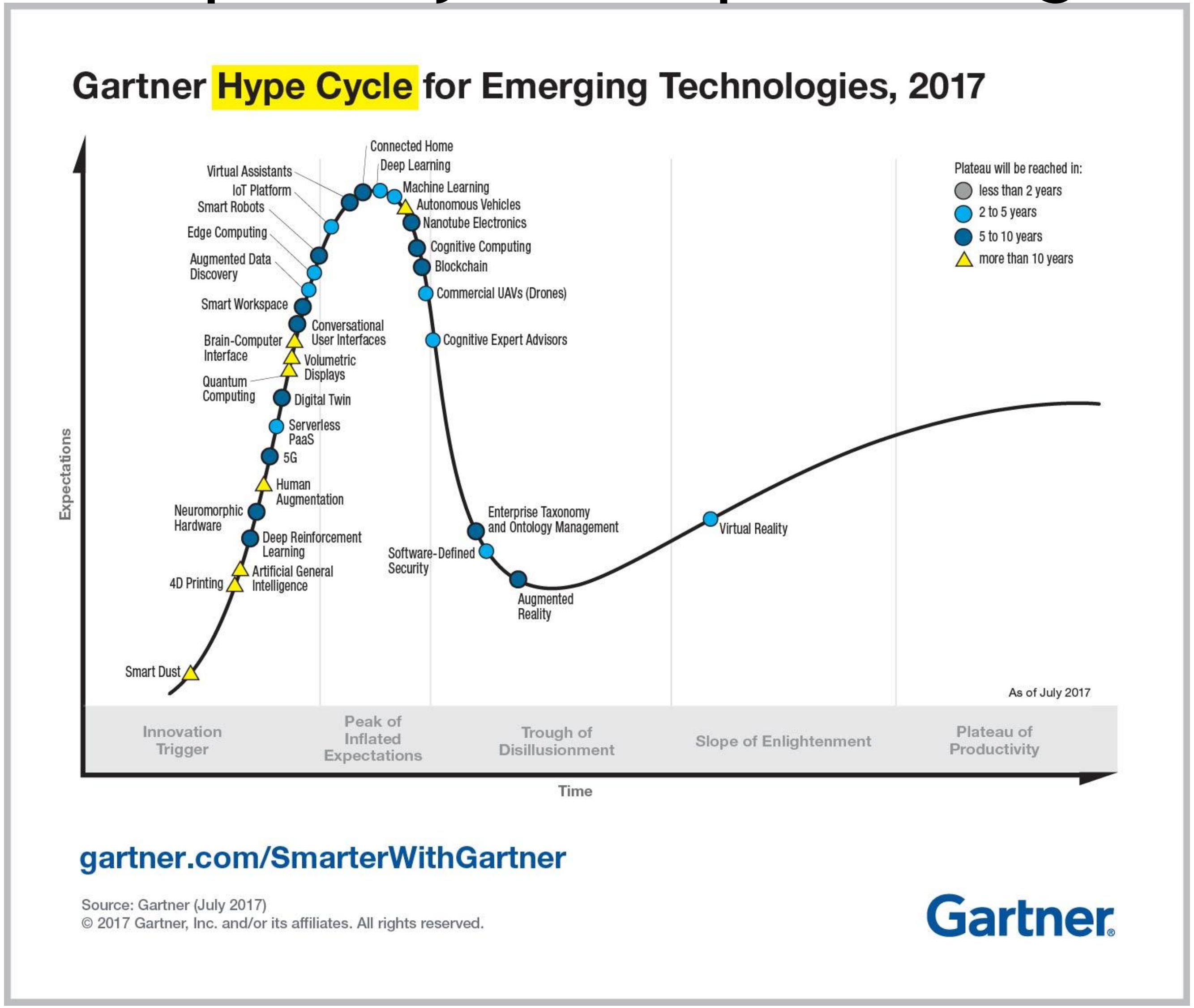
Disclaimer: Number of citations is not a metric that could or should be used to compare the scientific quality or significance of the papers. This slide is only intended to demonstrate the **popularity** of deep learning.

ODTÜ METU
DEEP LEARNING INSTITUTE

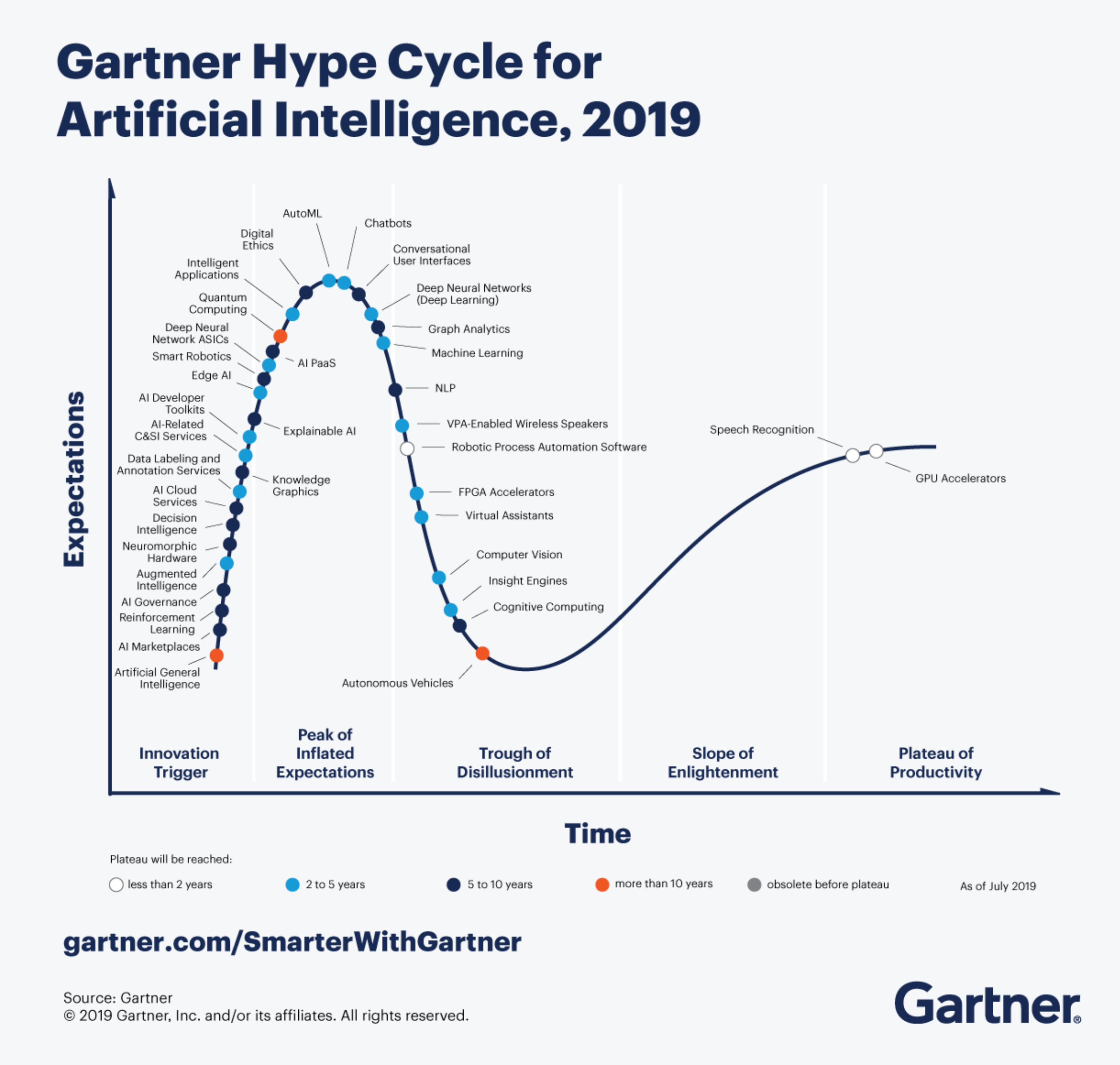
Popularity of Deep Learning



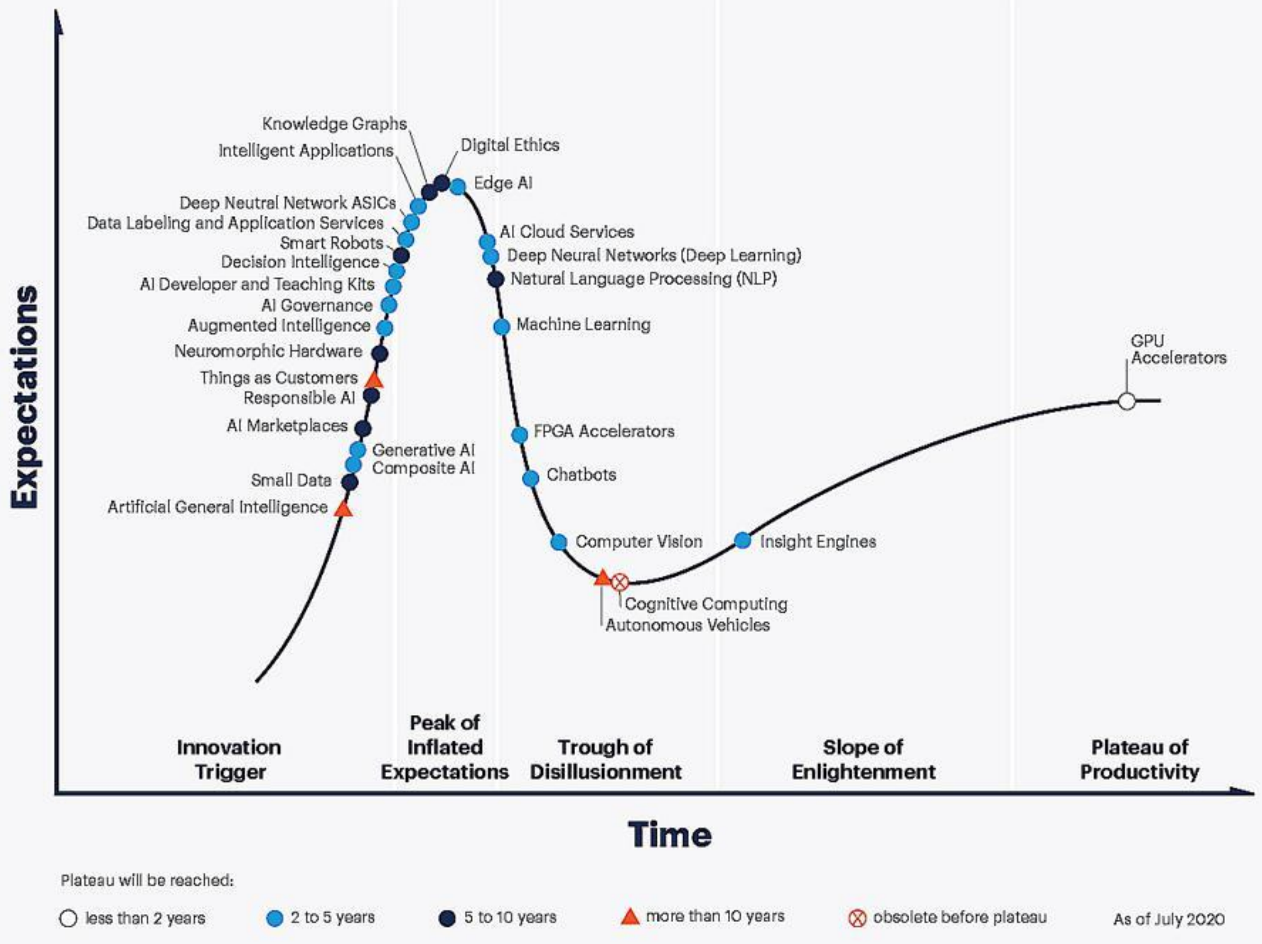
Popularity of Deep Learning



Popularity of Deep Learning



Hype Cycle for Artificial Intelligence, 2020



gartner.com/SmarterWithGartner

Source: Gartner

What Deep Learning Is NOT!

✗ Deep learning algorithms do not aim to accurately model the brain
✓ They are machine learning systems with **some** inspiration from neurons

✗ Deep learning systems do not have “deep” thoughts (or any thoughts!)
✓ The architecture has many layers, implying a **deep architecture**

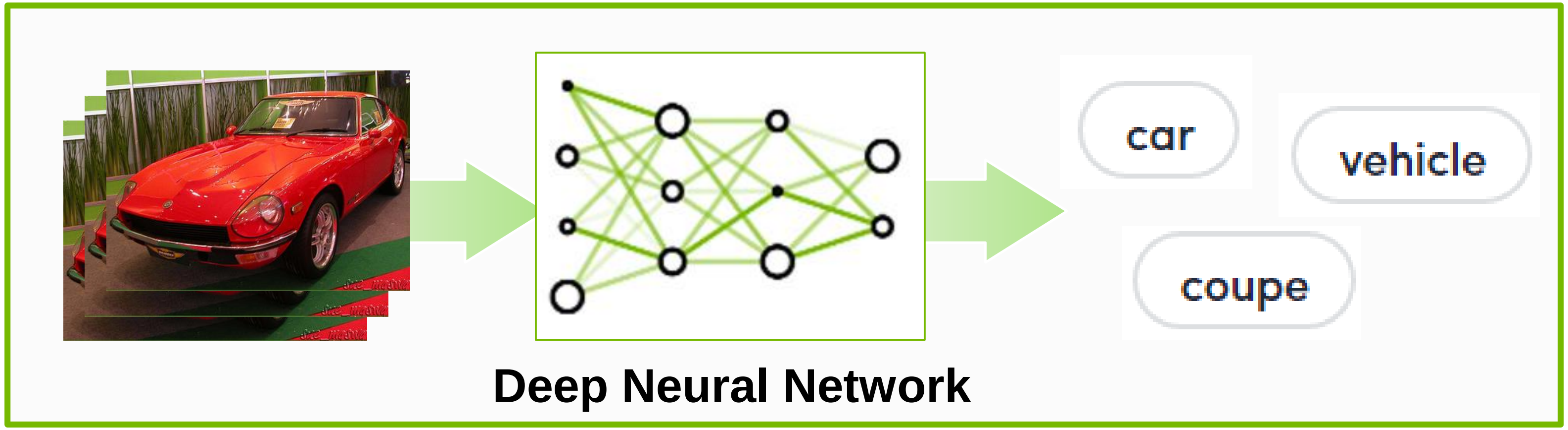
✗ Deep learning systems do not try to compete with human learning
✓ They aim to solve **particular problems** using statistical machine learning techniques and for some particular problems they have better accuracy than humans

A New Computing Model

Algorithms that Learn from Examples

Traditional Approach

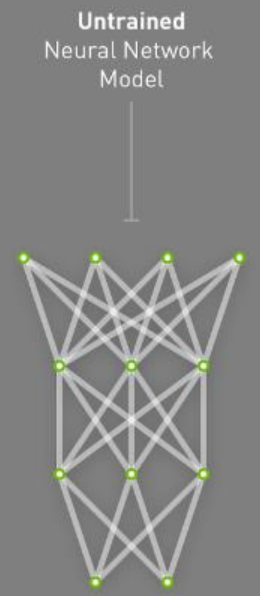
- ➢ Requires domain experts
- ➢ Time consuming
- ➢ Not scalable to new problems



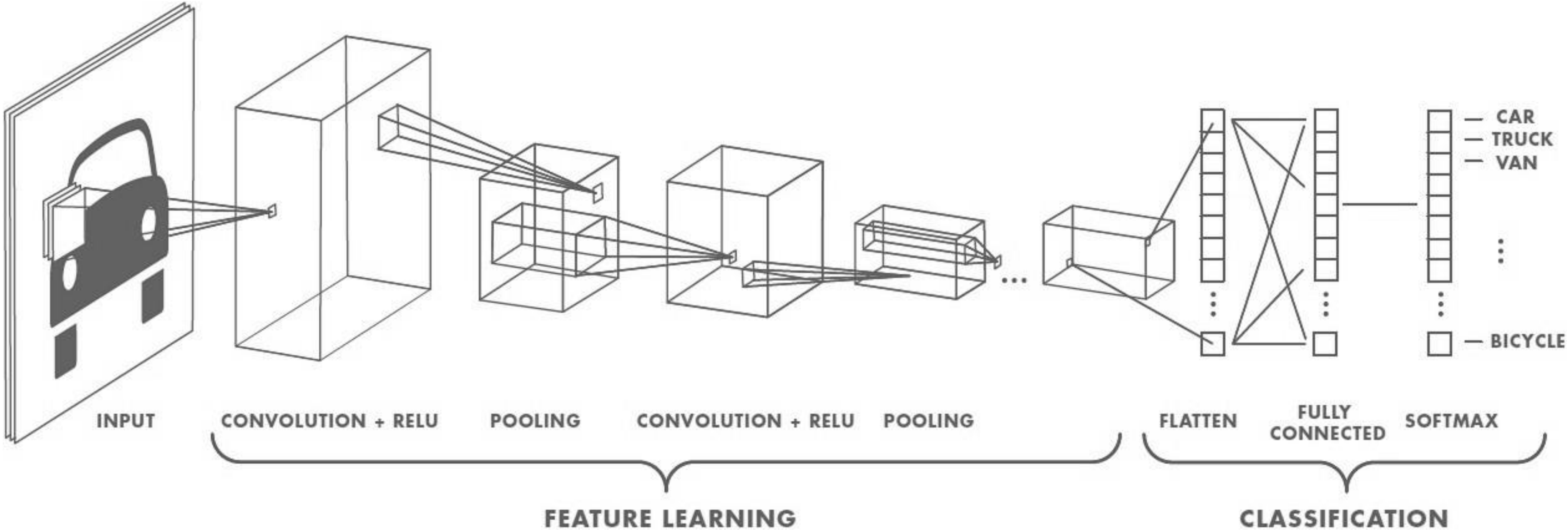
Deep Learning Approach

- ✓ Learn from data
- ✓ Easy to extend

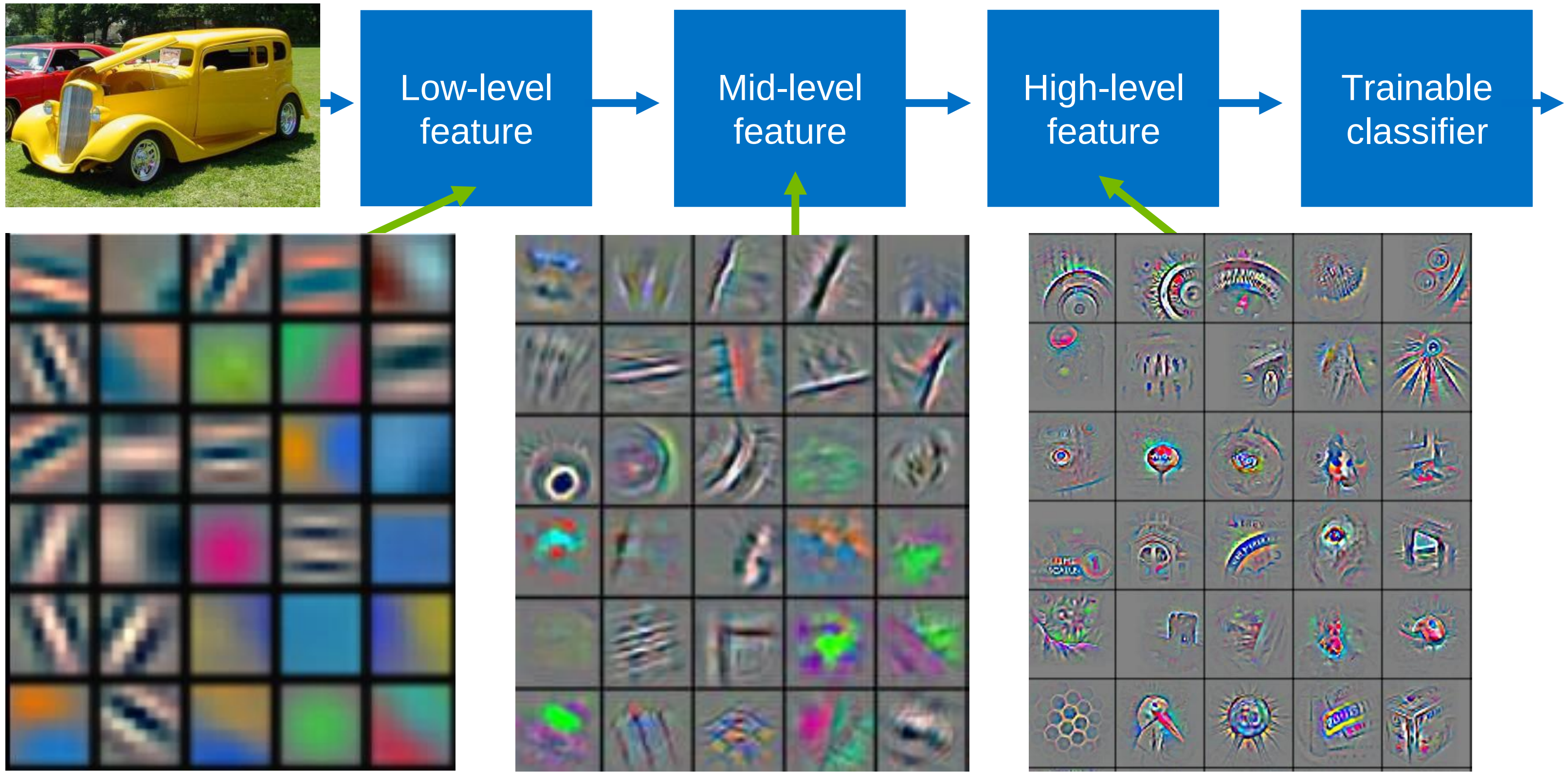
DEEP LEARNING



Deep Learning – Convolutional Networks



Deep Learning – Convolutional Networks



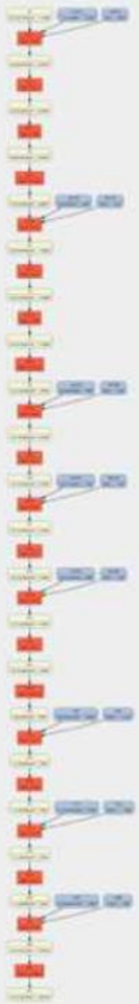
Feature visualization of convolutional net trained on ImageNet from [Zeiler & Fergus 2013]

How Deep is Enough?

AlexNet (2012)

5 convolutional layers

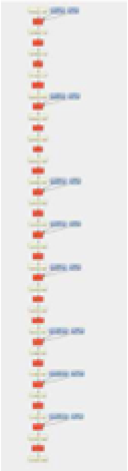
3 fully-connected layers



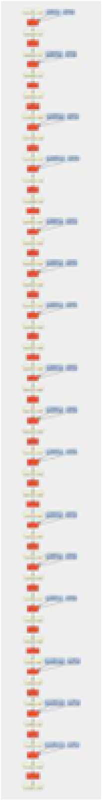
How Deep is Enough?

AlexNet (2012)

VGG-M (2013)



VGG-D (2013)

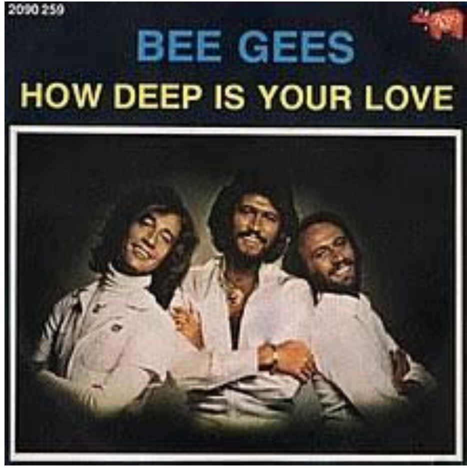


- AlexNet:5 convolutional layers,
- VGG-D: 16 layers
- VGG-E: 19 layers

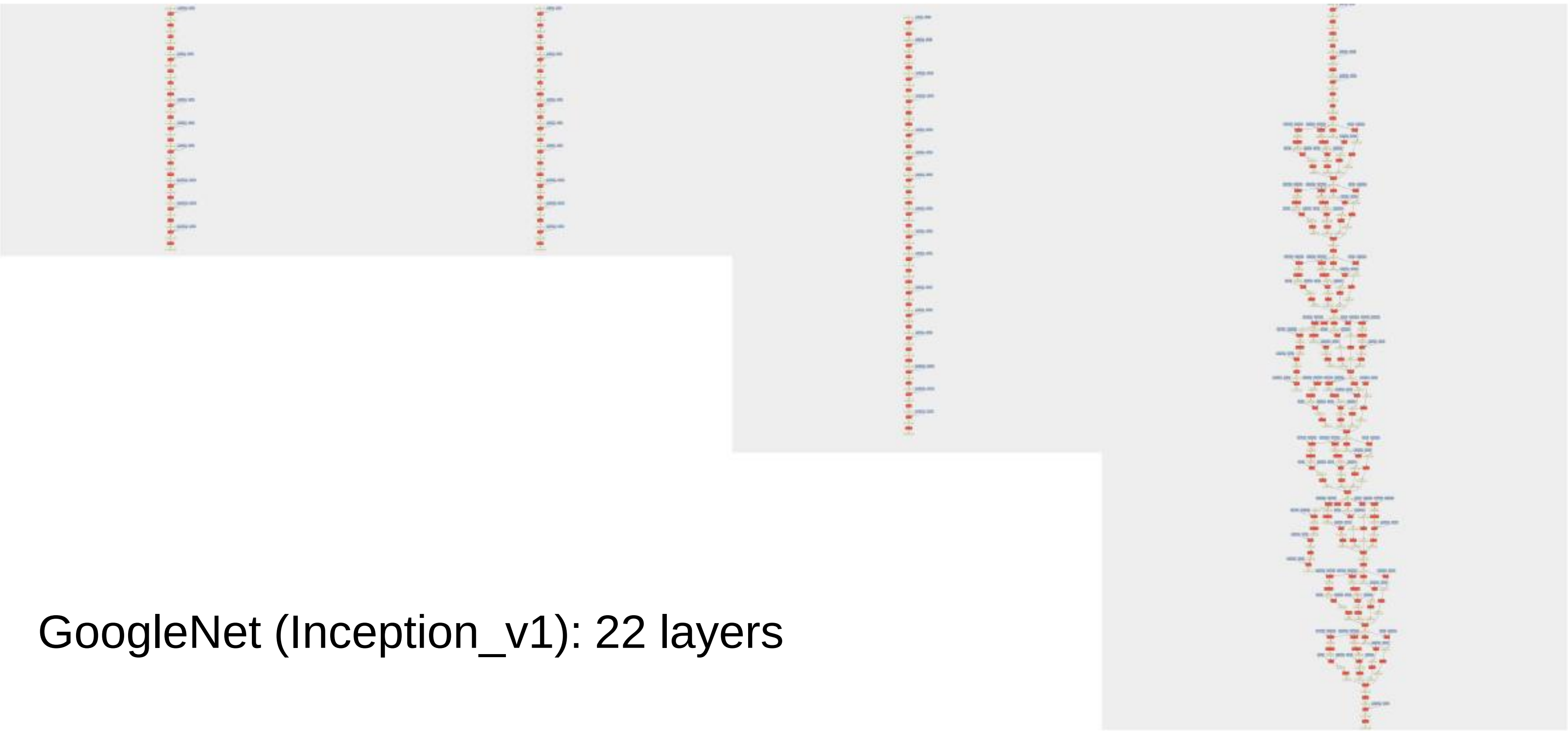
- VGG-E Top-5 error rate: 7.3%

- To reduce the number of parameters in such very deep networks, they used smaller 3×3 filters in all convolutional layers

How Deep is Enough?

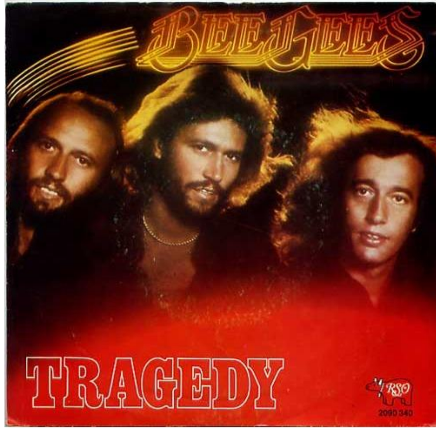


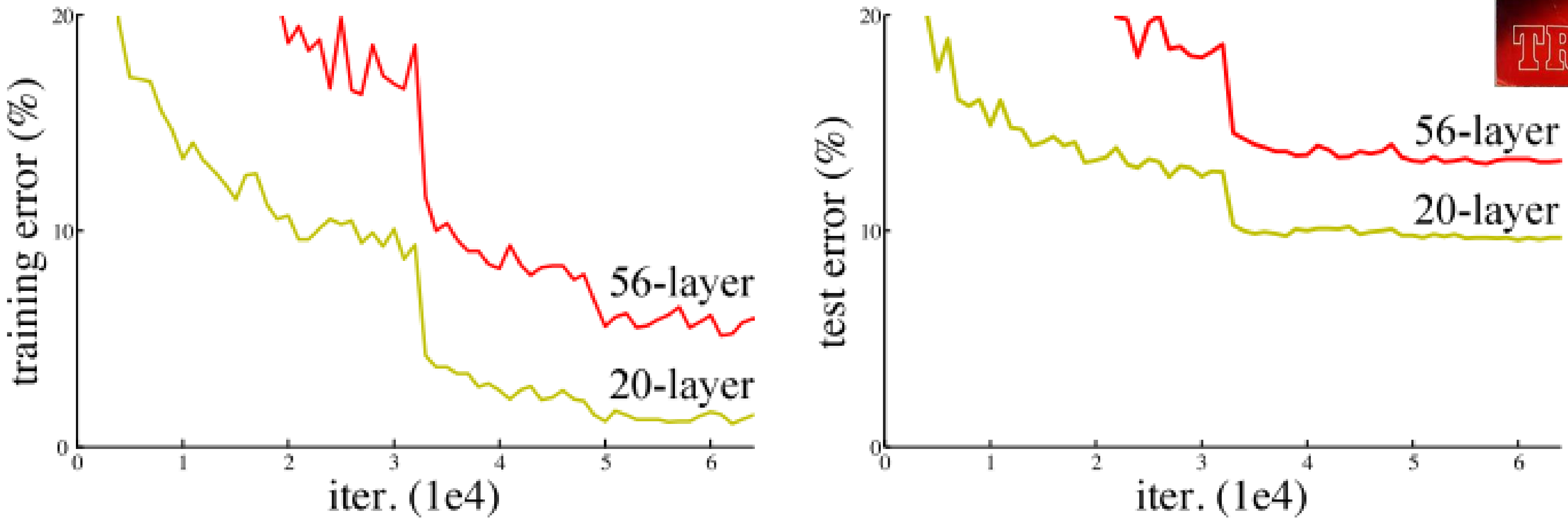
AlexNet (2012) VGG-M (2013) VGG-VD-16 (2014) GoogLeNet (2014)



GoogleNet (Inception_v1): 22 layers

How Deep is Enough?

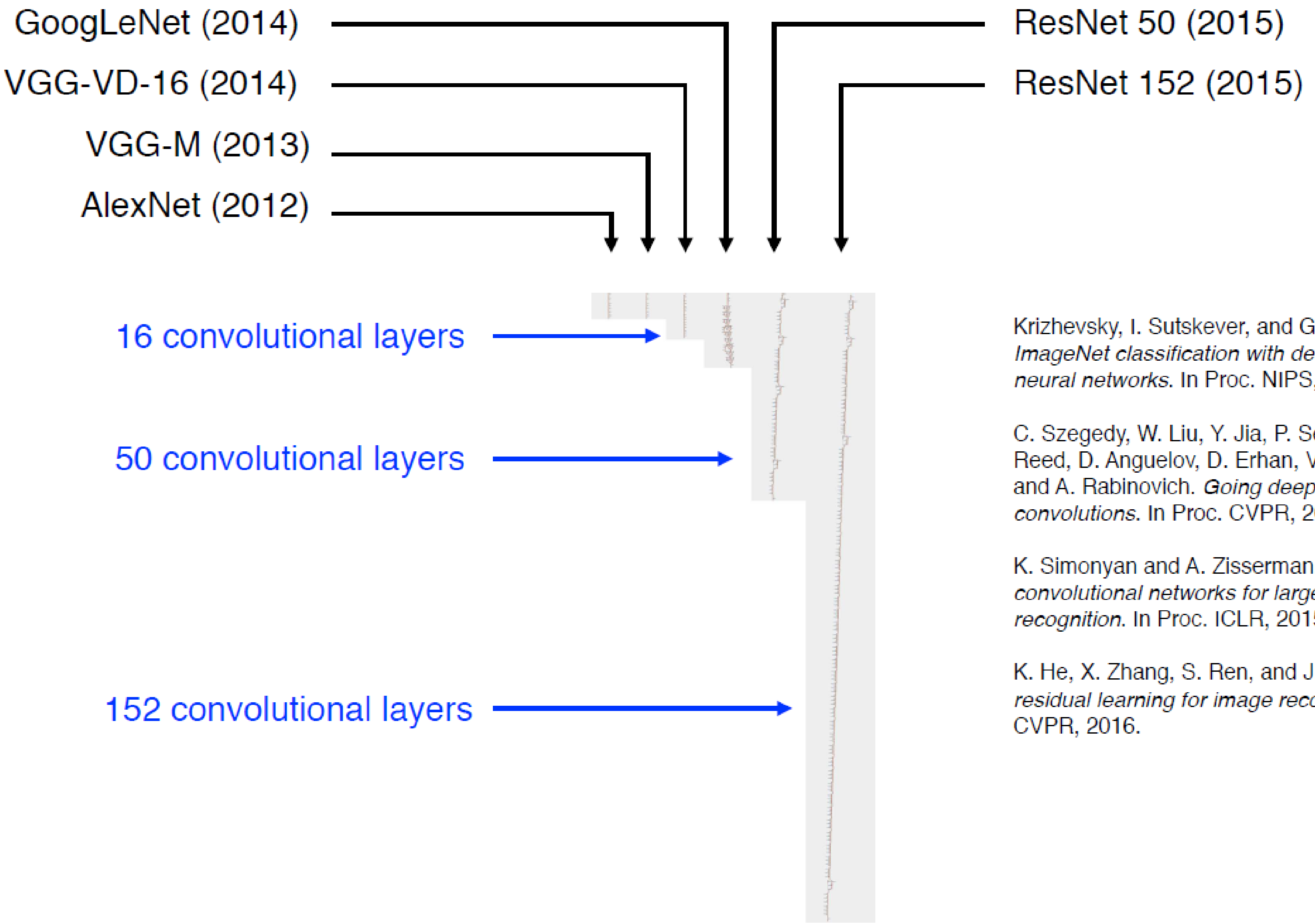


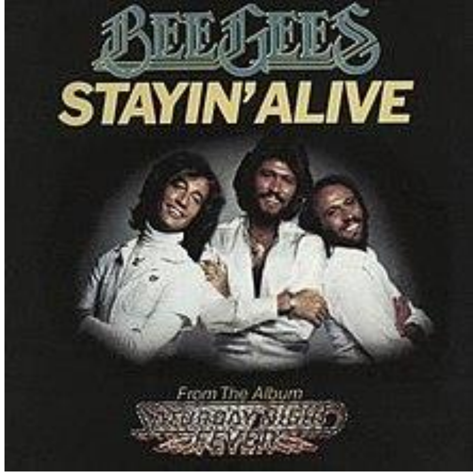


Increasing network depth leads to worse performance!

Vanishing gradients problem: gradients do not propagate through so many layers (they become smaller and smaller) to the earlier layers

How Deep is Enough?





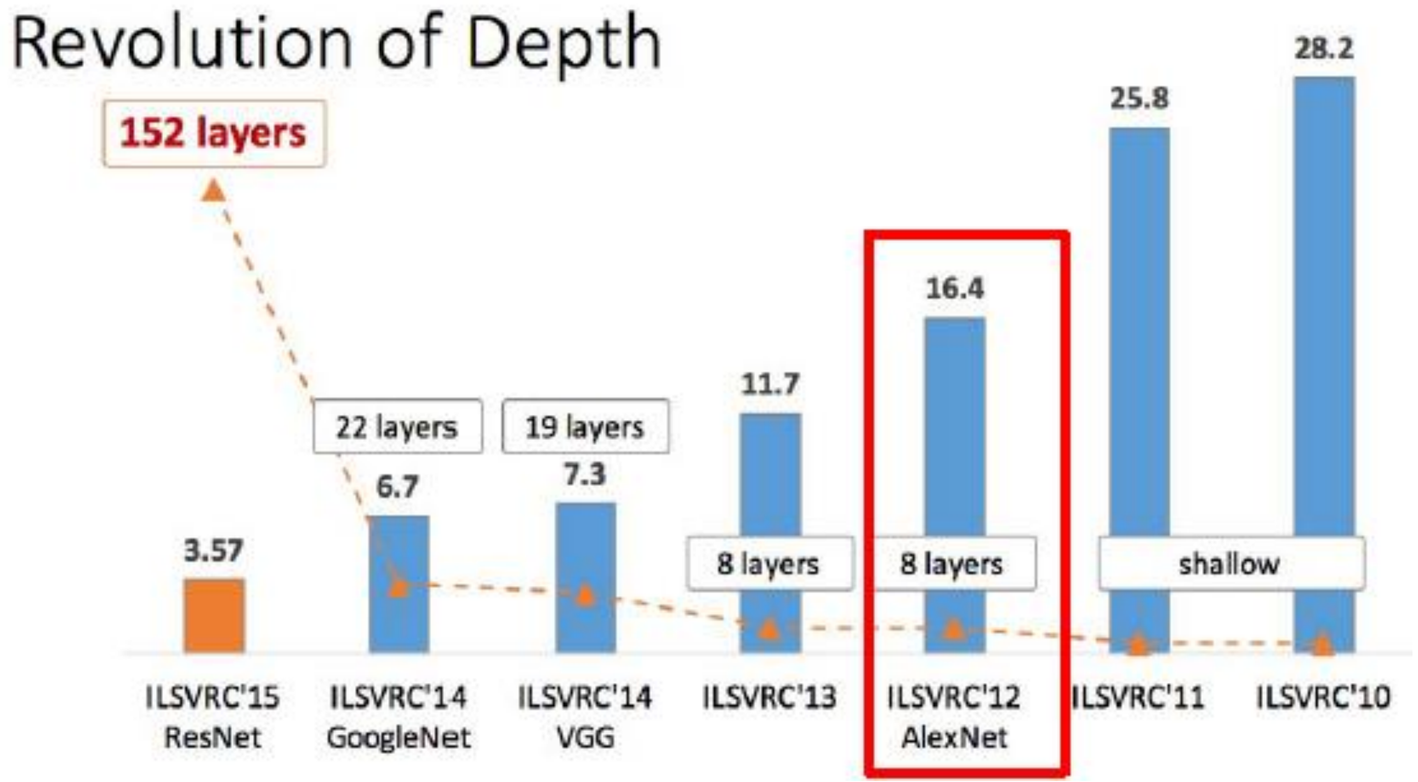
Krizhevsky, I. Sutskever, and G. E. Hinton. *ImageNet classification with deep convolutional neural networks*. In Proc. NIPS, 2012.

C. Szegedy, W. Liu, Y. Jia, P. Sermanet, S. Reed, D. Anguelov, D. Erhan, V. Vanhoucke, and A. Rabinovich. *Going deeper with convolutions*. In Proc. CVPR, 2015.

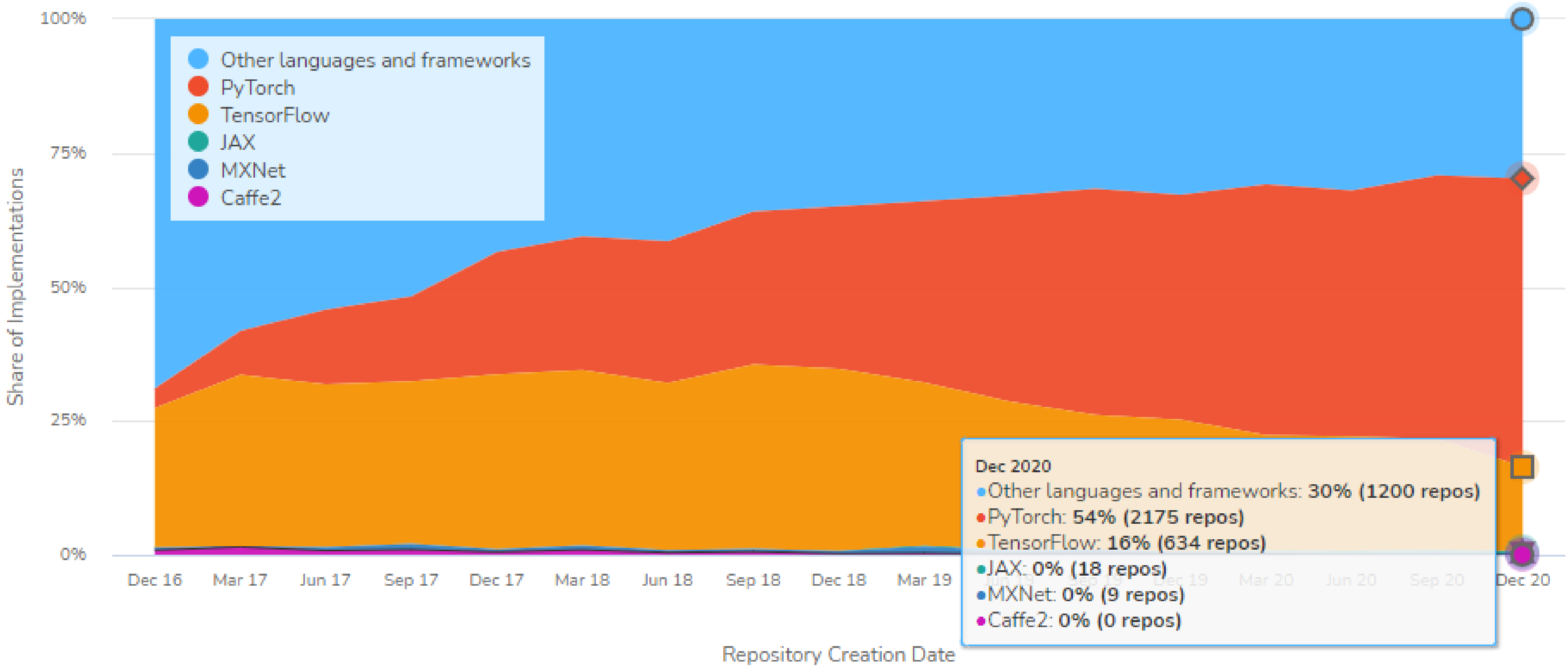
K. Simonyan and A. Zisserman. *Very deep convolutional networks for large-scale image recognition*. In Proc. ICLR, 2015.

K. He, X. Zhang, S. Ren, and J. Sun. *Deep residual learning for image recognition*. In Proc. CVPR, 2016.

How Deep is Enough?

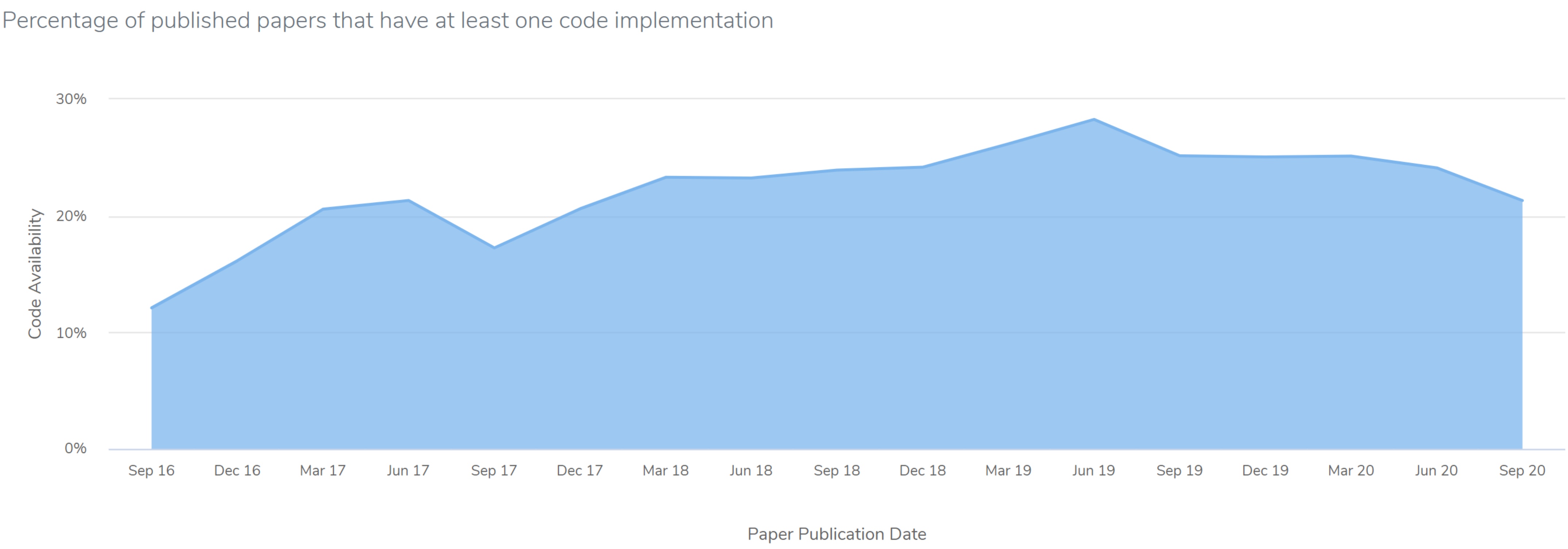


Deep Learning Software – Training



https://paperswithcode.com/trends

Deep Learning Software – Training

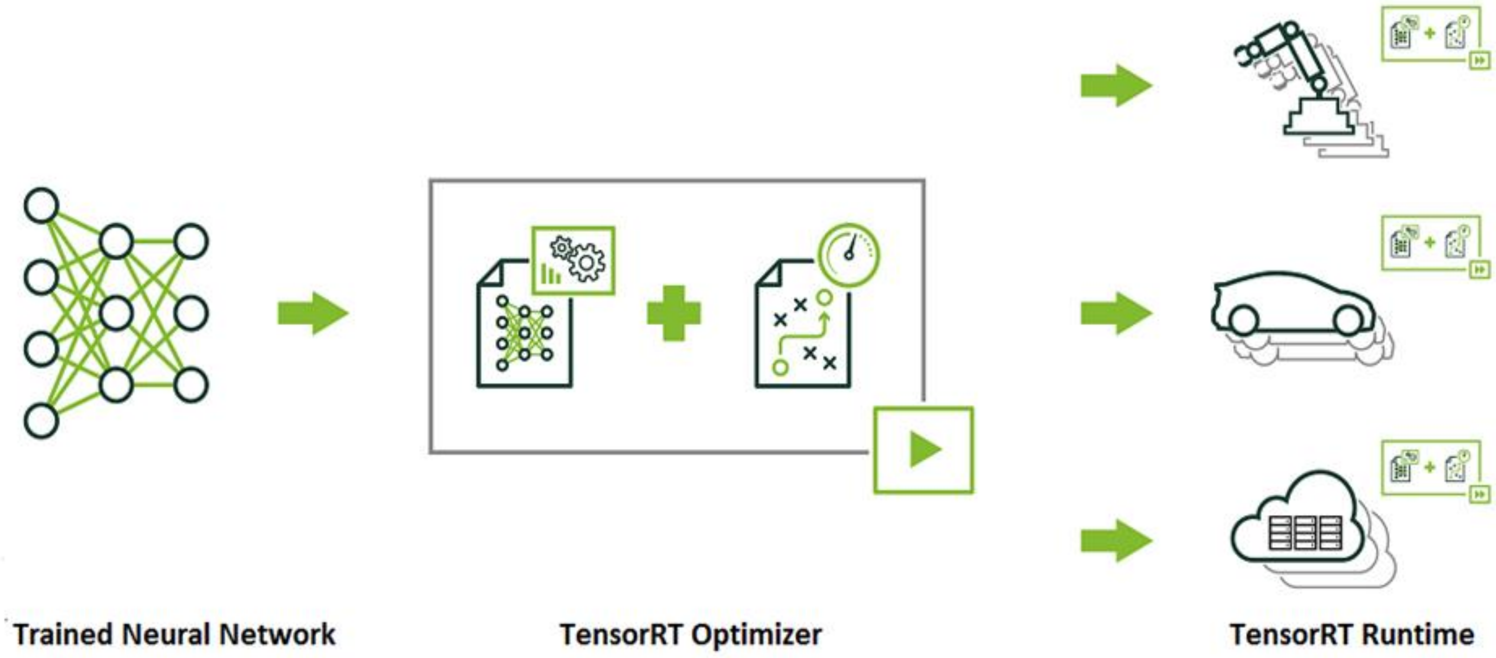


https://paperswithcode.com/trends

Deep Learning Software – Inference

TensorRT

Inference engine for production deployment of deep learning applications



Deep Learning Hardware

TRAINING

FULLY INTERGRATED DL SUPERCOMPUTER

DGX-1 & DGX Station

DESKTOP

RTX/GTX series

DATA CENTER

Tesla A100
Tesla V100

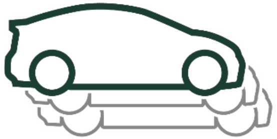
INFERENCE

DATA CENTER

Tesla A100/V100

Tesla T4

AUTOMOTIVE



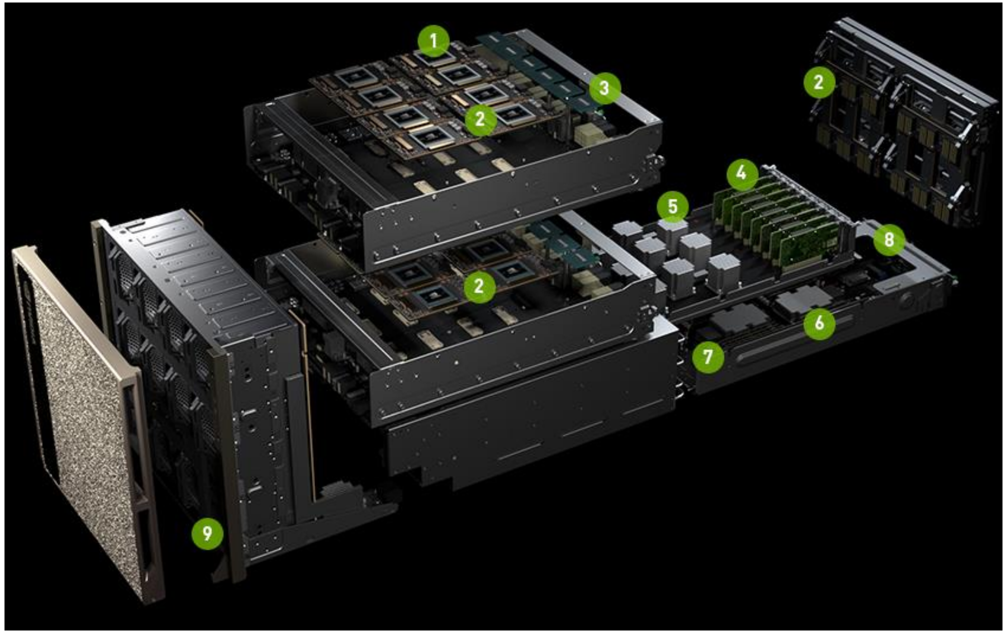
Drive PX2

EMBEDDED

Jetson TX2

Jetson Xavier

Deep Learning Hardware



NVIDIA® DGX™-2, the first 2 petaFLOPS system that combines 16 fully interconnected GPUs

Edge AI-NVIDIA Jetson Xavier

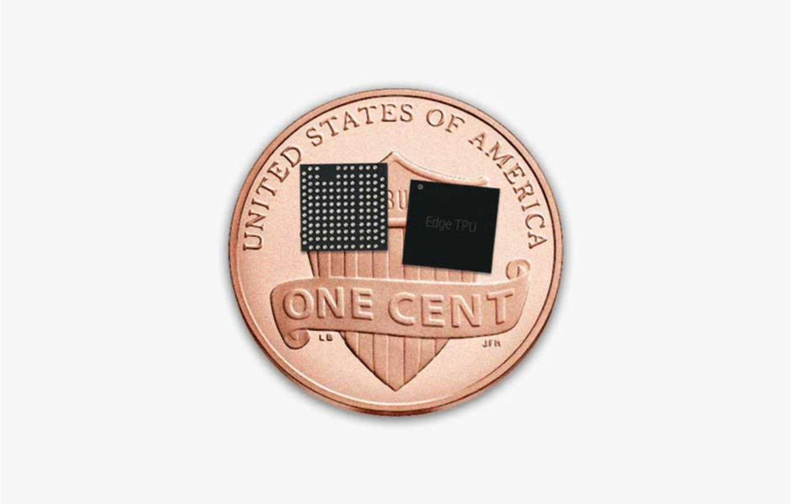
Jetson AGX Xavier	
GPU	512-core Volta GPU with Tensor Cores
CPU	8-core ARM v8.2 64-bit CPU, 8MB L2 + 4MB L3
Memory	16GB 256-Bit LPDDR4x | 137GB/s
Storage	32GB eMMC 5.1
DL Accelerator	(2x) NVDLA Engines*
Vision Accelerator	7-way VLIW Vision Processor*
Encoder/Decoder	(2x) 4Kp60 | HEVC/(2x) 4Kp60 | 12-Bit Support
Size	105 mm x 105 mm

ODTÜ METU

NVIDIA DEEP LEARNING INSTITUTE

Edge AI- Google TPU

- Edge TPU: a small ASIC designed by Google that provides high performance ML inferencing for low-power devices.
- It can execute state-of-the-art mobile vision models such as MobileNet V2 at 100+ fps, in a power efficient manner.
- Supports Tensorflow Lite



Two Edge TPU chips on the head of a US penny

USB Accelerator with Edge TPU

Edge AI– AWS Deeplens

The world's first deep learning enabled video camera for developers

AWS DeepLens helps put deep learning in the hands of developers, literally, with a fully programmable video camera, tutorials, code, and pre-trained models designed to expand deep learning skills.

Get started with your DeepLens

ACTIVITY RECOGNITION

Recognize more than 30 kinds of actions such as brushing teeth, applying lipstick, and playing guitar.

AWS DeepLens - Deep learning enabled video camera for developers

by Amazon Web Services

27 customer reviews | 27 answered questions

Price: $249.00 & **FREE Shipping**. Details

prime | Try Fast, Free Shipping

ODTÜ METU

NVIDIA DEEP LEARNING INSTITUTE

Edge AI– Mobile Nets

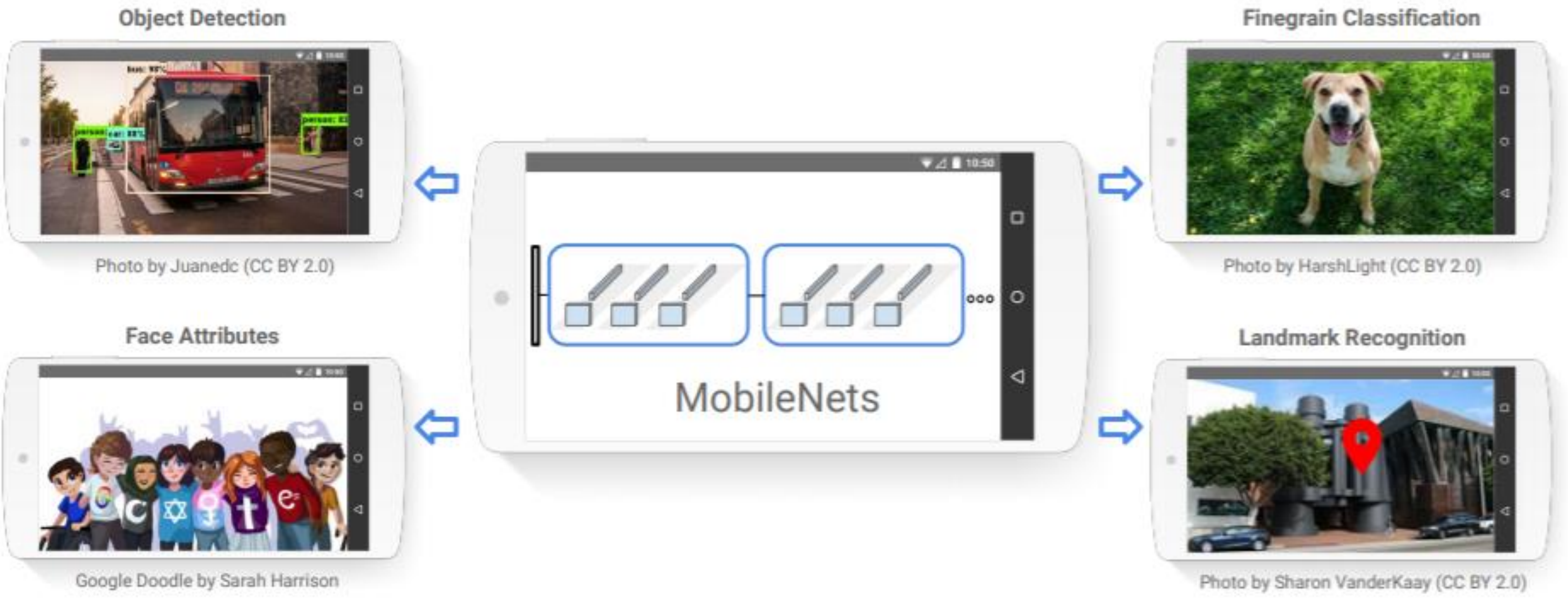


Figure 1. MobileNet models can be applied to various recognition tasks for efficient on device intelligence.

Howard, A.G., Zhu, M., Chen, B., Kalenichenko, D., Wang, W., Weyand, T., Andreetto, M. and Adam, H., 2017. Mobilenets: Efficient convolutional neural networks for mobile vision applications. *arXiv preprint arXiv:1704.04861*.

Deep Learning to Keep Cats from Pooping on Lawn!

Deep Learning to Keep Cats from Pooping on Lawn!

AI PaaS- Amazon AWS

AI Platform as a Service (AI PaaS): AI services provided by cloud vendors.

Vision Services

Amazon Rekognition Image
Deep learning-based image analysis
Learn more »

Amazon Rekognition Video
Deep learning-based video analysis
Learn more »

Conversational chatbots

Amazon Lex
Build chatbots to engage customers
Learn more »

Language Services

Amazon Comprehend
Discover insights and relationships in text
Learn more »

Amazon Translate
Fluent translation of text
Learn more »

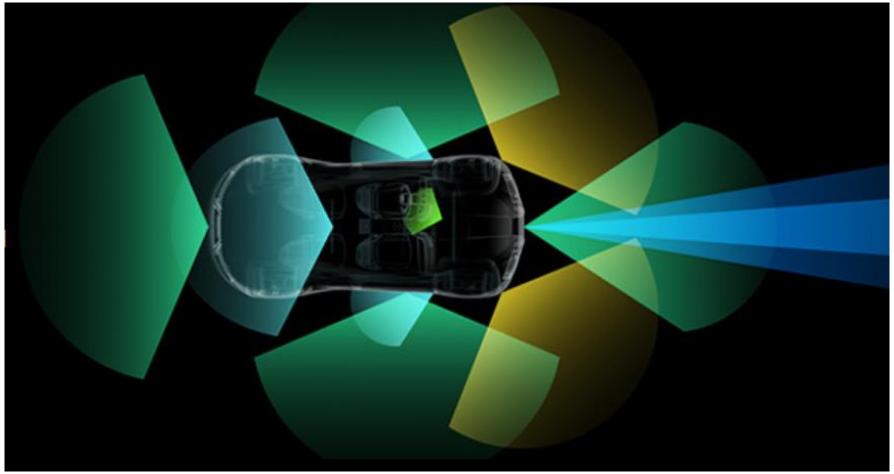
Amazon Transcribe
Automatic speech recognition
Learn more »

Amazon Polly
Natural sounding text to speech
Learn more »

Deep Learning for Autonomous Driving

NVIDIA DRIVE systems can fuse data from multiple cameras, as well as lidar, radar, and ultrasonic sensors.

This allows algorithms to accurately understand the full 360-degree environment around the car to produce a robust representation, including static and dynamic objects.



Deep Learning Project Checklist

1. What problem are you solving, what are the DL tasks?
2. What data do you have/need, and how is it labeled?
3. Which deep learning framework & tools will you use?
4. What already trained models are available?
5. On what platform(s) will you train and deploy?

What Problem Are You Solving?

Defining the AI/DL Tasks

INPUTS	QUESTION	AI/DL TASK	EXAMPLE OUTPUTS
Text Data, Images, Video, Audio	Is "it" present or not?	Detection	Cancer Detection
	What type of thing is "it"?	Classification	Tumor Identification
	To what extent is "it" present?	Segmentation	Tumor Size/Shape Analysis
	What is the likely outcome?	Prediction	Survivability Prediction
	What will likely satisfy the objective?	Recommendation	Therapy Recommendation

What Problem Are You Solving?

Can be a combination or chain of AI tasks to achieve more sophisticated outputs. Some examples:

Family photo: face **detection** followed by facial recognition (**classification**).

Translation: speech to text (**classification**) followed by translation (**prediction**) and then speech synthesis (**prediction**).

Google Maps: business type **detection**, open hours sign **detection** & hours **recognition**, published via Google Maps.

Generative Adversarial Networks (GAN)

JFT-300M dataset

- 512x512 3 channel resolution 300 million images 18000 classes
- Google's internal dataset, not public

48 hours of training on 512 TPUv3

Brock, A., Donahue, J. and Simonyan, K., 2018. Large scale GAN training for high fidelity natural image synthesis. arXiv preprint arXiv:1809.11096.

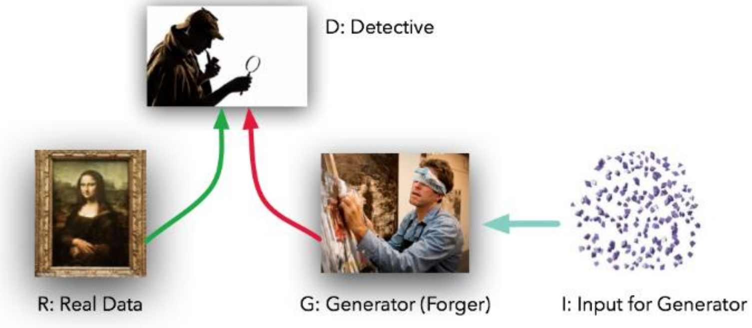
Generative Adversarial Networks (GAN)

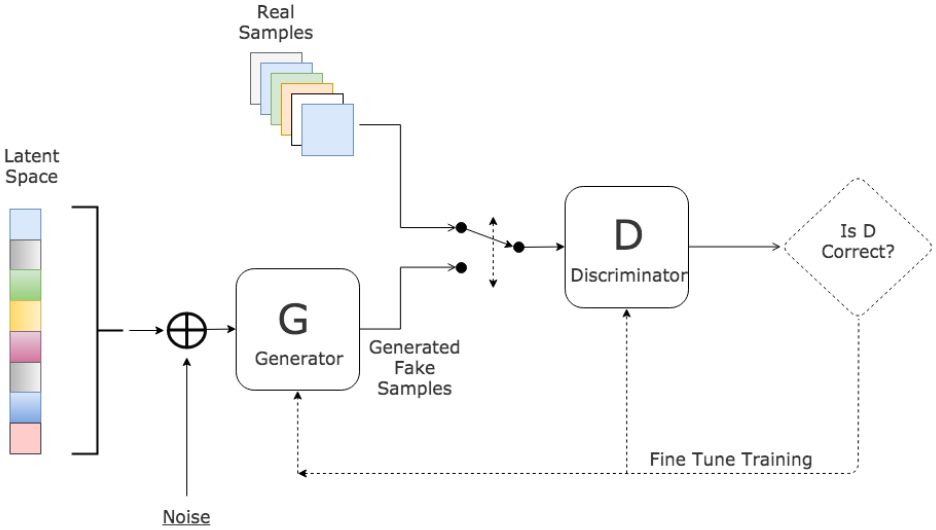
https://www.youtube.com/watch?v=kSLJriaOumA

Karras, T., Laine, S. and Aila, T., 2019. A style-based generator architecture for generative adversarial networks.
In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (pp. 4401-4410).

Generative Adversarial Networks (GAN)

- An adversarial process for estimating generative models
- Consists of 2 simultaneously trained models
  - a generative model **G**
  - a discriminator model **D**
- The generative model G takes random noise as input and generates data candidates
- Discriminator model D tries to distinguish which is real data

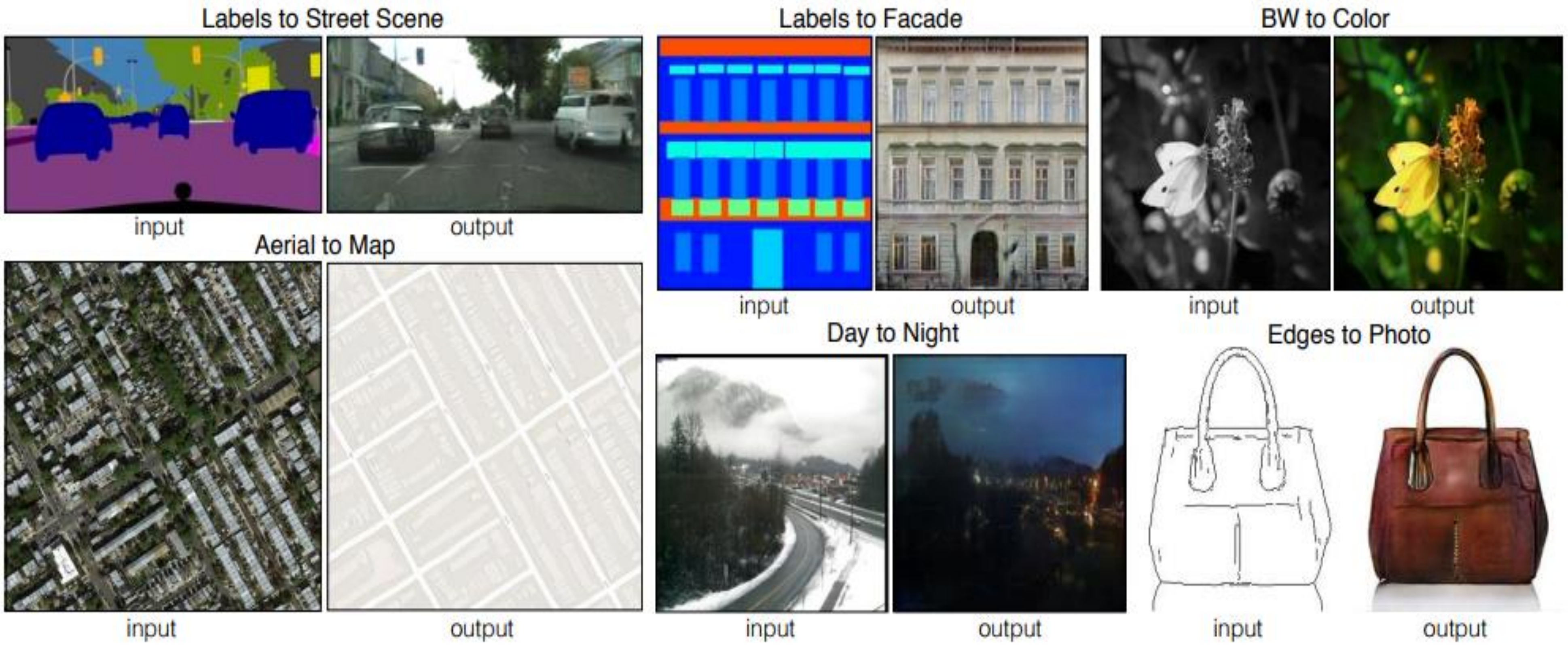




I. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. Courville, and Y. Bengio, 2014. Generative adversarial nets. In *Advances in neural information processing systems*

GAN Applications

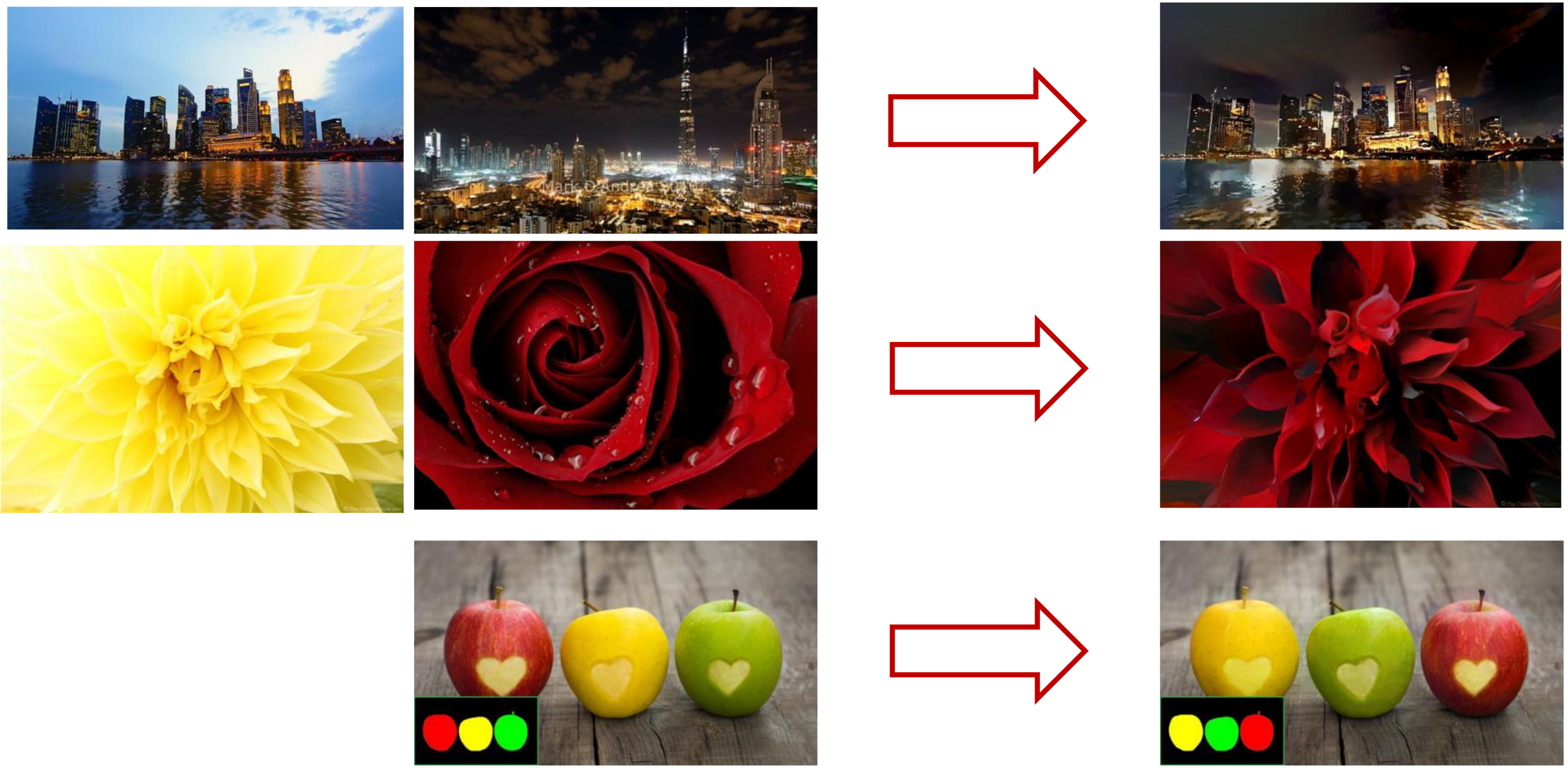
pix2pix A general purpose solution to image-to-image translation
Conditional GAN : generates an output from an image with given condition



P. Isola, J.-Y. Zhu, T. Zhou and A. A. Efros, "Image-to-Image Translation with Conditional Adversarial Networks," in *CVPR*, 2017.

GAN Applications

Style transfer for scene images of cities and wide area
Time of the day, weather, season and artistic edit.

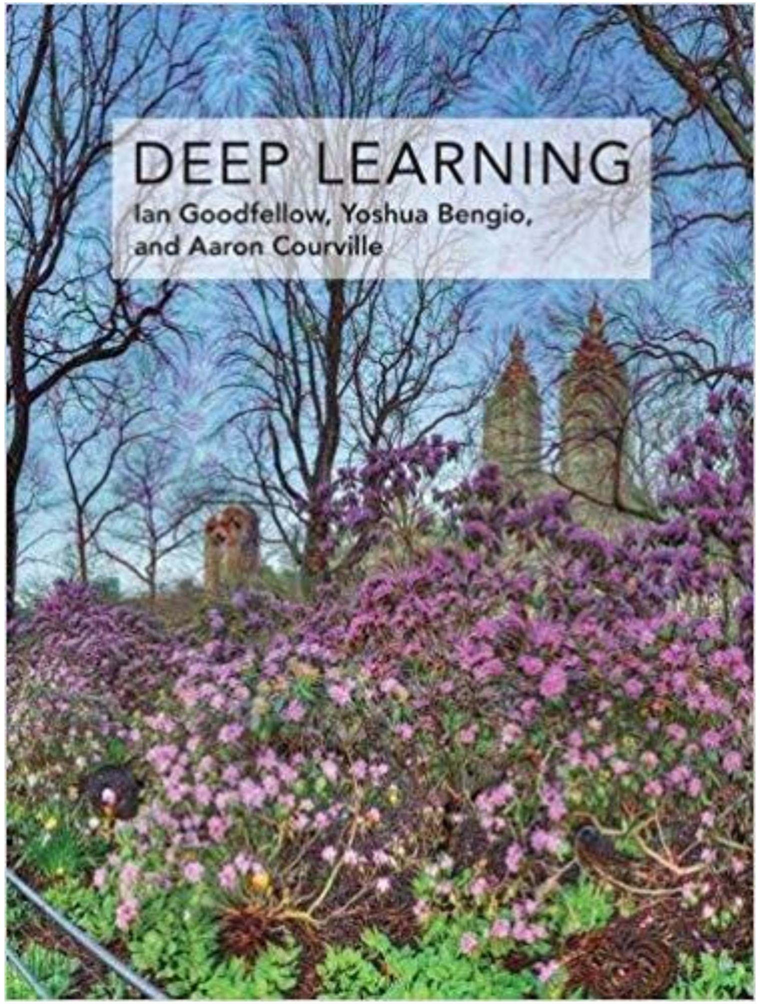


F. Luan, S. Paris, E. Shechtman and K. Bala, "Deep Photo Style Transfer," in *arXiv:1703.07511*, 2017.

Conclusions

- Mature tools, availability of pre-trained models, easy access to GPUs in training and hardware integration for a complete solution makes it easy to develop “intelligent” applications
    - Makes the entry barriers low not just for you but for everyone!
    - Domain expertise makes a big difference!
- Being an *AI expert* and releasing *AI enabled technologies* are very “hot”; but
    - AI experts are expected to have an in-depth understanding and experience in both theory and practice (and a track record)
    - AI enabled technologies are expected to satisfy “inflated” expectations of the users

Reference Book

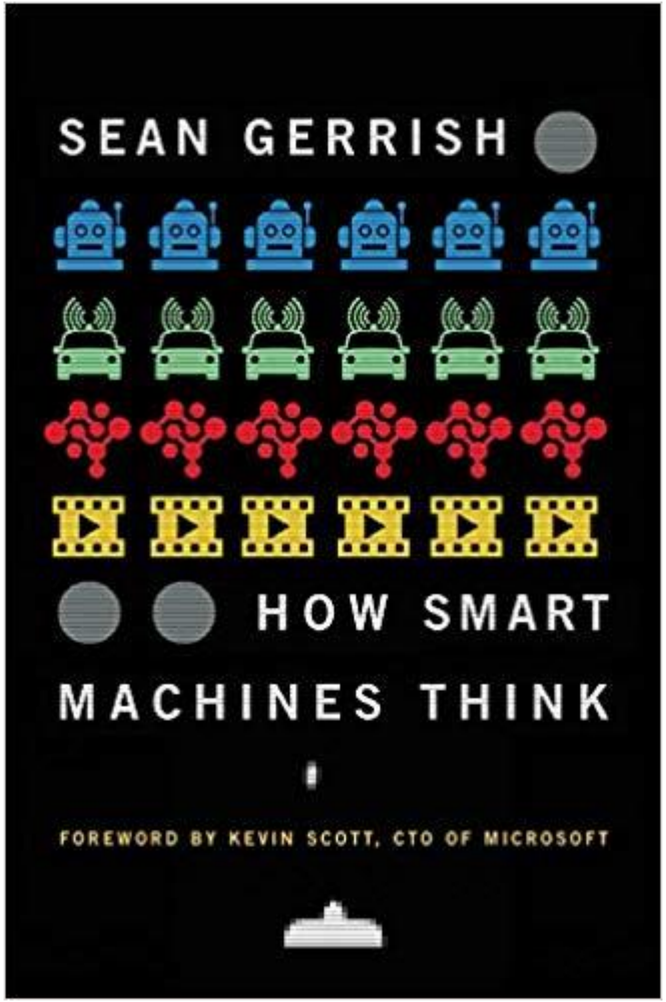


Deep Learning

Ian Goodfellow, Yoshua Bengio, Aaron Courville

ODTÜ METU
DEEP LEARNING INSTITUTE

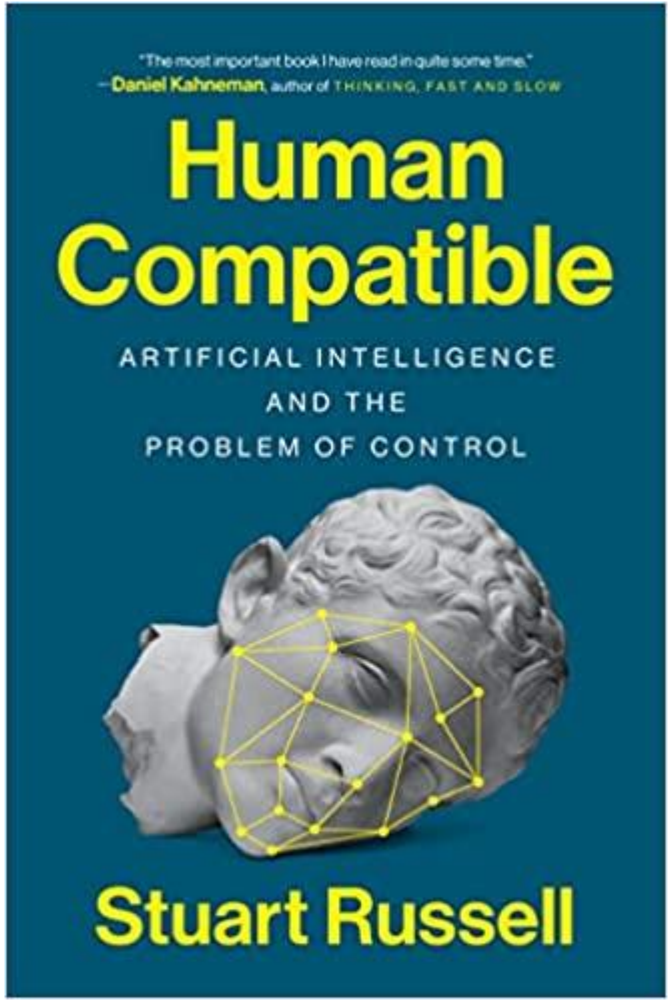
Reading Material



How Smart Machines Think (The MIT Press)
Hardcover – 6 Nov 2018

Sean Gerrish, Kevin Scott

Reading Material



Human Compatible: Artificial Intelligence and the Problem of Control

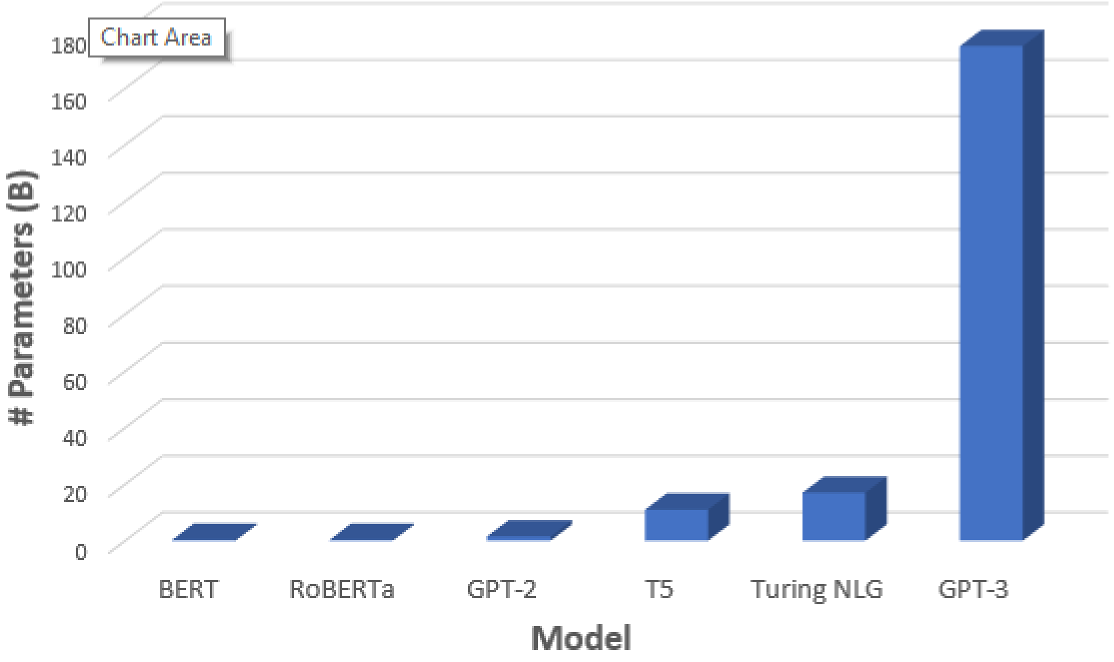
8 Oct 2019

Stuart Russell

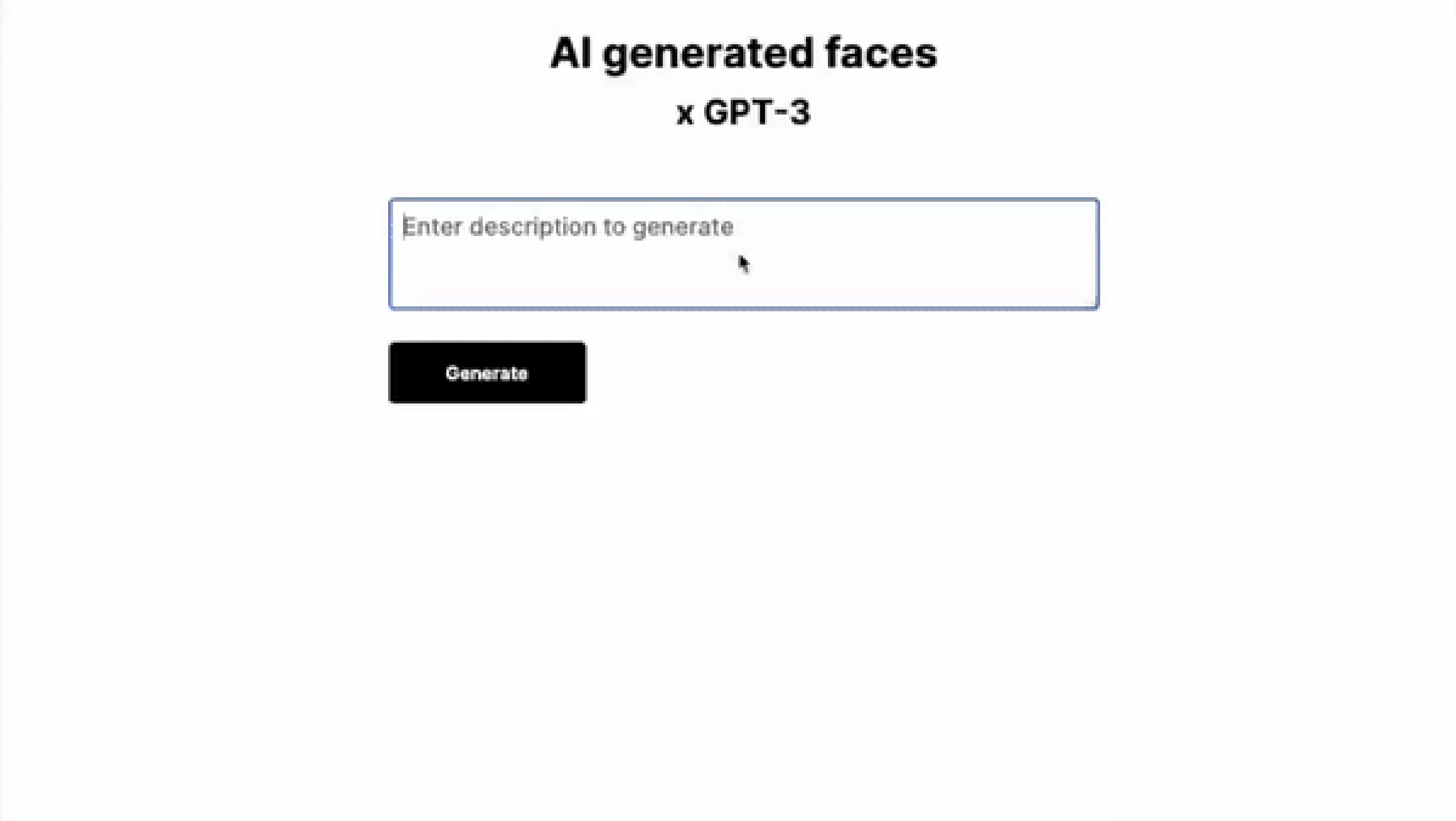
Deep Learning in NLP

Pre-trained language models trained on huge text corpus (in an unsupervised way) are later fine-tuned on specific tasks such as translation, question answering using much smaller task specific datasets.

GPT-3 is a task agnostic model, which needs limited examples to do well and achieve close to state of the art performance on a number of NLP tasks



Deep Learning in NLP



Source: https://www.auxiliary.tools/

IS 784 Deep Learning for Text Analytics

Natural language processing (NLP) concepts and NLP applications, deep learning methods for NLP, evaluation techniques, word embedding, Long-Short-Term Memory (LSTM) models, transformer models.

In addition, other application areas using the models initially proposed for NLP tasks will also be taught.

New Course Coming Next Semester!

MMI 7XX Machine Learning Systems Design and Deployment
Machine Learning (ML) Production Pipeline
Machine Learning Systems Design
Data Engineering
Model Development and Training
Scaling Up Training
Model Evaluation
Experiment Tracking and Versioning
Deployment
Deployment Platforms and Frameworks
TinyML
Integrating ML into Business
Future ML Systems