Athens Living Lab Ideathon

Machine Vision : Artificial Intelligence, Machine Learning & Computer Vision Challenges

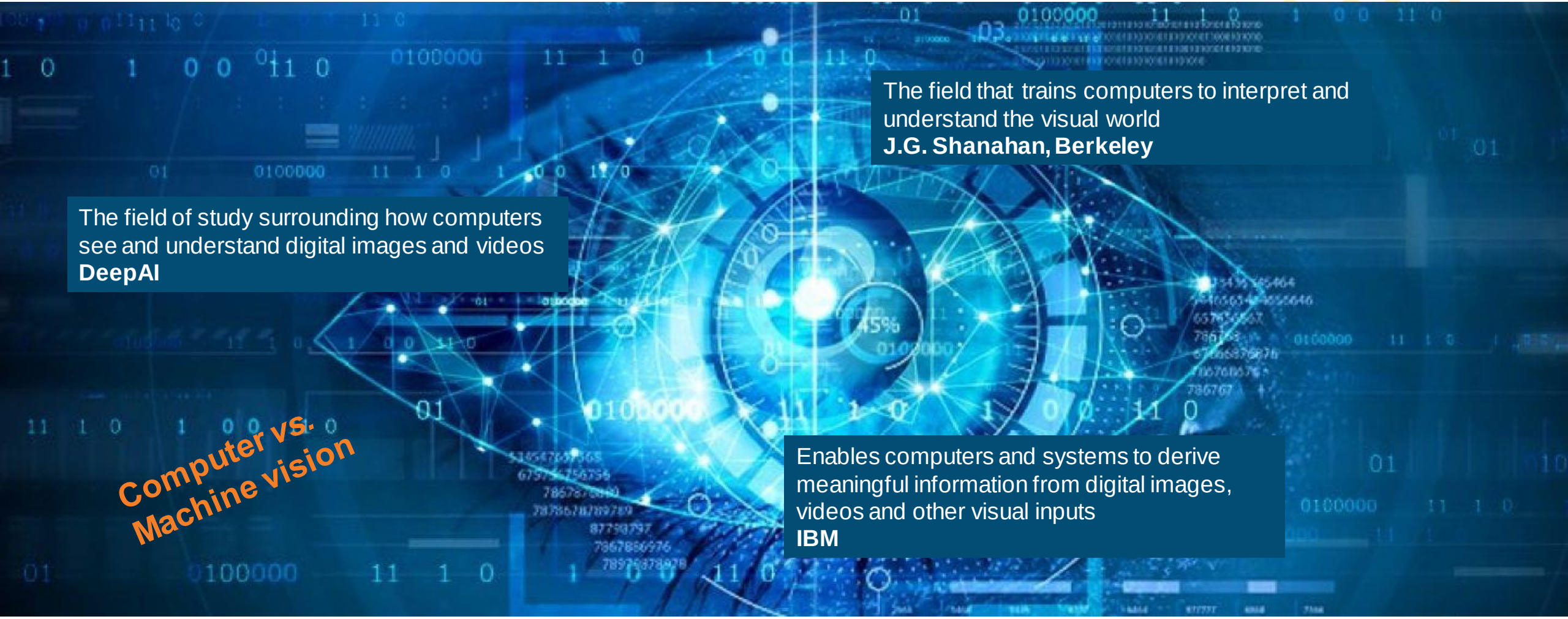
Institute of Communication and Computer Systems (**ICCS**)

10 October 2022

Georgios Drainakis
Software Engineer, Researcher
ISENSE Group / ICCS

5GLOGINNOV

Intro to Computer Vision



The field of study surrounding how computers see and understand digital images and videos
DeepAI

The field that trains computers to interpret and understand the visual world
J.G. Shanahan, Berkeley

Computer vs. Machine vision

Enables computers and systems to derive meaningful information from digital images, videos and other visual inputs
IBM

Computer Vision - Evolution

1959 — First digital image scanner transformed images into grids of numbers

Creation of perceptron — **1960s**

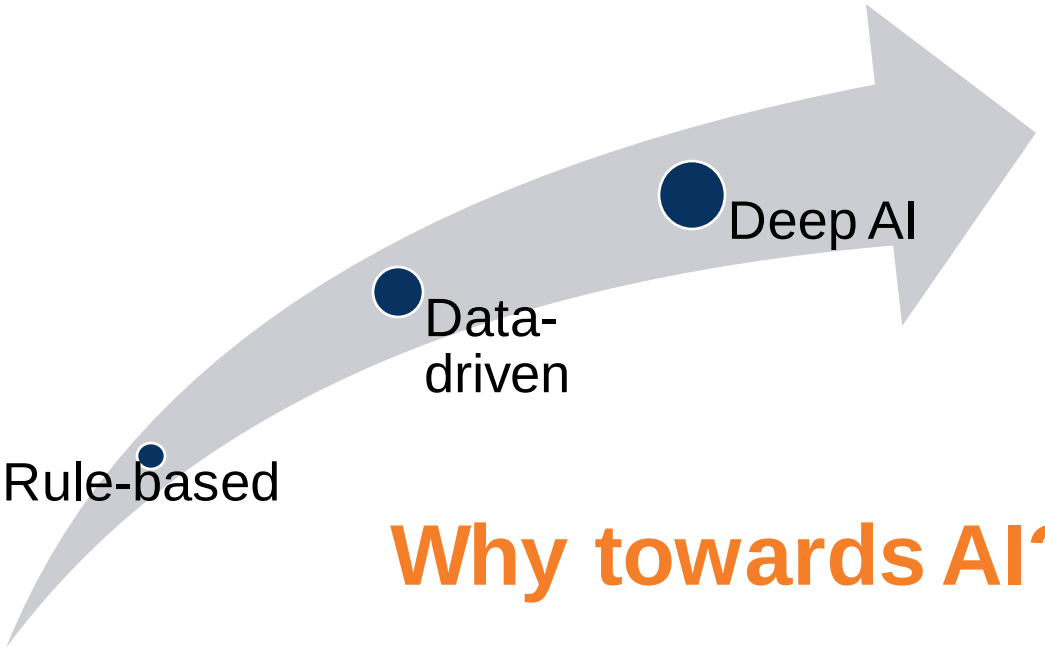
1970 — Back-propagation

Deep Learning — **2006**

2010 — Facebook facial recognition

Google Tensor Flow — **2015**

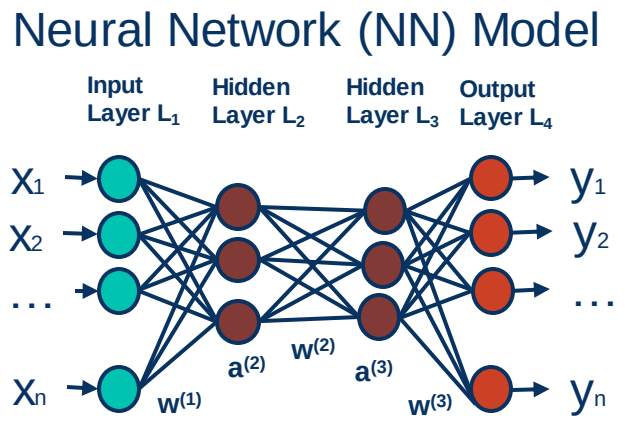
2020+ — CNNs

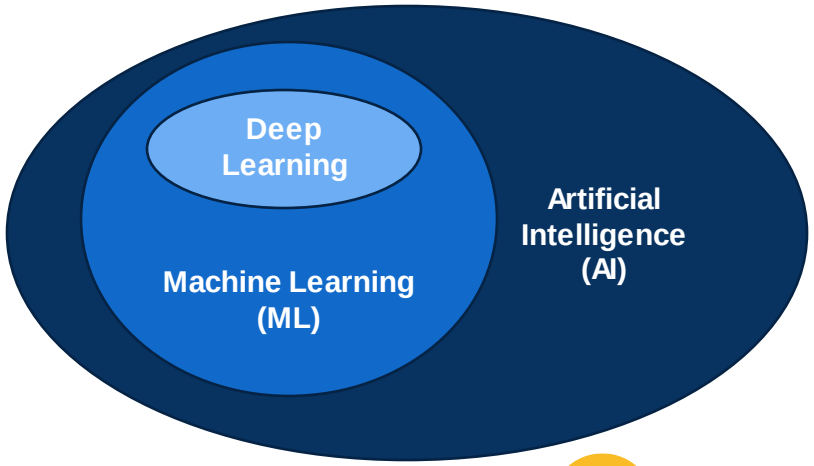


Why towards AI?

AI/ML simplified

- **Neurons**: mathematical functions
- **Weights**: connections between neurons (numerical)
- **Prediction**: $Y=X^TW$
- **Loss function**: Comparison metric between prediction and grown truth
- **Deep Learning**: Like simple ML but Model is chaotically big
- **Gradient Descent**: Optimization algorithm to find the minimum of the loss function
- **Training**: The process of feeding the ML model with data to allow the learning algorithm to “learn” data patterns (by adjusting the weights)
- **Inference**: The process of running new (live) data into a trained ML model and calculate output aka prediction

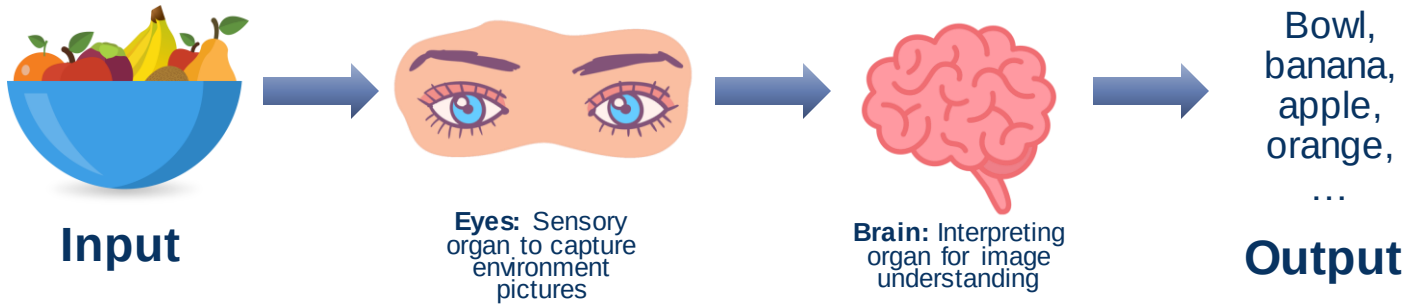




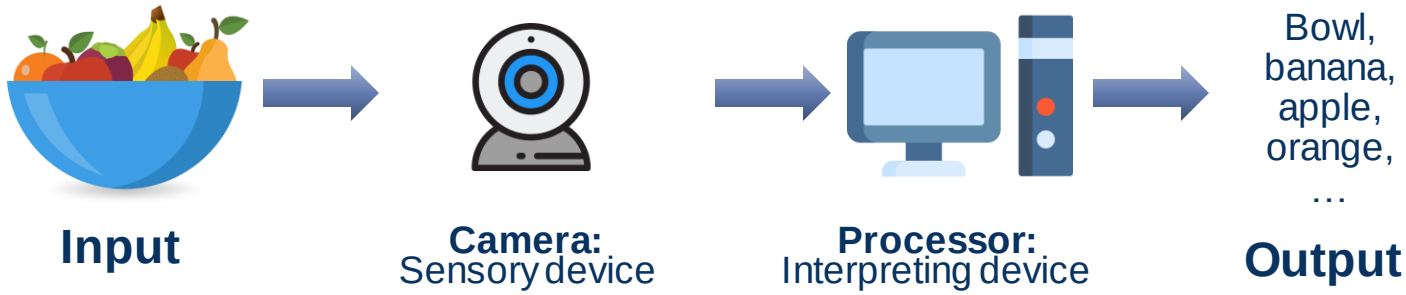
5GLOGINNOV

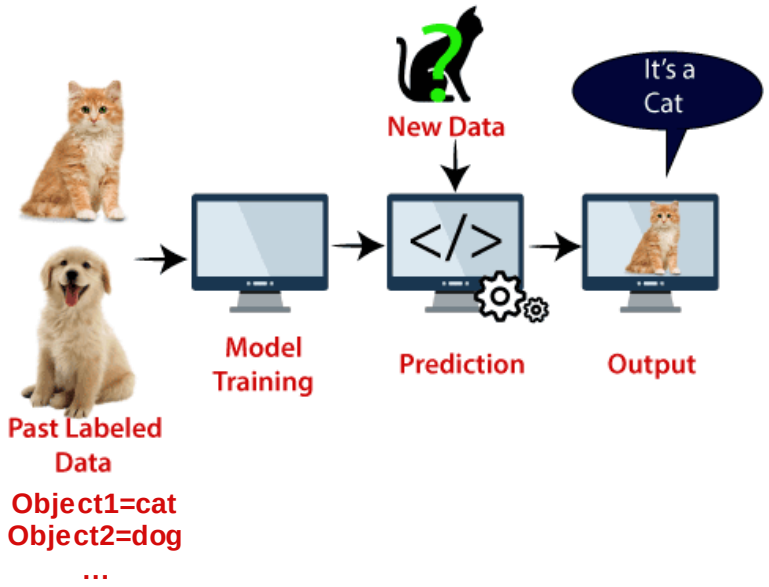
AI/ML & Computer Vision Convergence

Human vision system



Computer vision system





5GLOGINNOV

AI/ML Toolbox & Building Blocks

Software Frameworks

PyTorch

TensorFlow

pandas

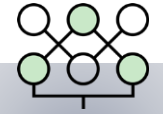
Keras

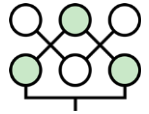
Hardware

Google TPU

NVIDIA Tesla V100

NVIDIA RTX

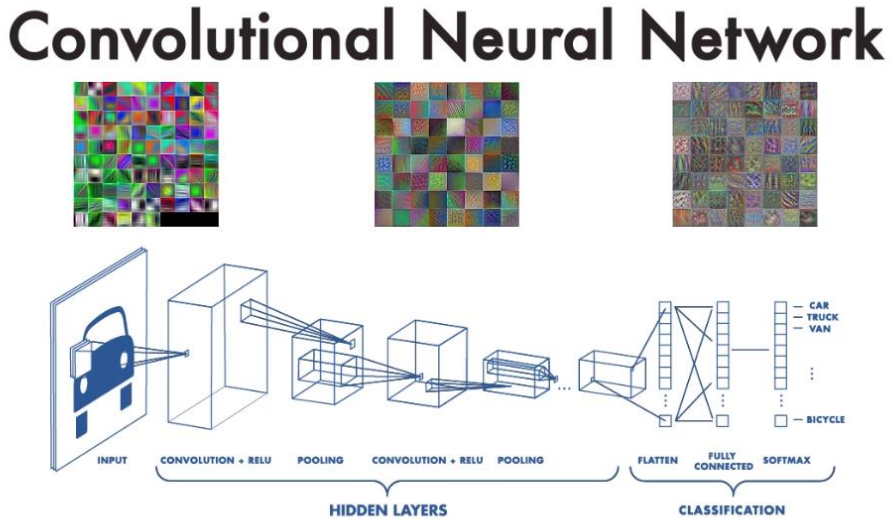




5GLOGINNOV

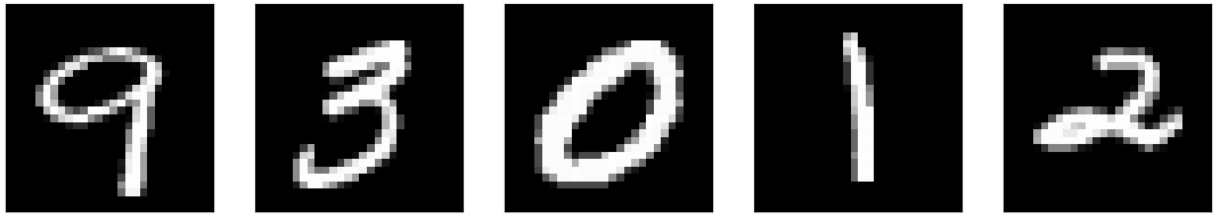
Demo: Digit Recognition (1)

- **Dataset: MNIST, Gray-scale digits 0-9**
- **ML Model: CNN (2 convolutional+2 linear layers)**
- **Tuning: Learning rate=0.001, Training epochs=10**
- **Image processing: Blurring factor (bf) of Gaussian Noise**

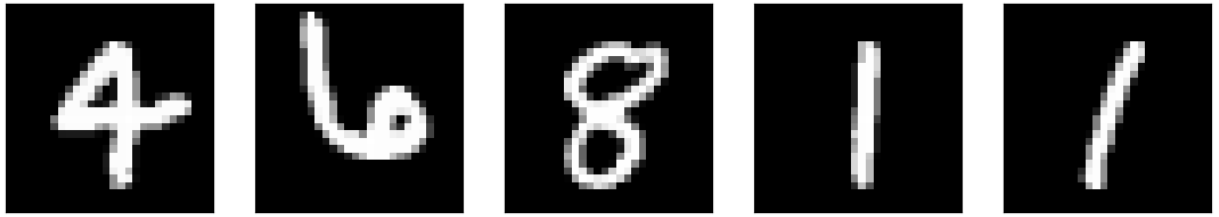


Demo: Digit Recognition (2)

bf=0



Truth	9	3	0	1	2
Prediction	9	3	0	1	2



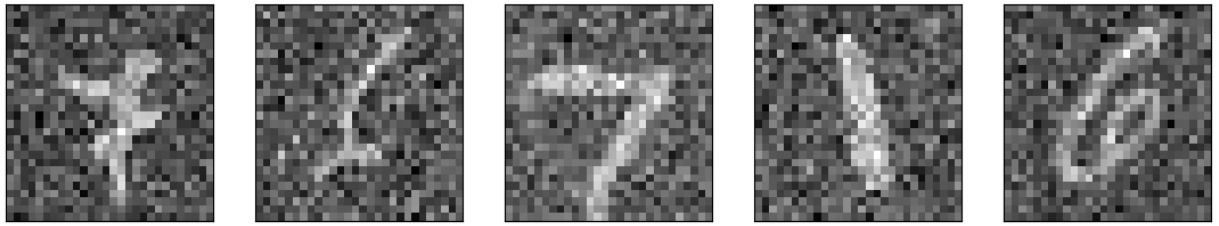
Truth	4	6	8	1	1
Prediction	4	6	8	1	1

NN score: 98%

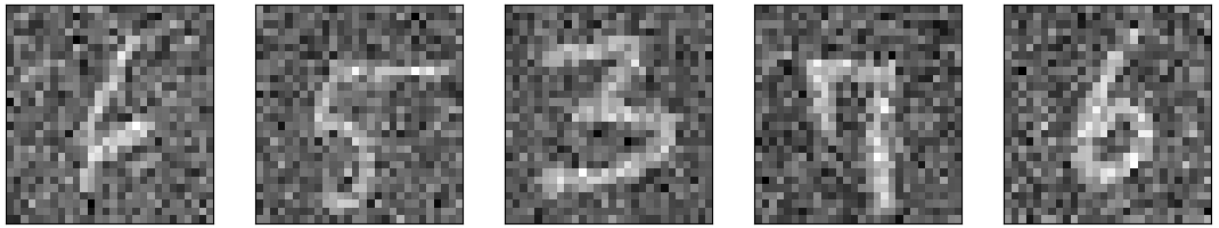
5GLOGINNOV

Demo: Digit Recognition (3)

bf=1



Truth	7	1	7	1	6
Prediction	7	1	7	1	6



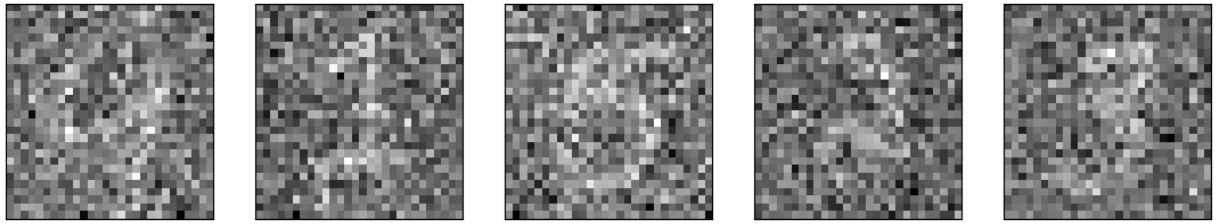
Truth	6	5	3	7	6
Prediction	6	5	3	7	6

NN score: 97%

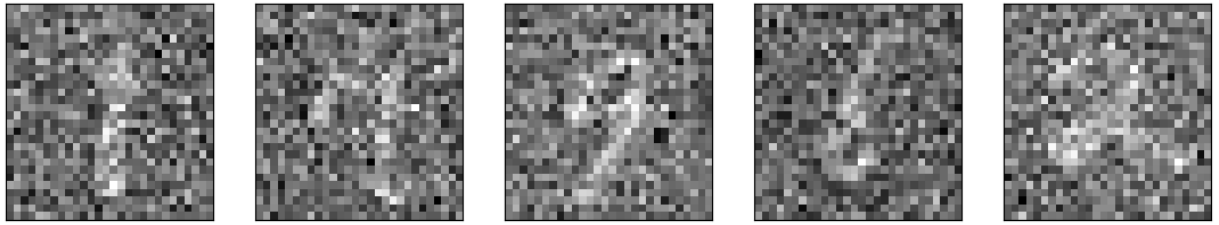
5GLOGINNOV

Demo: Digit Recognition (4)

bf=2



Truth	4	1	5	2	3
Prediction	9	1	5	2	3

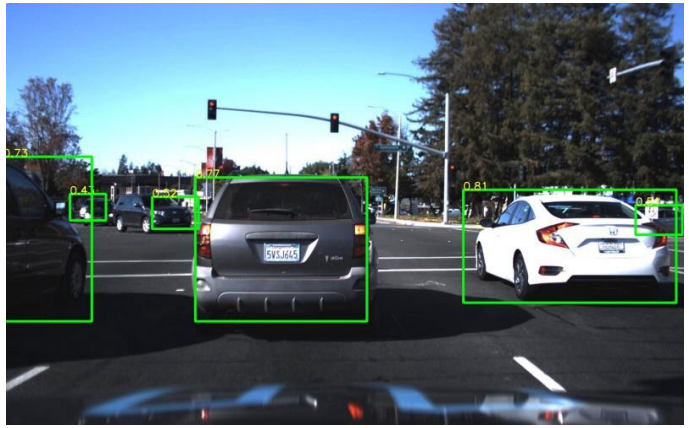


Truth	1	7	9	6	2
Prediction	1	7	9	6	2

NN score: 93.5%

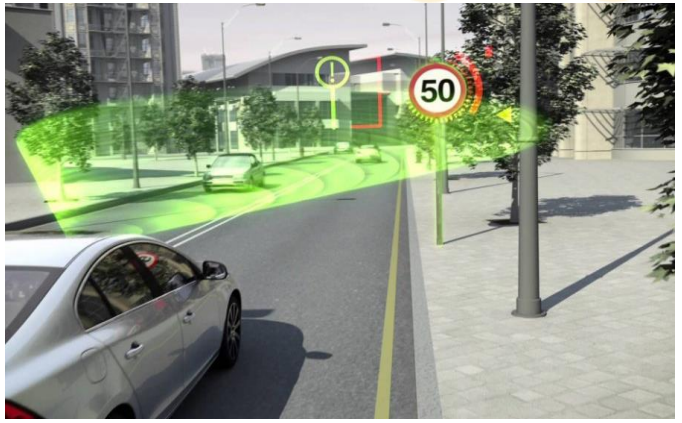
Can you beat it?

Computer vision apps: ITS



Vehicle & pedestrian detection

Lane detection

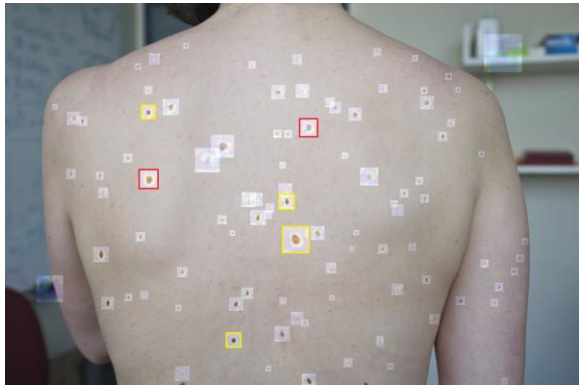


Traffic sign recognition

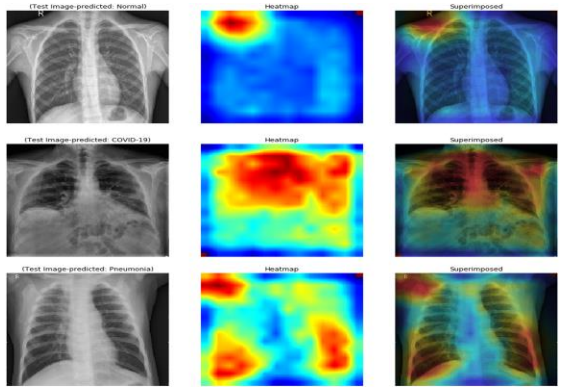
Key Requirements

- ✓ **Accuracy**
- ✓ **Minimal delays**
- ✓ **ML Model personalization**

Computer vision apps: Medical



Skin cancer detection



X-ray analysis e.g. COVID detection

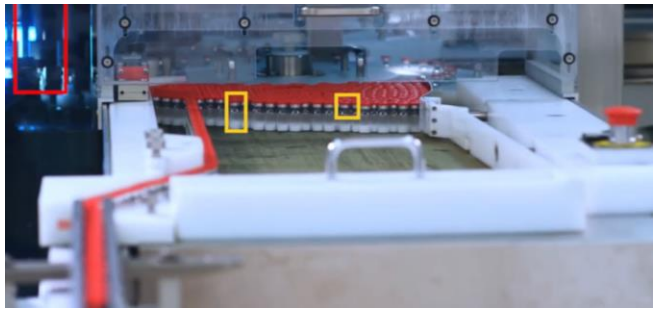
Surgical robotics

Key Requirements

- ✓ **Accuracy**
- ✓ **Precision & Recall i.e., no false-positives/negatives**
- ✓ **Data privacy**

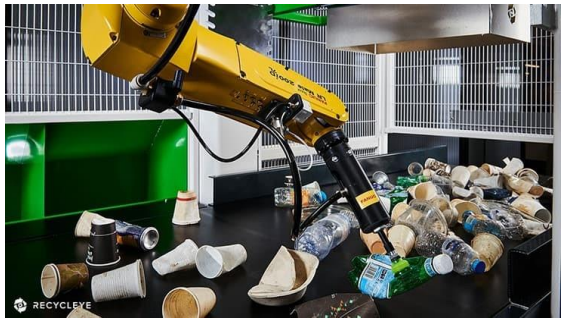
5GLOGINNOV

Computer vision apps: Industrial IoT



Defect detection

Inventory management



Material sorting

Key Requirements

- ✓ **Accuracy**
- ✓ **Minimal delays**
- ✓ **Processing-light ML tasks (if running on embedded devices)**
- ✓ **Energy-aware ML**

5GLOGINNOV

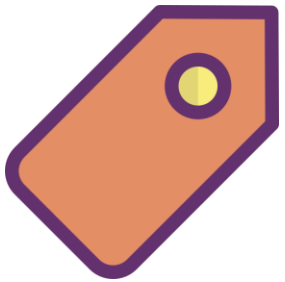
Computer vision challenges

DATA SHORTAGE

- Inadequate data quantity
- Restrictions due to privacy
- Low-quality data e.g. shades

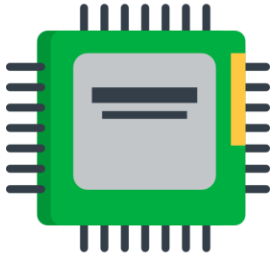
BLACK-BOX AI

- No existing theory on AI/ML
- No convergence guarantees
- Works only by trial and error (result-oriented)
- Based on existing practical guidelines/rules



MANUAL LABELLING

- Requires human interference
- Long process
- Prone to errors



PROCESSING CAPACITY

- Heavy ML models for HD images, videos, etc. usually require GPUs
- Most ML tasks can be only performed in High Availability Data Centers

OVER-FITTING

- ML model is trained but cannot generalize over new data
- Requires analysis of ML algorithm, parameters, etc.

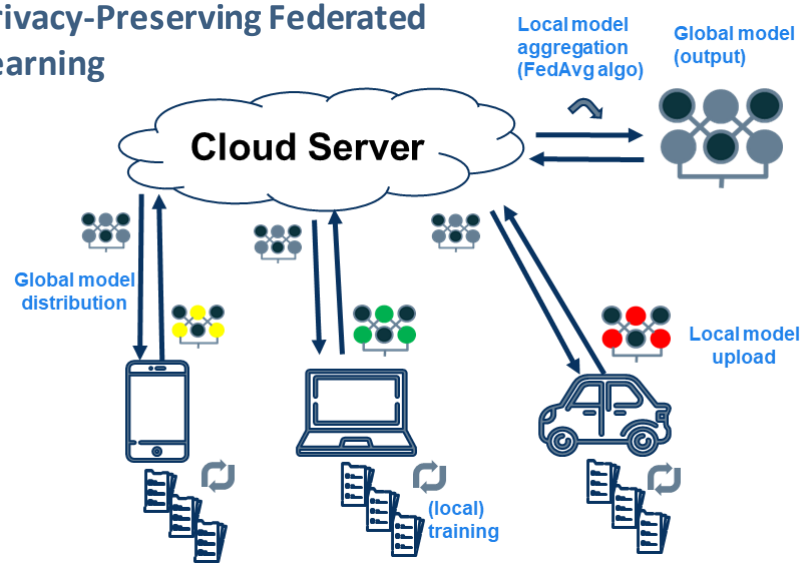
ENERGY COSTS

- GPU operations inflict major energy expenditure

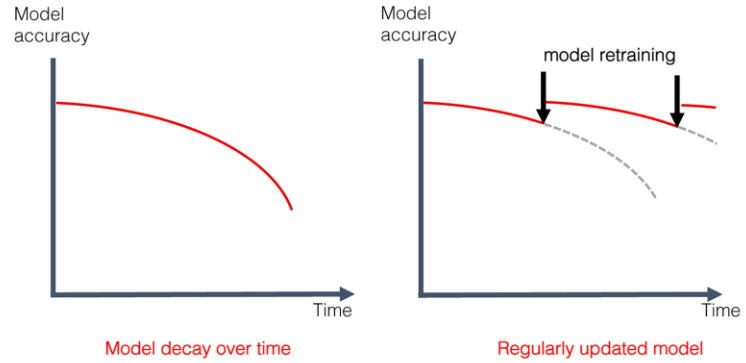
5GLOGINNOV

Future Directions: a glimpse

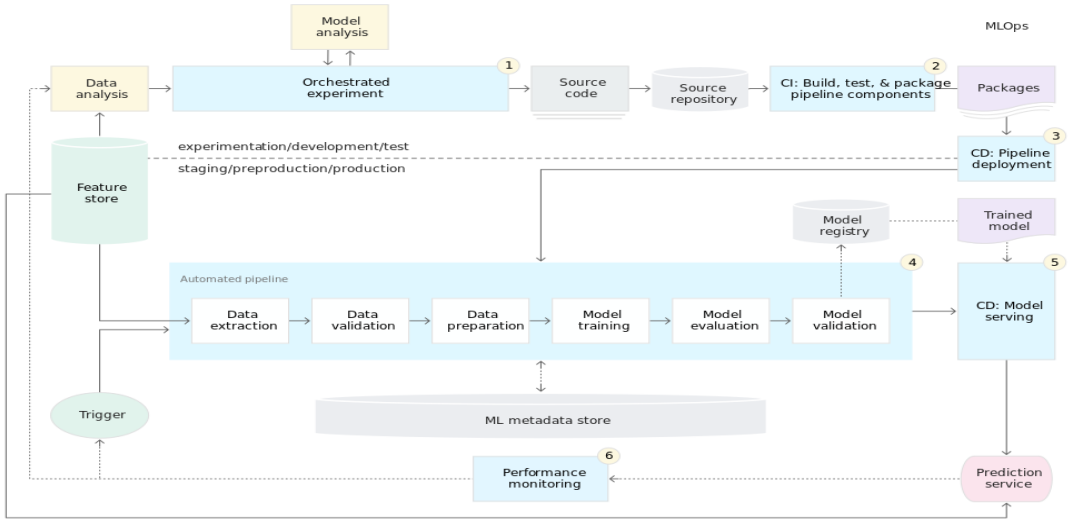
Privacy-Preserving Federated Learning



Continuous Learning



MLops



Thank you!

Questions?

5GLOGINNOV

Co-financed by the European Union
Connecting Europe Facility