tinyML® Talks

Enabling Ultra-low Power Machine Learning at the Edge

"AI at the Edge: Enabling Vision for low-power devices"

Anna Petrovicheva - OpenCV.AI

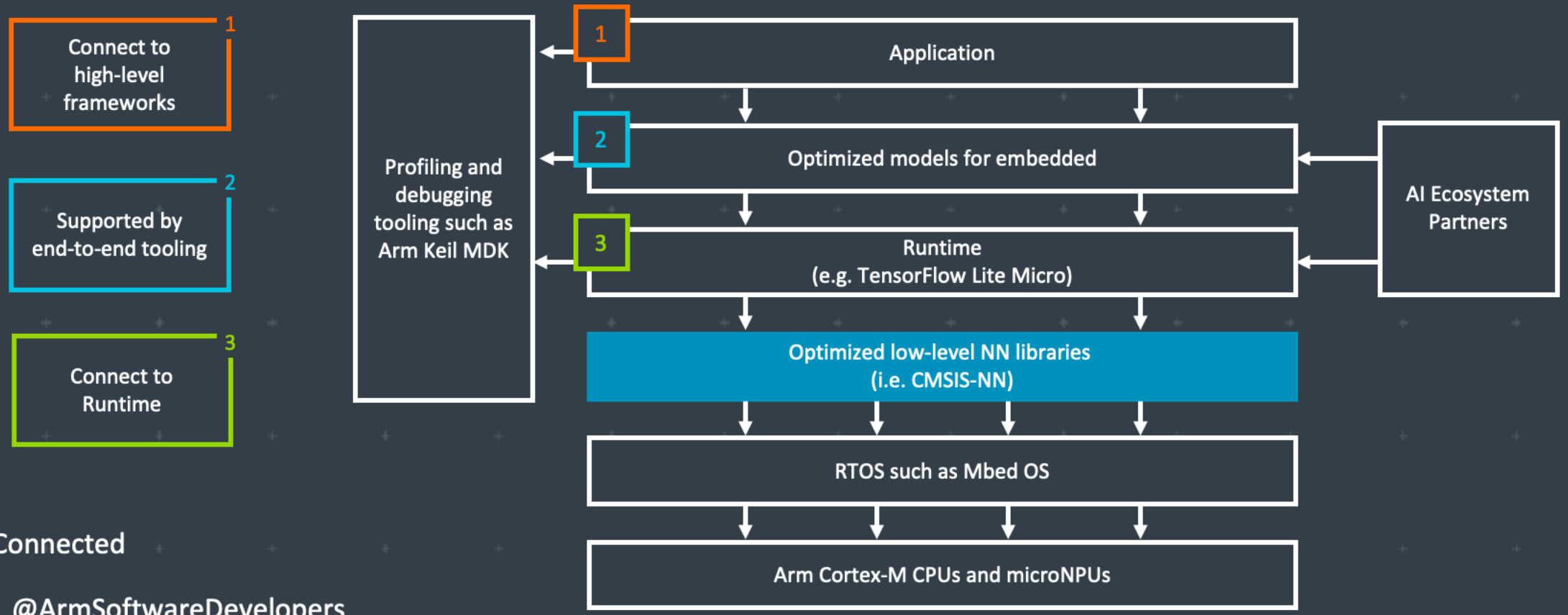
November 25, 2021

www.tinyML.org

tinyML Talks Strategic Partners

Additional Sponsorships available – contact Olga@tinyML.org for info

Arm: The Software and Hardware Foundation for tinyML



Stay Connected

@ArmSoftwareDevelopers

@ArmSoftwareDev

Resources: developer.arm.com/solutions/machine-learning-on-arm

arm

Deeplite

WE USE AI TO MAKE OTHER AI FASTER, SMALLER AND MORE POWER EFFICIENT

Automatically compress SOTA models like MobileNet to <200KB with **little to no drop in accuracy** for inference on resource-limited MCUs

Reduce model optimization trial & error from weeks to days using Deeplite's **design space exploration**

Deploy more models to your device without sacrificing performance or battery life with our **easy-to-use software**

BECOME BETA USER **bit.ly/testdeeplite**

mobilityXlab arm AI 100 CBINSIGHTS

TinyML for all developers

C++ library

Arduino library

WebAssembly

Dataset

Acquire valuable training data securely

Enrich data and train ML algorithms

Edge Device

Real sensors in real time

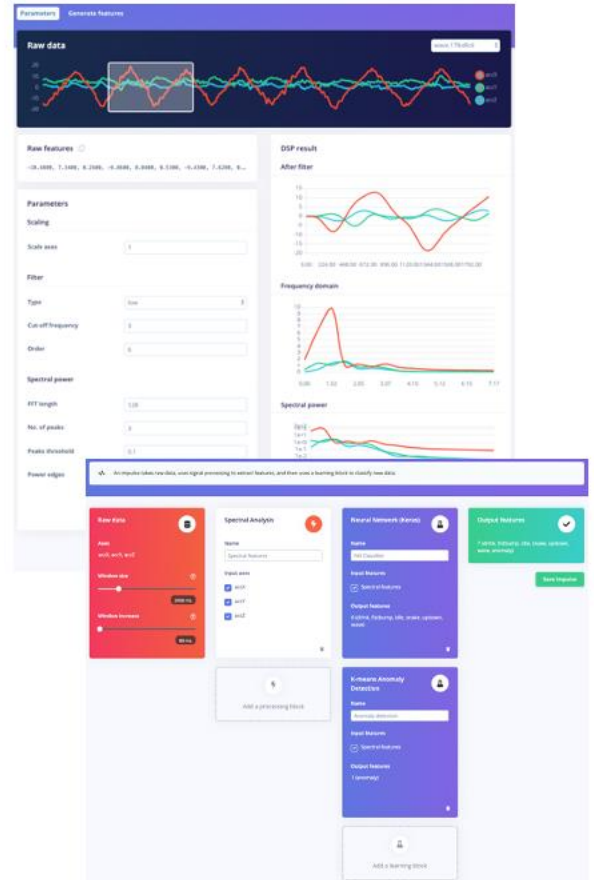
Open source SDK

Impulse

Embedded and edge compute deployment options

Test impulse with real-time device data flows

Test

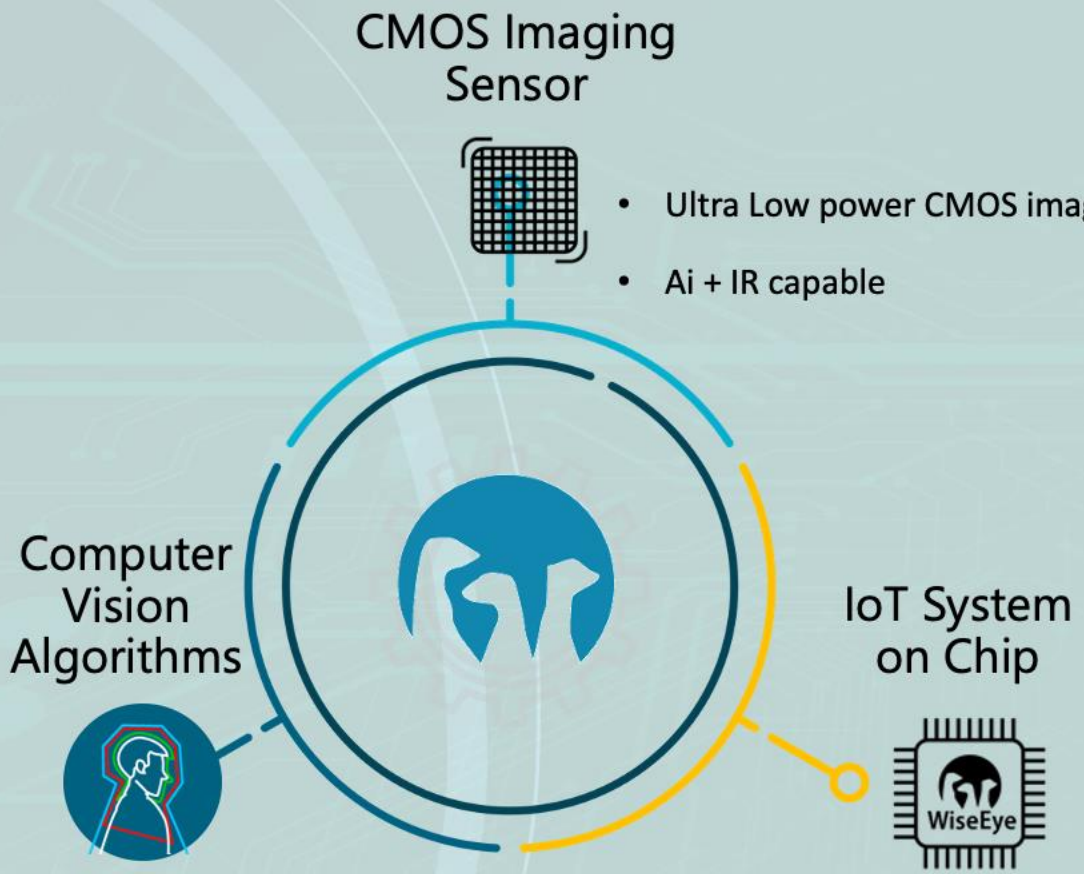


www.edgeimpulse.com

emza
visual sense

The Eye in IoT

Edge AI Visual Sensors



- Ultra Low power CMOS imager
- Ai + IR capable

- Machine Learning algorithm
- <1MB memory footprint
- Microcontrollers computing power
- Trained algorithm
- Processing of low-res images
- Human detection and other classifiers

- Machine Learning edge computing silicon
- <1mW always-on power consumption
- Computer Vision hardware accelerators

info@emza-vs.com

Enabling the next generation of Sensor and Hearable products to process rich data with energy efficiency

Visible Image

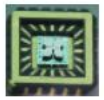
Sound

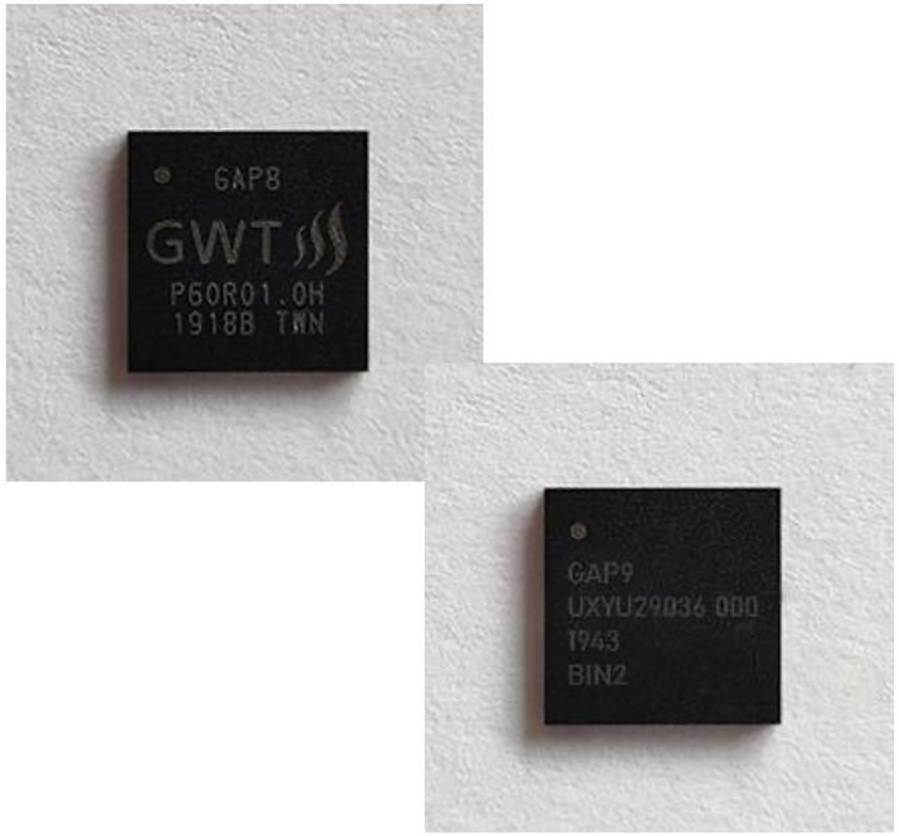
IR Image

Radar

Bio-sensor

Gyro/Accel

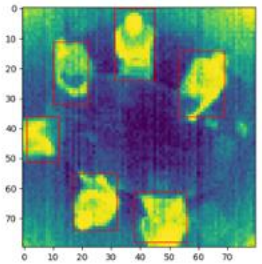




Wearables / Hearables

Battery-powered consumer electronics

IoT Sensors



GREENWAVES TECHNOLOGIES

Grovety Inc.

SOFTWARE DEVELOPMENT SERVICES FOR TINYML SOLUTIONS

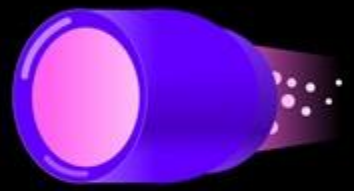
1 **Development tools**
SDK, IDE, compilers, leveraging on TVM, uTVM & LLVM

2 **Firmware**
Drivers, BSP, protocols, etc.

arm AI PARTNER

Distributed infrastructure for TinyML apps

HOTG
Decoupling intelligence



Develop at warp speed

Automate deployments

Device orchestration

HOTG is building the **distributed infrastructure** to pave the way for **AI** enabled **edge applications**

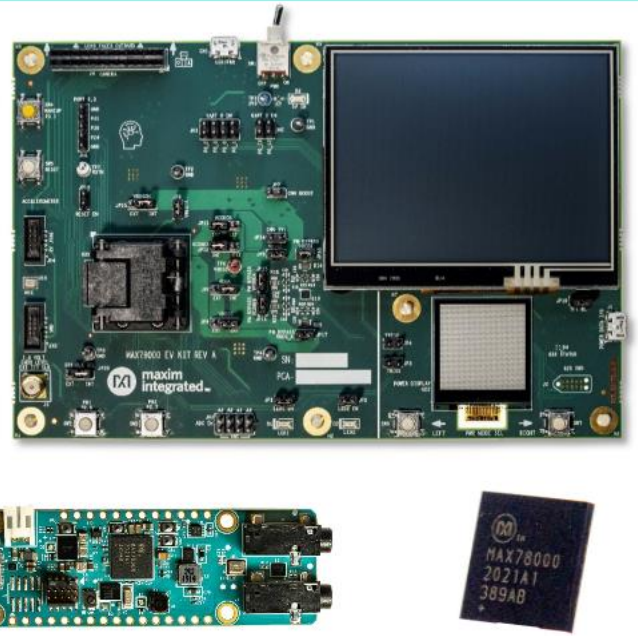
LatentAI

Adaptive AI for the Intelligent Edge

Latentai.com

Maxim Integrated: Enabling Edge Intelligence

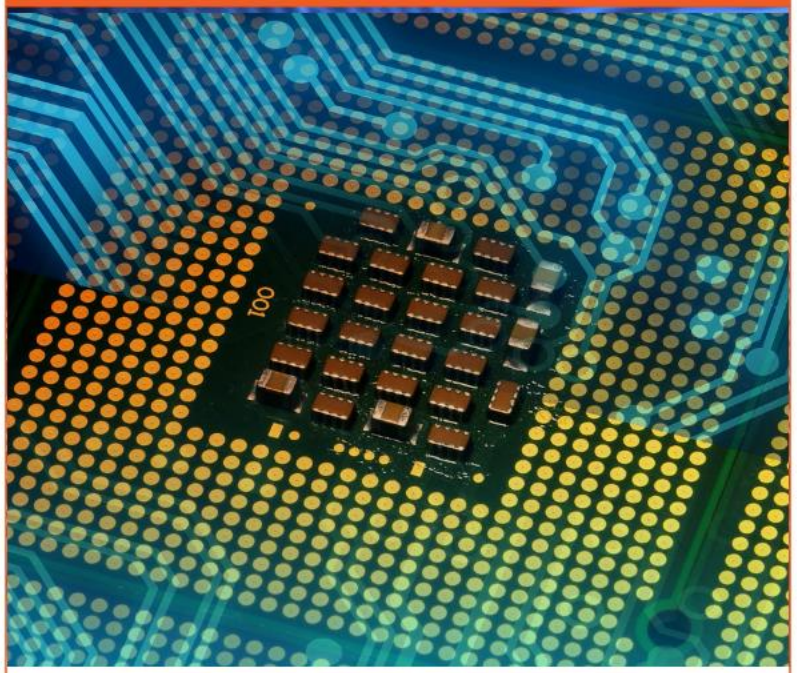
Advanced AI Acceleration IC



The new MAX78000 implements AI inferences at low energy levels, enabling complex audio and video inferencing to run on small batteries. Now the edge can see and hear like never before.

www.maximintegrated.com/MAX78000

Low Power Cortex M4 Micros



Large (3MB flash + 1MB SRAM) and small (256KB flash + 96KB SRAM, 1.6mm x 1.6mm) Cortex M4 microcontrollers enable algorithms and neural networks to run at wearable power levels.

www.maximintegrated.com/microcontrollers

Sensors and Signal Conditioning

Health sensors measure PPG and ECG signals critical to understanding vital signs. Signal chain products enable measuring even the most sensitive signals.

www.maximintegrated.com/sensors

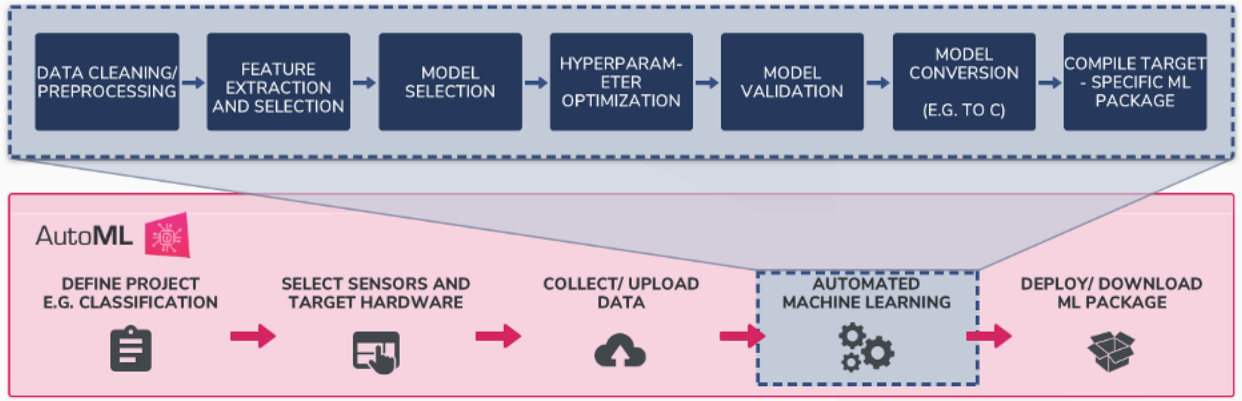
Qeexo AutoML

Automated Machine Learning Platform that builds tinyML solutions for the Edge using sensor data

Key Features

- Supports 17 ML methods:
  - Multi-class algorithms: GBM, XGBoost, Random Forest, Logistic Regression, Gaussian Naive Bayes, Decision Tree, Polynomial SVM, RBF SVM, SVM, CNN, RNN, CRNN, ANN
  - Single-class algorithms: Local Outlier Factor, One Class SVM, One Class Random Forest, Isolation Forest
- Labels, records, validates, and visualizes time-series sensor data
- On-device inference optimized for low latency, low power consumption, and small memory footprint applications
- Supports Arm® Cortex™- M0 to M4 class MCUs

End-to-End Machine Learning Platform



For more information, visit: www.qeexo.com

Target Markets/Applications

- Industrial Predictive Maintenance
- Smart Home
- Wearables
- Automotive
- Mobile
- IoT

Qualcomm
AI research

Advancing AI research to make efficient AI ubiquitous

Power efficiency

Model design, compression, quantization, algorithms, efficient hardware, software tool

Personalization

Continuous learning, contextual, always-on, privacy-preserved, distributed learning

Efficient learning

Robust learning through minimal data, unsupervised learning, on-device learning

A platform to scale AI across the industry

Perception

Object detection, speech recognition, contextual fusion

Reasoning

Scene understanding, language understanding, behavior prediction

Action

Reinforcement learning for decision making

Edge cloud

Cloud

IoT/IIoT

Automotive

Mobile

Reality AI®

Add Advanced Sensing to your Product with Edge AI / TinyML

https://reality.ai info@reality.ai @SensorAI Reality AI

Pre-built Edge AI sensing modules, plus tools to build your own

Reality AI solutions

Prebuilt sound recognition models for indoor and outdoor use cases

Solution for industrial anomaly detection

Pre-built automotive solution that lets cars "see with sound"

Reality AI Tools® software

Build prototypes, then turn them into real products

Explain ML models and relate the function to the physics

Optimize the hardware, including sensor selection and placement

SensiML™

Build Smart IoT Sensor Devices From Data

SensiML pioneered TinyML software tools that auto generate AI code for the intelligent edge.

- End-to-end AI workflow
- Multi-user auto-labeling of time-series data
- Code transparency and customization at each step in the pipeline

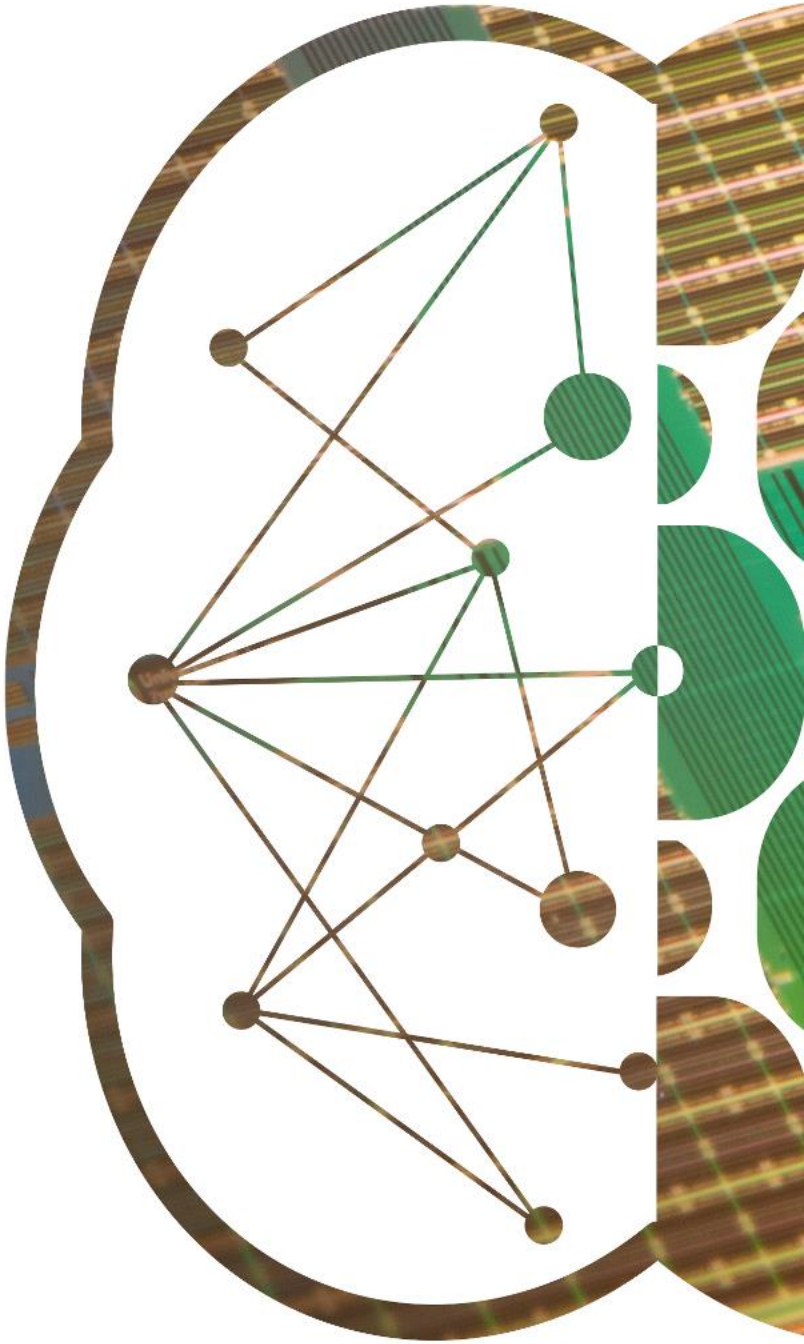
We enable the creation of production-grade smart sensor devices.

sensiml.com

SynSense

SynSense builds **sensing and inference** hardware for **ultra-low-power** (sub-mW) **embedded, mobile and edge** devices. We design systems for **real-time always-on smart sensing**, for audio, vision, IMUs, bio-signals and more.

https://SynSense.ai

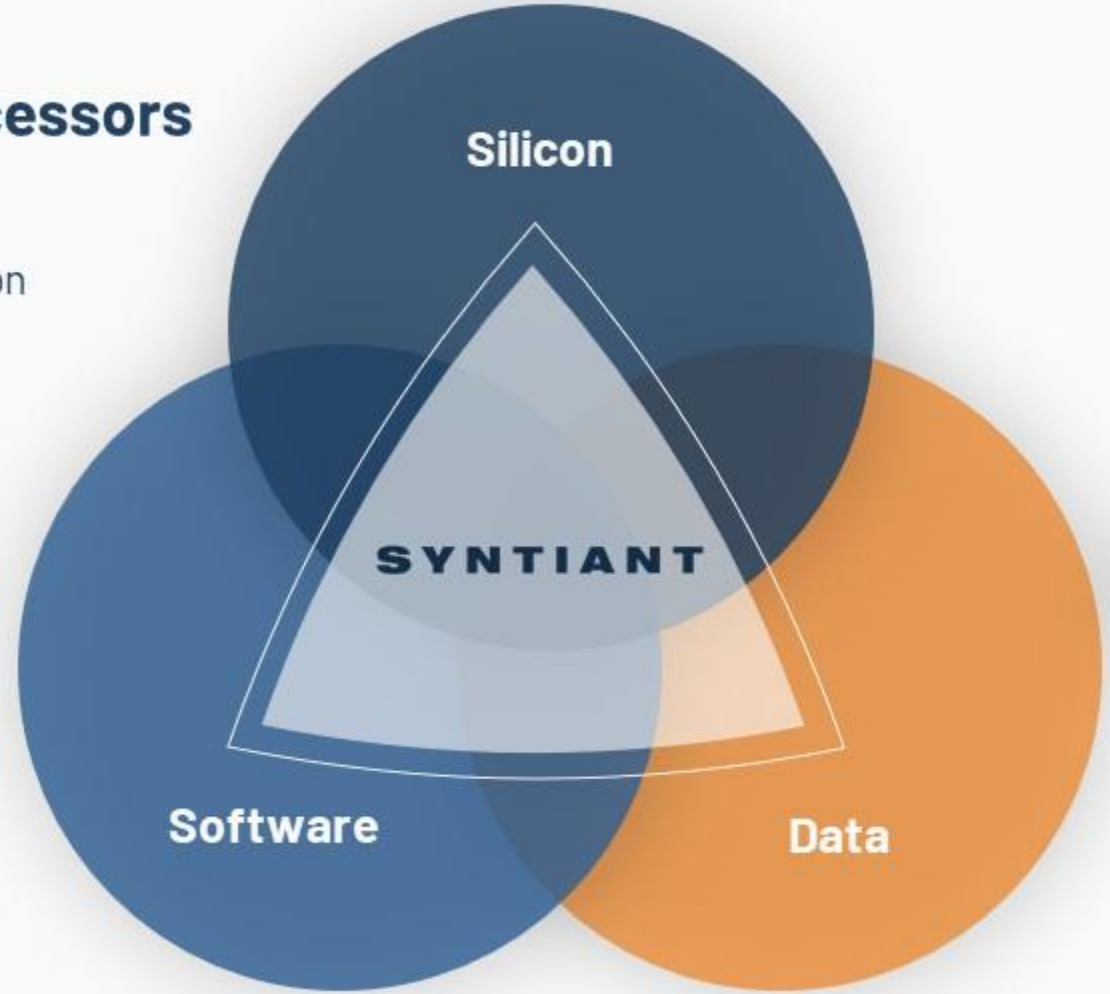


SYNTIANT

End-to-End Deep Learning Solutions for TinyML & Edge AI

Neural Decision Processors

- At-Memory Compute
- Sustained High MAC Utilization
- Native Neural Network Processing



ML Training Pipeline

- Enables Production Quality Deep Learning Deployments

Data Platform

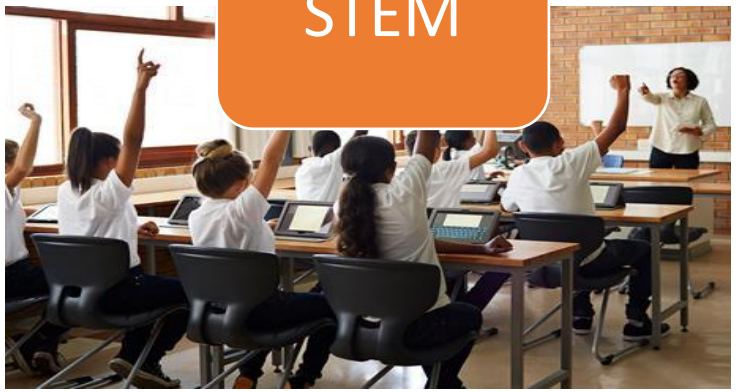
- Reduces Data Collection Time and Cost
- Increases Model Performance

partners@syntiant.com

www.syntiant.com

tinyML for Good – Workshop, November 17th(7 am PDT)

STEM



TINY ML For Good

Healthcare

Earth
Climate
Conservation

Contact: 4good@tinyML.org

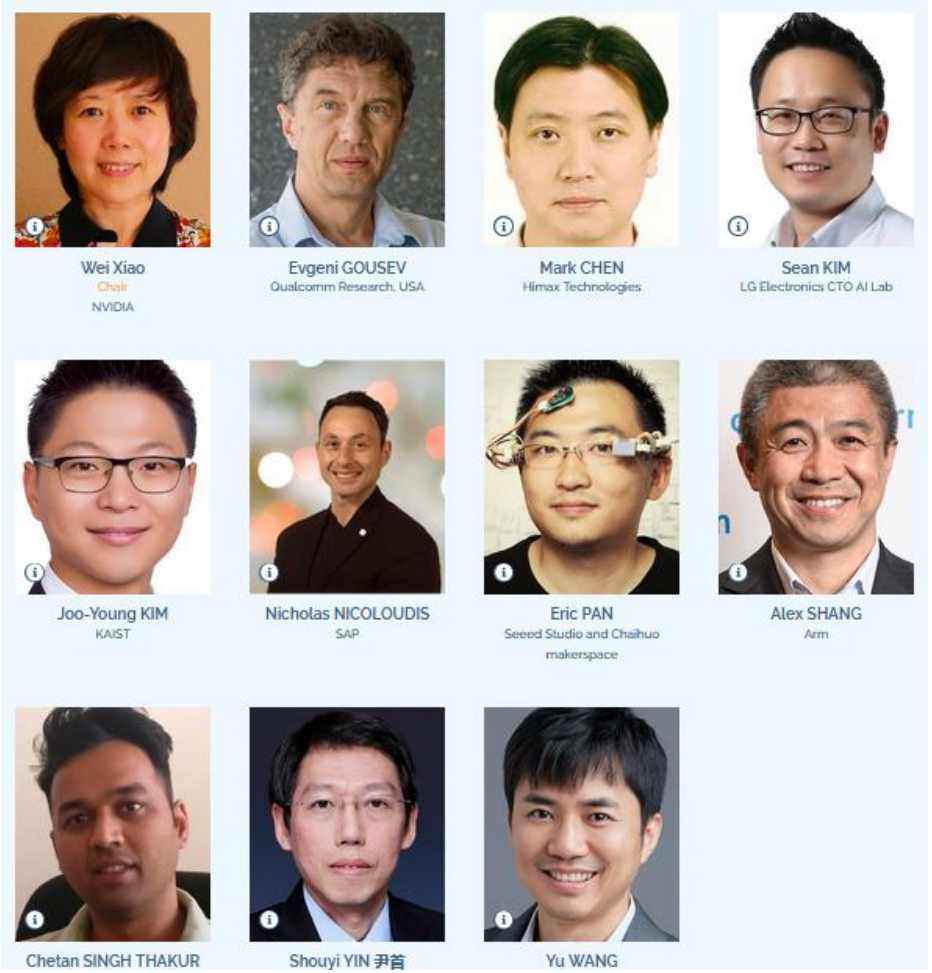
TINY ML ASIA

LIVE ONLINE **November 2-5, 2021**

(9-11:30 am China Standard time)

https://www.tinyml.org/event/asia-2021/

Technical Program Committee



Register today!

Free event courtesy of our sponsors and strategic partners

More sponsorships are available: sponsorships@tinyML.org

TINY ML VISION Challenge

THE 2021 WINNERS ARE

collaboration with hackster.io AN AVNET COMMUNITY

Team Sol

RANKED WINNERS: 1ST PLACE

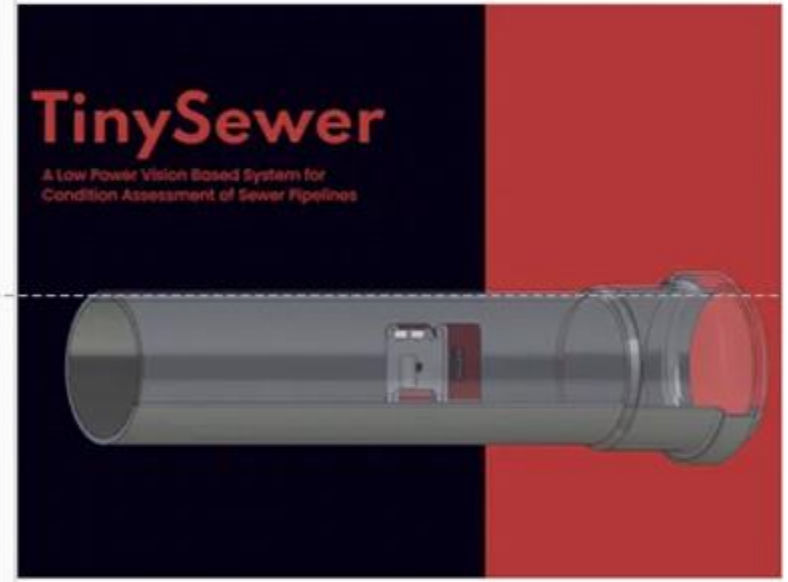
TinyML Aerial Forest Fire Detection

TheBlue Phoenix

RANKED WINNERS: 2ND PLACE

WorkSafe: Computer Vision based multiparameter monitor with

Huy Mai



RANKED WINNERS: 3RD PLACE

TinySewer - Low Power Sewer Faults Detection System

Honorable mention prize winners:

[Flat Tire Detection Using Machine Vision](#) by [Bob Hammell](#)

[Smart Bird Feeder](#) by **Ariela, Anna, Audrey, Nathan, Tianlang, Haoming, Eric, Edward and Tera** Guided by: [Chen Feng](#)

More details: tinyml.org/news/tinyml-vision-challenge-winners

arm EDGE IMPULSE Himax intel LATTICE SEMICONDUCTOR LUXonjs PixArt Imaging Inc. Qualcomm SynSense

Next tinyML Talks

Date	Presenter	Topic / Title
Wednesday, December 1	Pierre Gembaczka, Program manager, Fraunhofer IMS	AIfES - an open-source standalone AI framework for almost any hardware

Webcast start time is 9:30 am Pacific time

Please contact talks@tinyml.org if you are interested in presenting

TINY ML TALKS webcast

Reminders

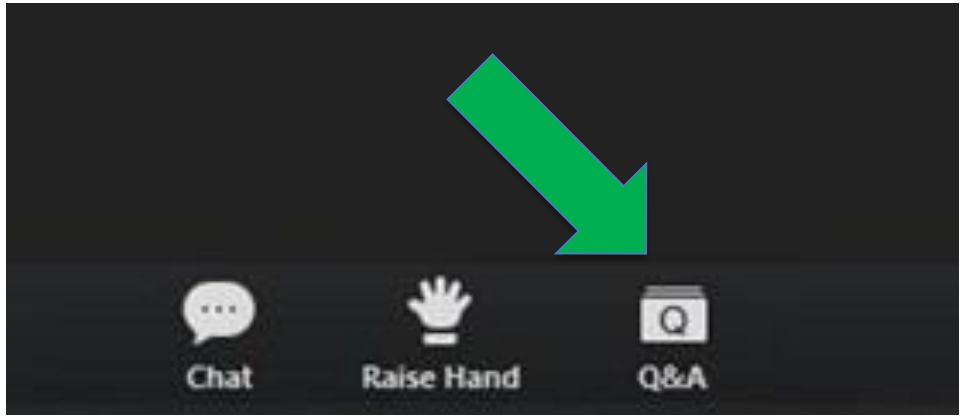
Slides & Videos will be posted tomorrow

PDF

tinyml.org/forums

youtube.com/tinyml

Please use the Q&A window for your questions



TINY ML TALKS webcast

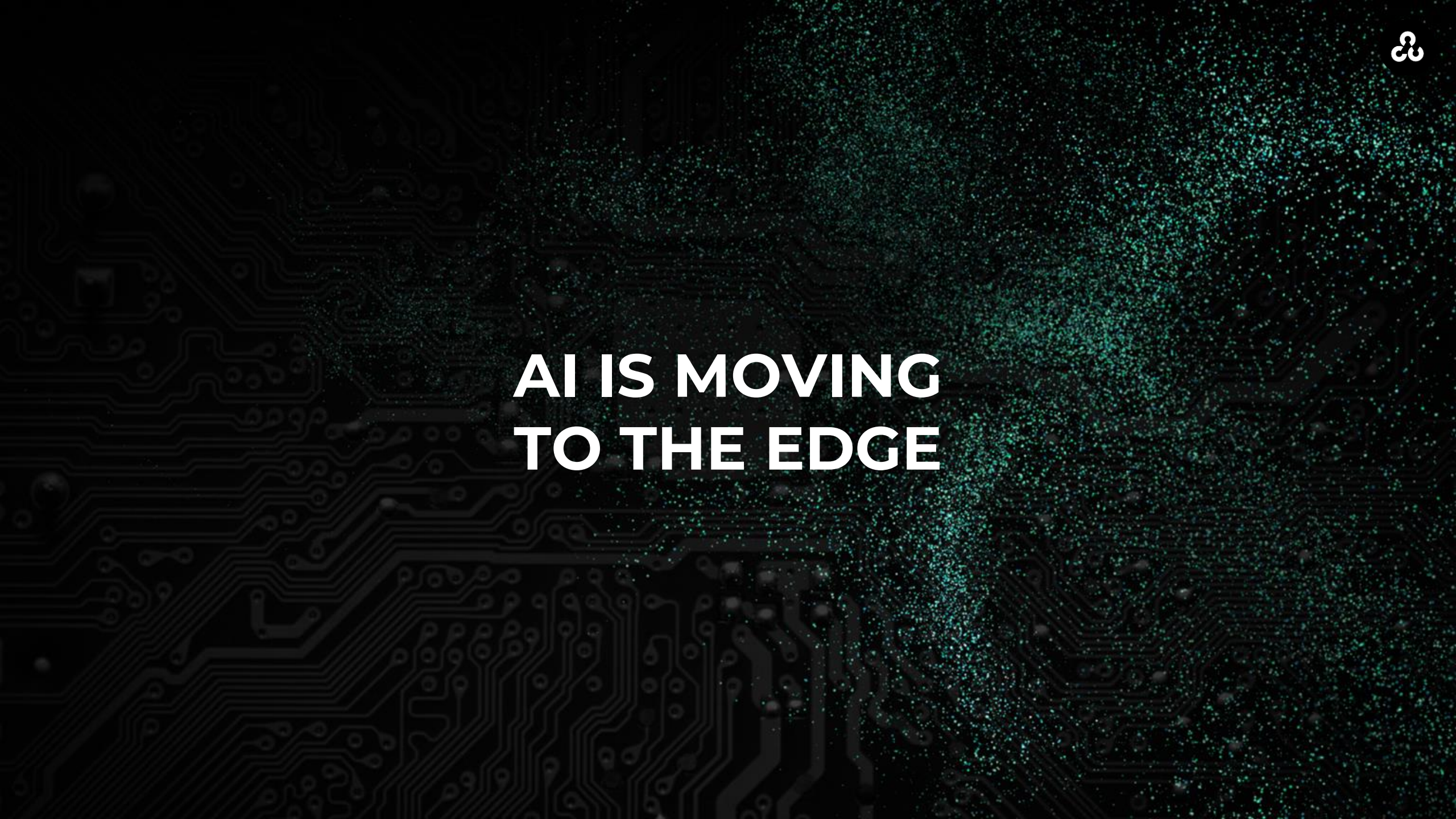
Anna Petrovicheva

Anna has been in the Computer Vision industry for 11 years now: first as an engineer, then as a technical lead. Her team works on solutions for a wide range of applications: from medicine and diagnostics to sports and education. The projects she lead were highlighted at Apple Event, in CNBC, and Wall Street Journal. Anna is a member of the OpenCV Foundation.

AI ON EDGE DEVICES

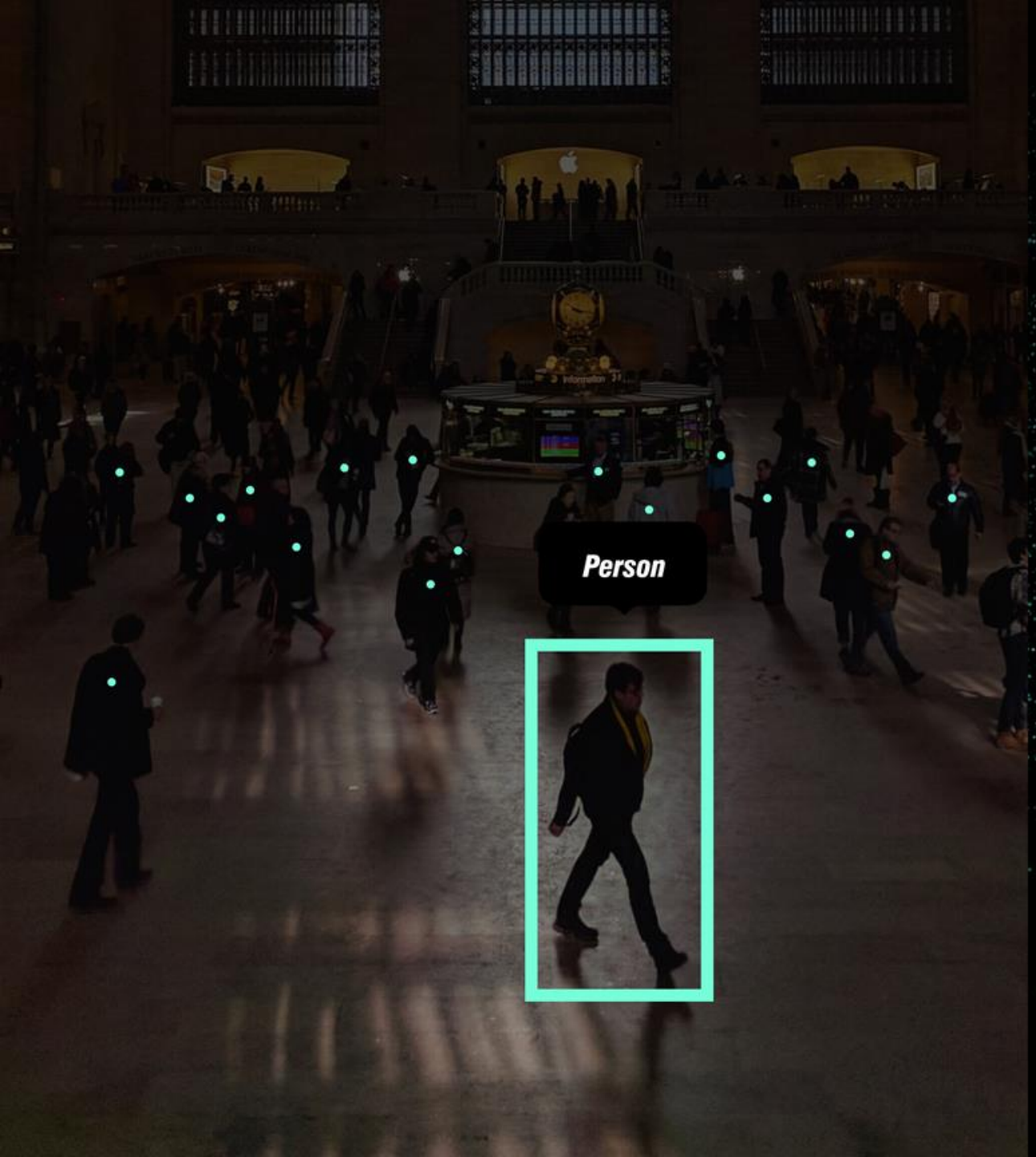
Anna Petrovicheva
CEO, Xperience.AI
Board Member, OpenCV.org


ANNA
PETROVICHEVA
intel
itseez
— Vision that works!
xperience/ai


AI IS MOVING
TO THE EDGE


INTERNET
TRAFFIC


SERVER
COSTS



SECURITY & PRIVACY

THE LOOP OF THE INDUSTRY

2010 - 2014: Classic CV on the edge

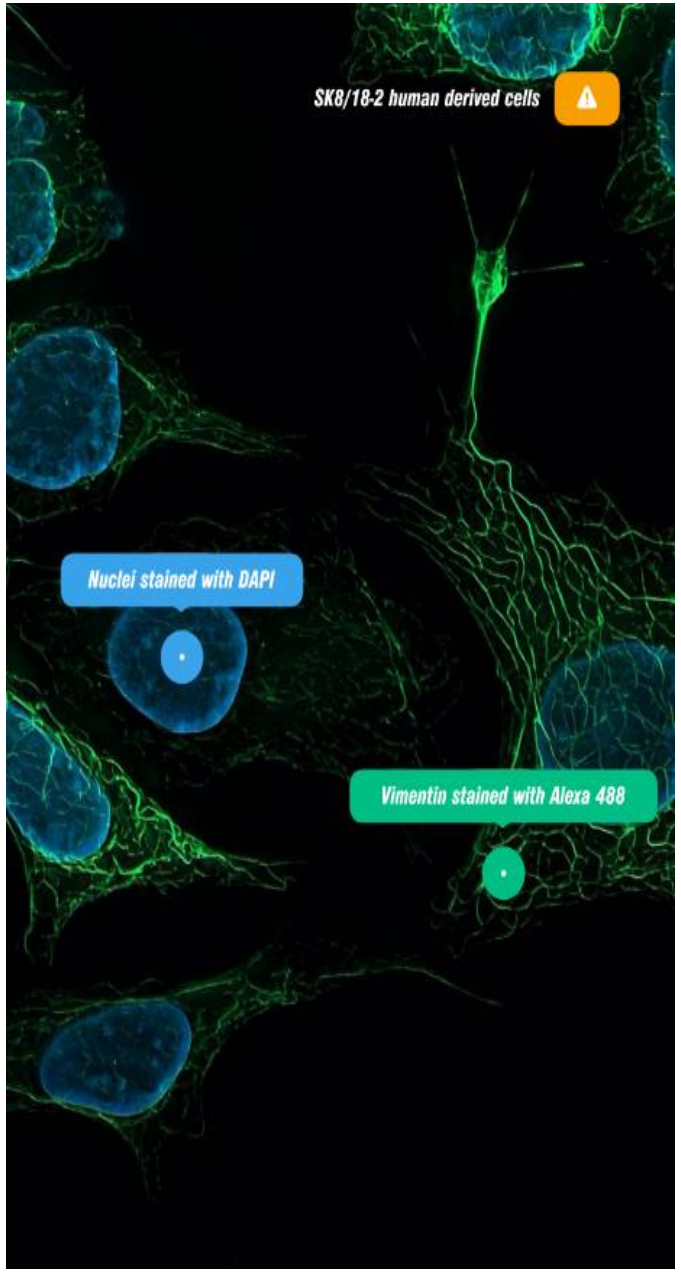
2014 - 2016: DL on the servers

2017 - now: back to the edge devices

Future: robotics, AR/VR; transformers

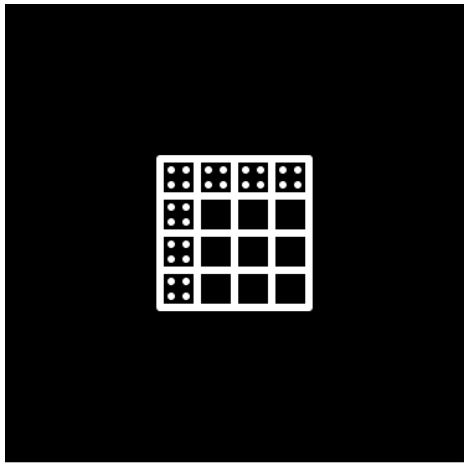
HOW TO SOLVE A PROBLEM


llinois
Land of Lincoln
568 612
MAR '89 IL
TEXAS
611 BU
I
MY TRU

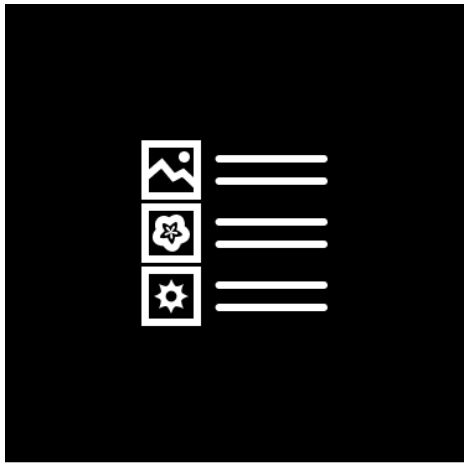

SK8/18-2 human derived cells
Nuclei stained with DAPI
Vimentin stained with Alexa 488


09.00-21.00

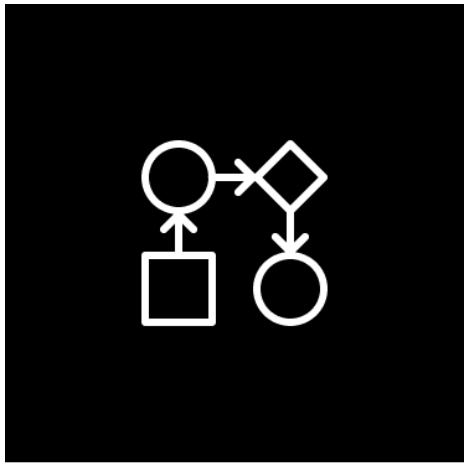
A DATA-BOUND PROBLEM



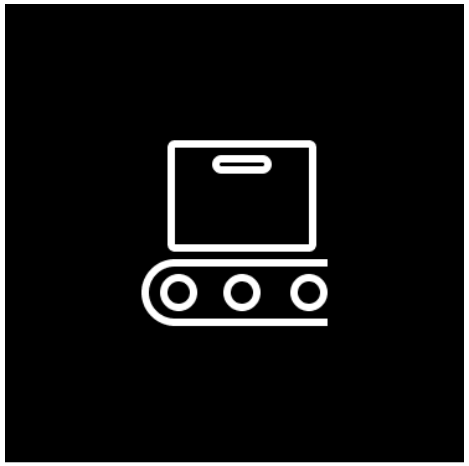
COLLECT
THE DATA



ANNOTATE
THE DATA



TRAIN
A MODEL



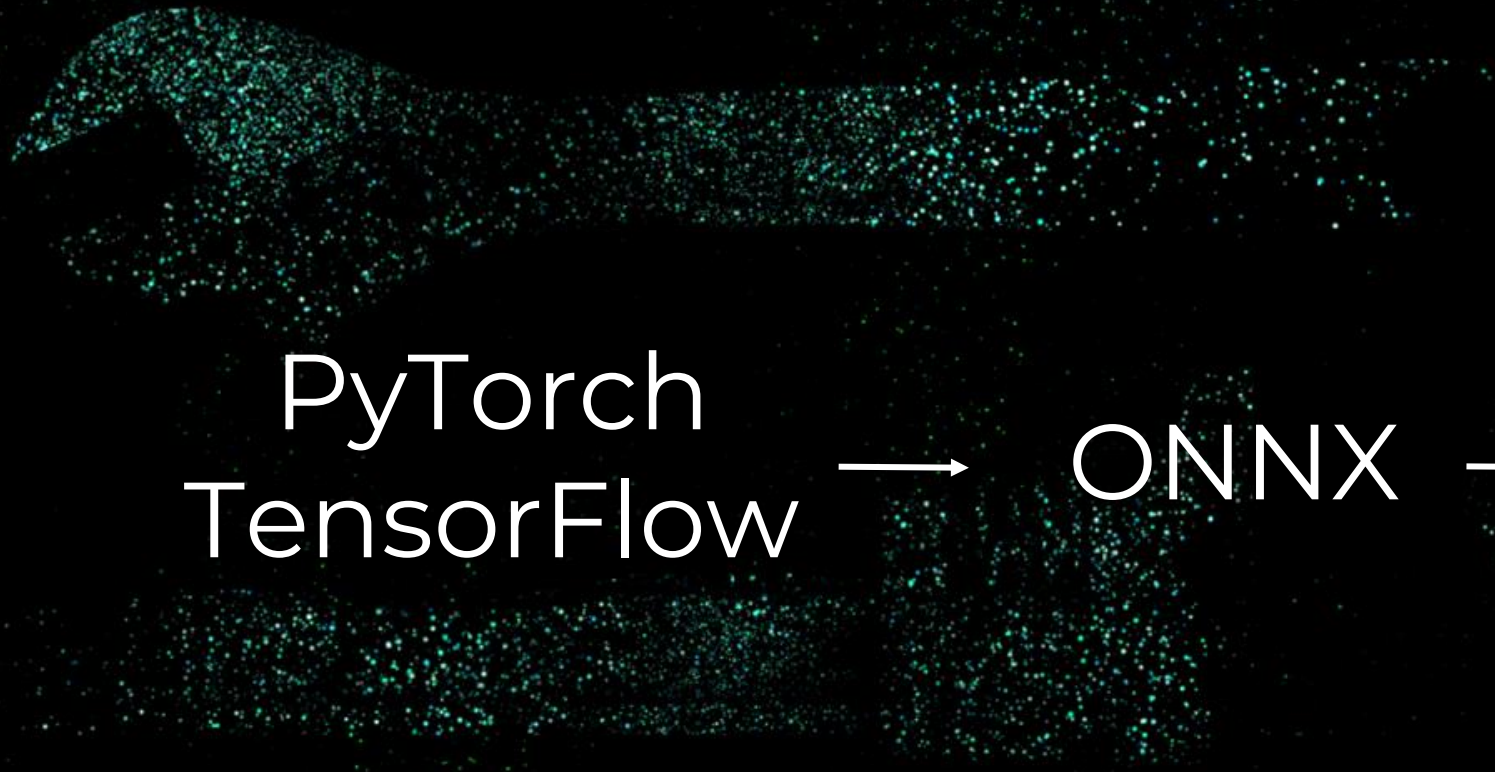
DEPLOY
TO TARGET ENV

SOFTWARE + HARDWARE

HARDWARE

SOFTWARE:
TOOLS and ALGORITHMS

TOOLS



PyTorch
TensorFlow

→

ONNX

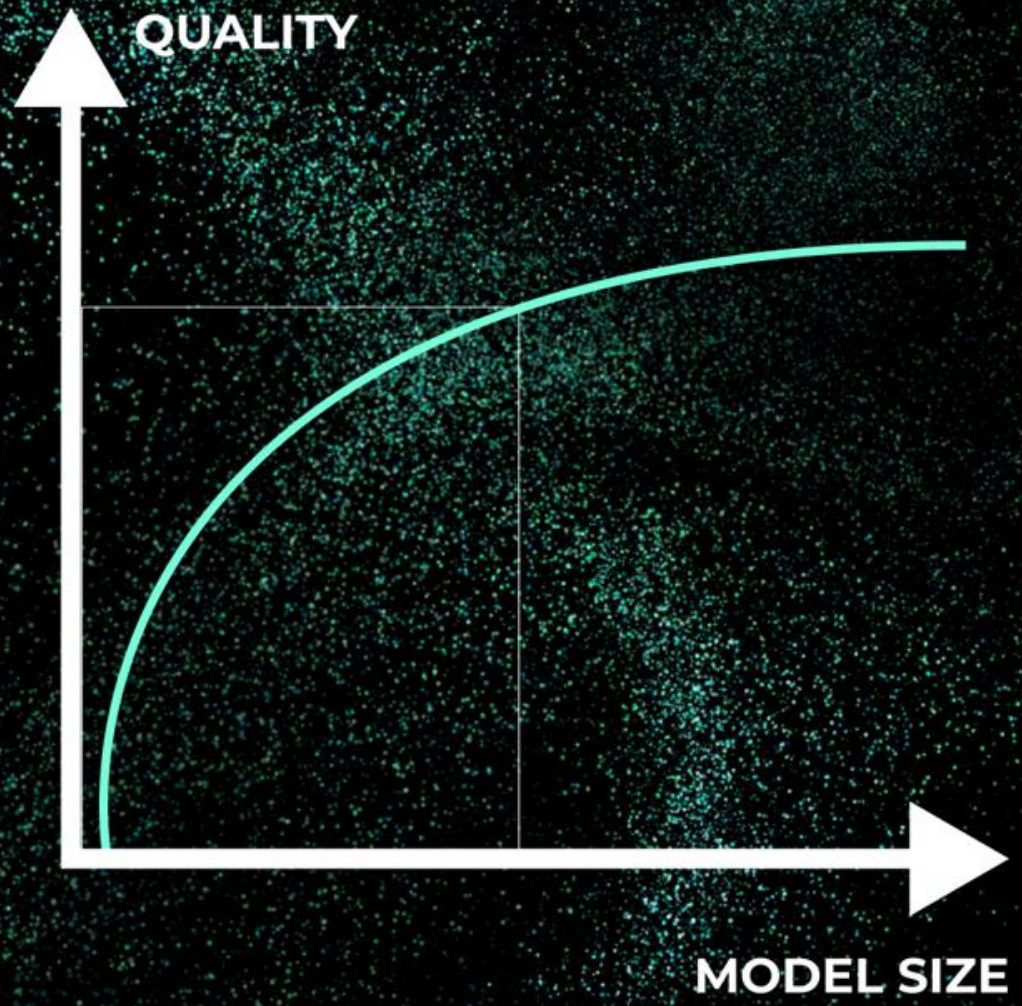
→

TFLite
CoreML
OpenVINO
TensorRT
PyTorch Mobile

ALGORITHMS

AI MODELS ARE HEAVY

- Size: up to 1 Gb
- RAM needed: up to 4 Gb
- Computations: hundreds of TFLOPs



AI MODELS ARE HIGHLY REDUNDANT

- Smaller models as baselines
- Quantization
- Pruning

50Mb -> 600Kb

PEOPLE DETECTION MODEL FOR FPGA

SUMMARY

anna@xperience.ai

Copyright Notice

www.tinyML.org