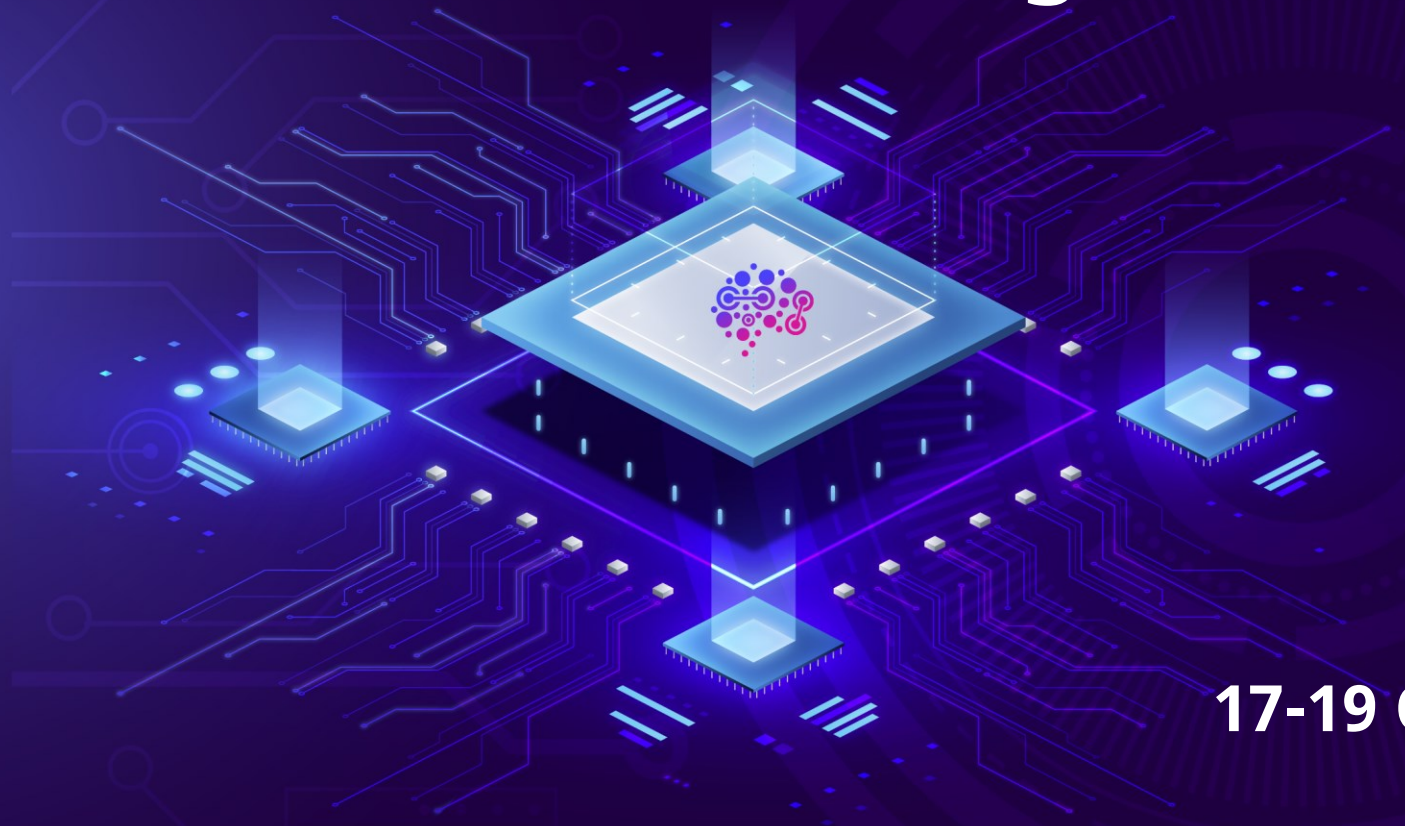
European Conference on EDGE AI Technologies and Applications - EEAI



Advancing emerging edge AI technologies and driving next generation intelligent applications.

17-19 October 2023, Athens, Greece

European Conference on EDGE AI Technologies and Applications - EEAI

Disease Identification for Vineyards Fields

Dr Lucas Mohimont, URCA [LICIIS], France

17-19 October 2023 Athens, Greece

Copyright © 2023

Introduction

- Plant diseases is a major cause of yield loss in the world
- Preventive spraying is used as a solution (not optimized)
- Research includes Computer Vision for:
  - Automatic spraying
  - Early diseases detection
  - Automatic disease mapping for visualization

Flavescence Dorée

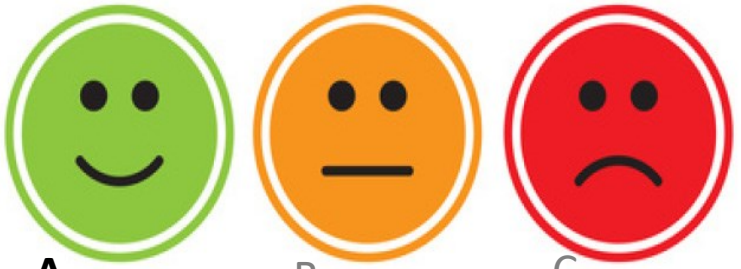
Esca

Copyright © 2023

Our previous work

- AI4DI – Grapes Disease Detection at the press
- Real time detection on the edge (Jetson Nano)

Quality grade



UNIVERSITÉ DE REIMS CHAMPAGNE-ARDENNE

VRANKEN POMMERY MONOPOLE

life.augmented

Copyright © 2023

State of the art 1/2

- Grape Leaf Disease classification
- Overused benchmark: PlantVillage

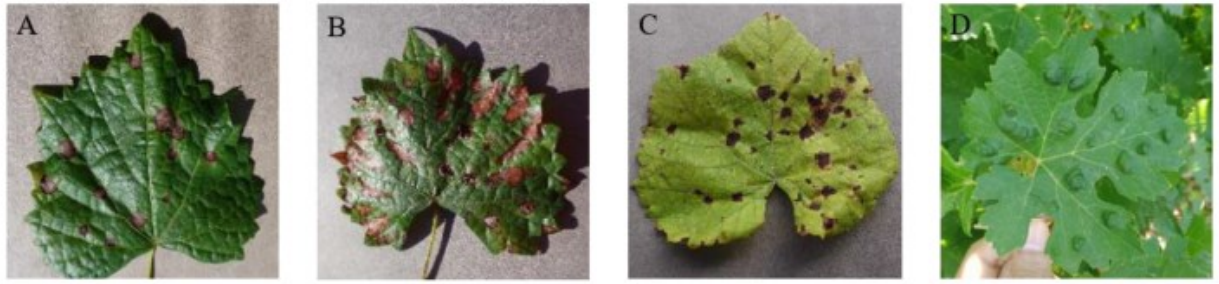


FIGURE 2 | Four common types of grape leaf diseases. **(A)** Black rot. **(B)** Black measles. **(C)** Leaf blight. **(D)** Mites of grape.

Xie, Xiaoyue, et al. "A deep-learning-based real-time detector for grape leaf diseases using improved convolutional neural networks." *Frontiers in plant science* 11 (2020): 751.

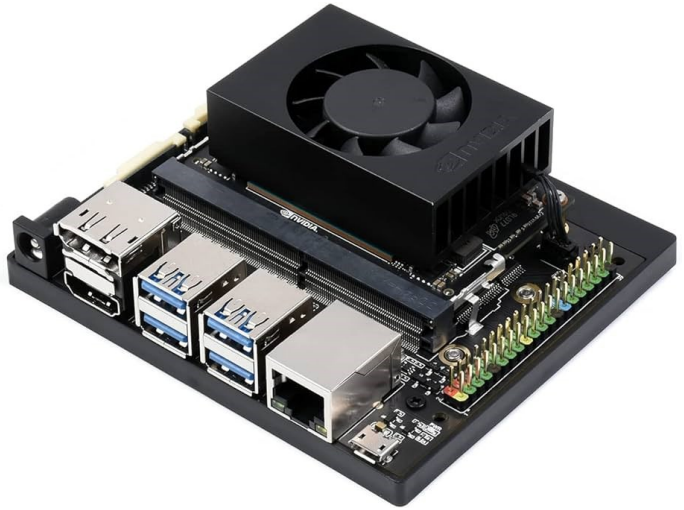
- Real-time Esca Classification on STM32
- 15 fps on a ARM Cortex-M7 MCU
- Limited to classification
- Precise detection is still a challenge

L. Falaschetti *et al*., "A Low-Cost, Low-Power and Real-Time Image Detector for Grape Leaf Esca Disease Based on a Compressed CNN," in *IEEE Journal on Emerging and Selected Topics in Circuits and Systems*, vol. 11, no. 3, pp. 468-481, Sept. 2021, doi: 10.1109/JETCAS.2021.3098454.

Copyright © 2023

State of the art 2/2

- Automatic symptoms detection & classification (Esca, Flavescence Dorée)
- 3 fps on “fat edge” Jetson Xavier (2048 x 2448)



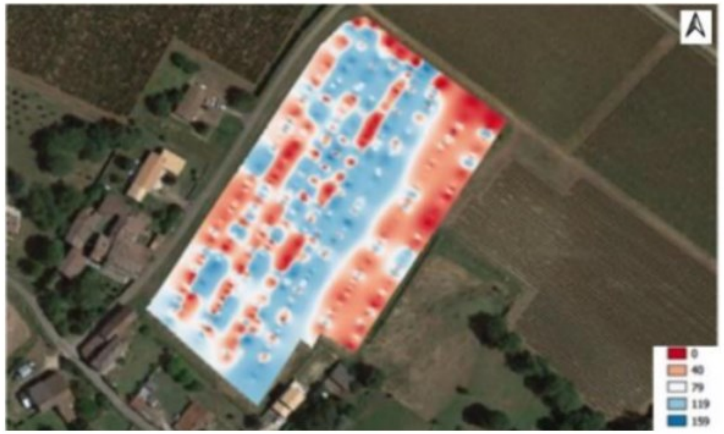
- GPU 512 cores Nvidia
- 16GB of ram
- Closer to a small workstation/laptop than edges devices

Malo Tardif, Ahmed Amri, Barna Keresztes, Aymeric Deshayes, Damian Martin, et al.. Two-stage automatic diagnosis of Flavescence Dorée based on proximal imaging and artificial intelligence: a multi- year and multi-variety experimental study.

Copyright © 2023

Long term objectives

- Adapt existing disease detection methodology to the Champagne case with our partner Vranken-Pommery
- Detection of downy mildew and deficiencies
- Inference on the edge with low-powered devices
  - STM32 with AI acceleration, Jetson Nano, FPGA?
- Automatic disease mapping with GPS
  - Many published results are made with GIS software



Nava et al., Big Wine Optimization Use Case: An IoT European Large Scale Pilot in Viticulture, 2022, Workshops at 18th International Conference on Intelligent Environments

Copyright © 2023

Our main research goal

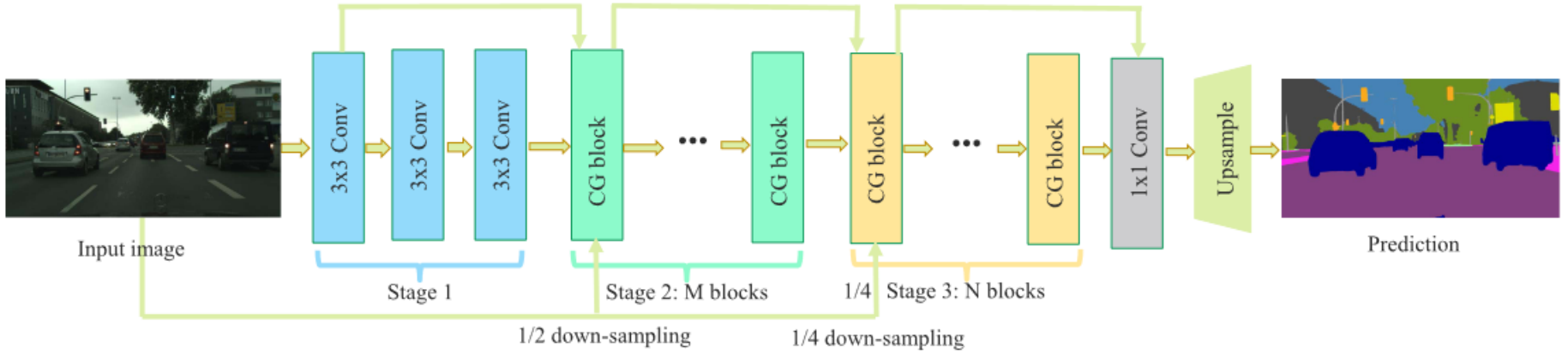
- Designing deep learning models for real-time image processing on the edge
  - Shared with the Yield Prediction demonstrator
- Useful for practical disease mapping without heavy cloud resources
- Can be transferred to many other applications
- Very difficult:
  - CNN are very expensive in term of memory and computation
  - Image size = quadratic complexity for both computation/memory
  - Example: InternImage “tiny” model = 30M parameters
  - With current methods: real time processing = loss in accuracy

HOW TO PROCESS HIS-RES IMAGES IN REAL-TIME ON EDGES DEVICES WITH HIGH ACCURACY?

Copyright © 2023

Output size

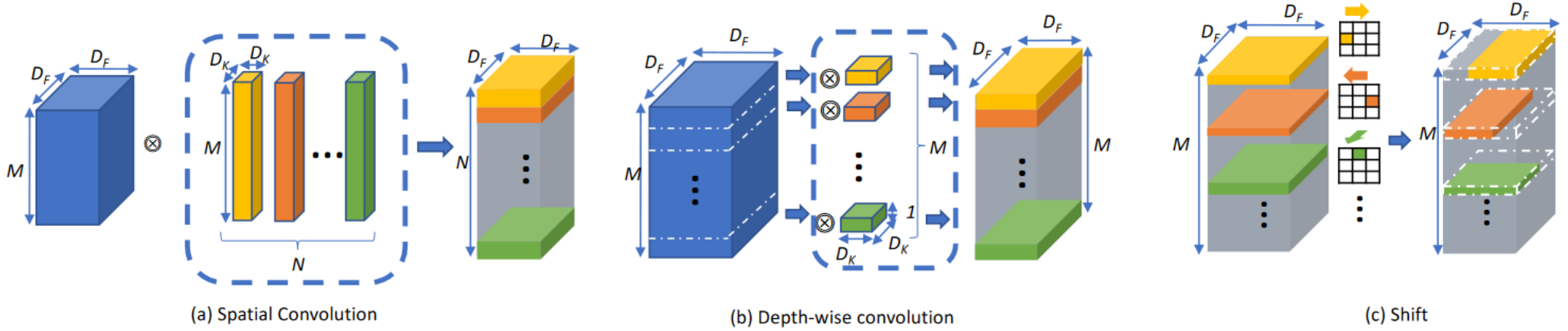
- Disease detection requires high-res images
- Object detection, segmentation = dense prediction
  - High memory requirement
- Compromise:
  - Input: high-res image (> 1024x1024p)
  - Output: lower-res masks and/or bounding box (ex: 256x256)
- Example: CGNet model(Wu et al. 2020)



Copyright © 2023

Base operations

- Classical 2D convolutions can be expensive
  - Example: 256 filters with 256 channels = ~600k parameters
- Cheaper alternatives exist
  - Depthwise Separable Convolutions
    - Example: 256 filters with 256 channels = ~68k parameters
  - Shifting: 0 parameters, 0 flops!
  - L1 distance: only uses addition ops
  - Etc.



Wu, Bichen, et al. "Shift: A zero flop, zero parameter alternative to spatial convolutions." *Proceedings of the IEEE conference on computer vision and pattern recognition*. 2018.

Copyright © 2023

Numeric precision

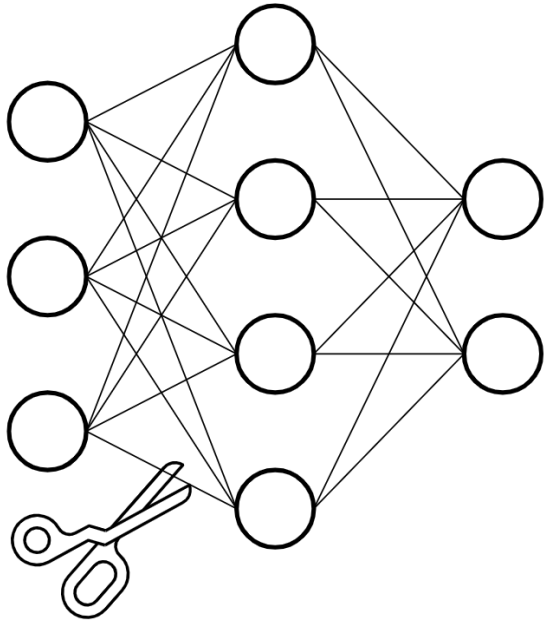
- CNN parameters don’t need Float32 precision
- Efficient alternatives depends on the hardware:
  - GPU with cuda cores: mixed precision / Float16
  - GPU with tensor cores: integer 8bits
  - TPU (ex: Google Coral): integer 8bits
  - FPGA: integer 8 bits or less
  - Etc.
- Binary, ternary coding is also possible
- Quantization Aware Training

Float32 = 4 Billions possibilities → **Float16 = 65k** → **Int8 = 256**

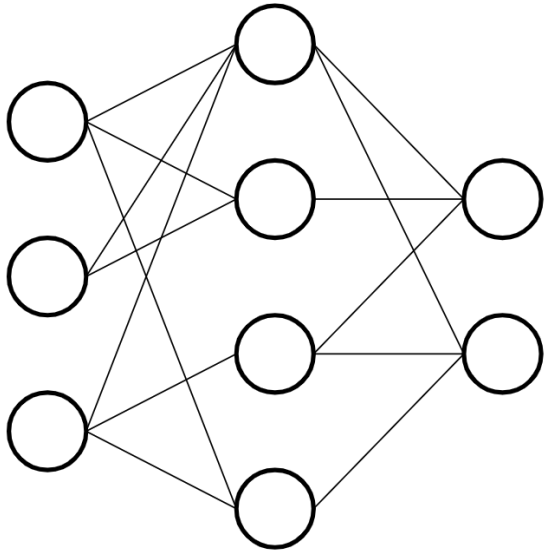
Copyright © 2023

Model pruning

- You can remove some filters
- Filters thresholding
- Implemented in many standard API



Before pruning



After pruning

Copyright © 2023

Model compression

- Filters approximation with CANDECOMP/PARAFAC (CP) tensor decomposition
- The results: more layers with less parameters
- 15fps for Esca Disease Classification on STM32 HW

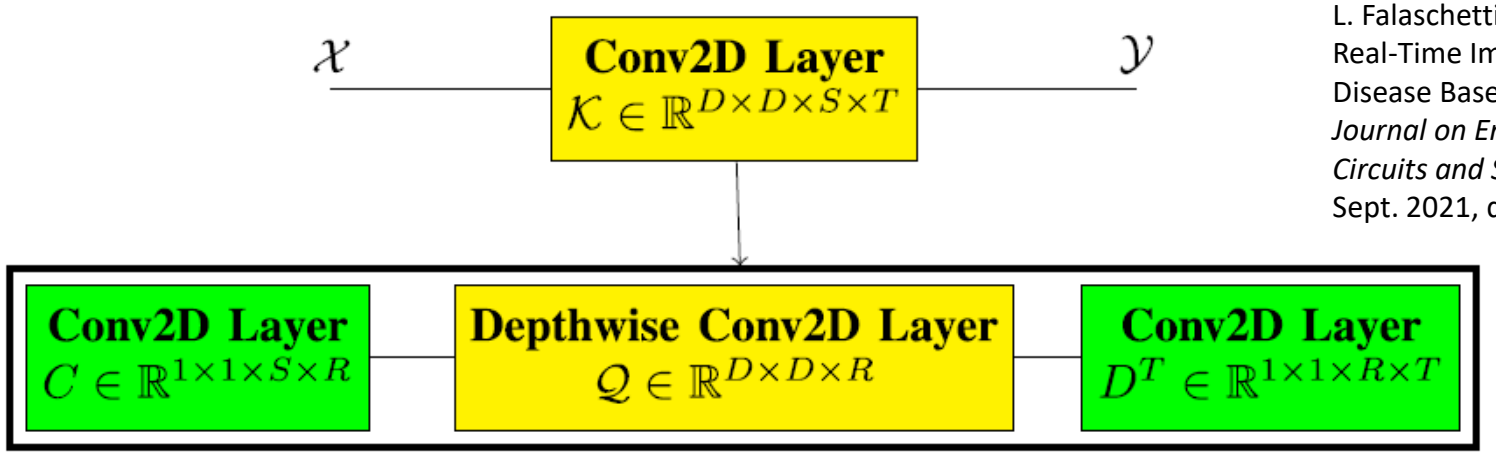


Fig. 6. The effect of CP decomposition on Conv2D Layer.

L. Falaschetti *et al.*, "A Low-Cost, Low-Power and Real-Time Image Detector for Grape Leaf Esca Disease Based on a Compressed CNN," in *IEEE Journal on Emerging and Selected Topics in Circuits and Systems*, vol. 11, no. 3, pp. 468-481, Sept. 2021, doi: 10.1109/JETCAS.2021.3098454.

Copyright © 2023

Conclusion

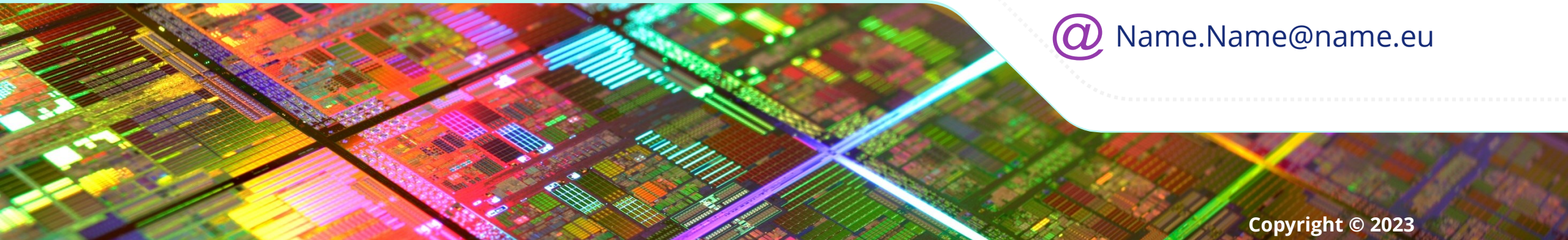
- Efficient CNN models will be transfered to vine disease detection
- Current development/research is focused on yield prediction
- Data collection campaign for vine disease will begin in Summer 2024
- Early results: 0.5s inference time on 4000x3000 images
- Tested on Jetson Nano 2GB with Onnx Runtime
- Accuracy is not always satisfying!

Copyright © 2023

EDGE AI

Thank You

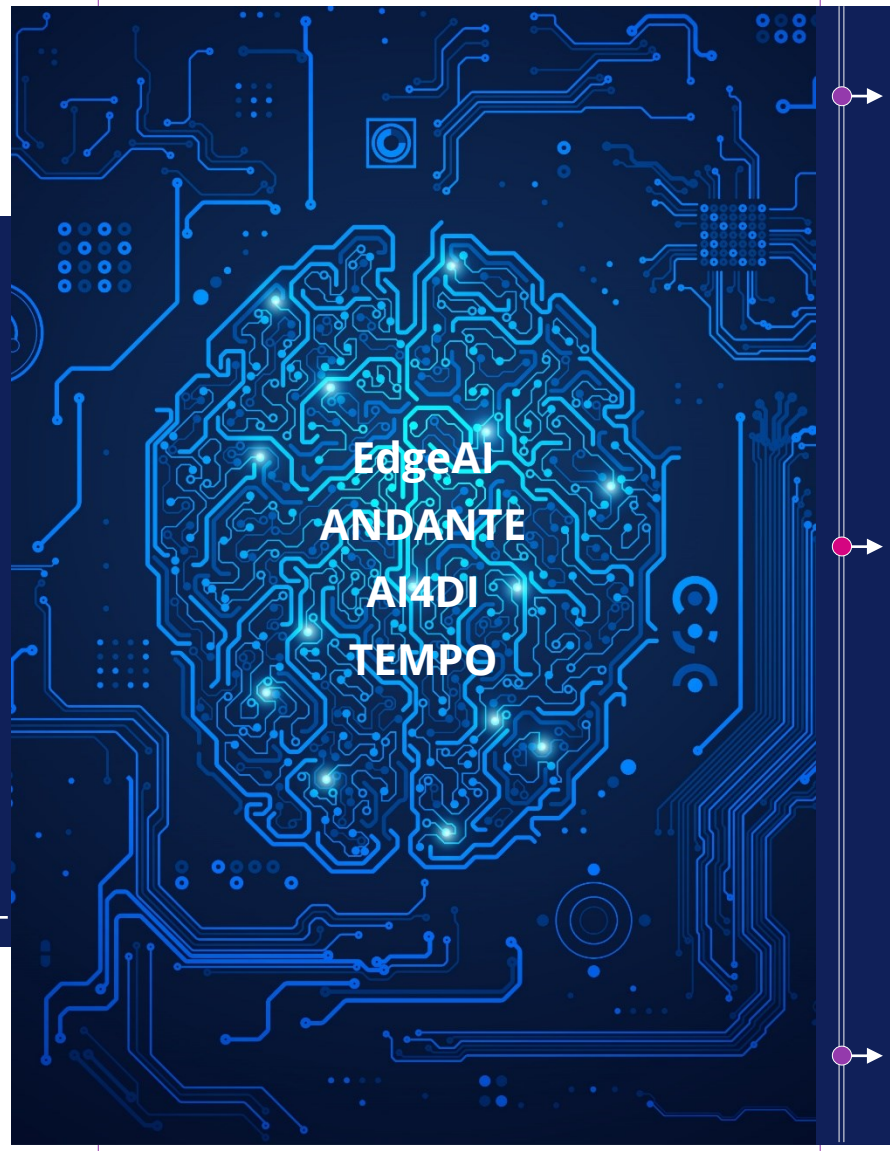
For your attention



@ Name.Name@name.eu

Copyright © 2023

Event Organisers



EdgeAI (Edge AI Technologies for Optimised Performance Embedded Processing) develops new electronic components and systems, processing architectures, connectivity, software, algorithms, and middleware through the combination of microelectronics, edge AI, embedded systems, and edge computing. www.edge-ai-tech.eu

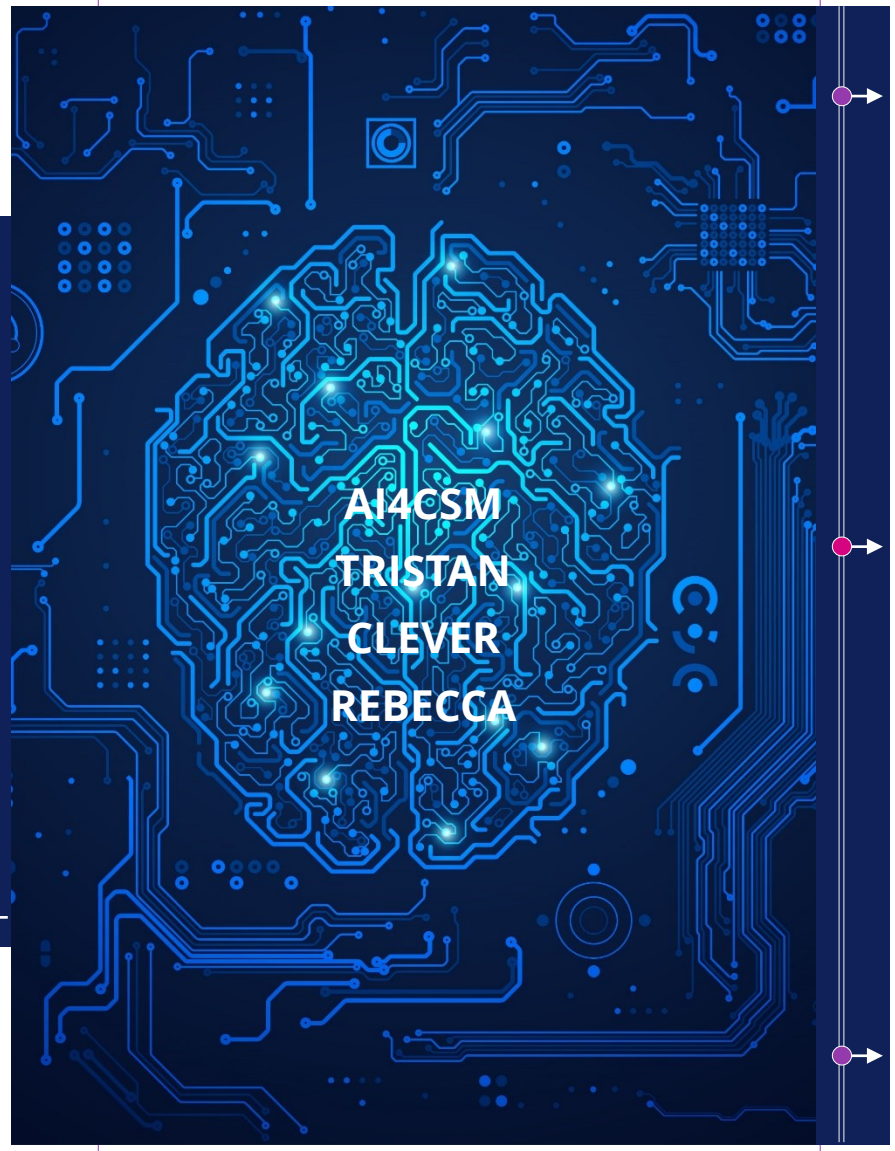
The ANDANTE project has received funding from the ECSEL Joint Undertaking (JU) under grant agreement No 876925. The JU receives support from the European Union's Horizon 2020 research and innovation programme and Belgium, France, Germany, The Netherlands, Portugal, Spain, Switzerland. www.andante-ai.eu

The TEMPO project has received funding from the ECSEL Joint Undertaking (JU) under grant agreement No 826655. The JU receives support from the European Union's Horizon 2020 research and innovation programme and Belgium, France, Germany, The Netherlands, Switzerland. www.tempo-ecsel.eu

The AI4DI project has received funding from the ECSEL Joint Undertaking (JU) under grant agreement No 826060. The JU receives support from the European Union's Horizon 2020 research and innovation programme and the national authorities. www.ai4di.eu

Copyright © 2023

Event Organisers



AI4CSM (Automotive Intelligence for Connected Shared Mobility) develops advanced electronic components and systems (ECS) and architectures for future mass-market ECAS vehicles based on sensor fusion and perception platforms, efficient propulsion and energy modules, advanced connectivity and trustworthy AI techniques and methods. www.ai4csm.eu

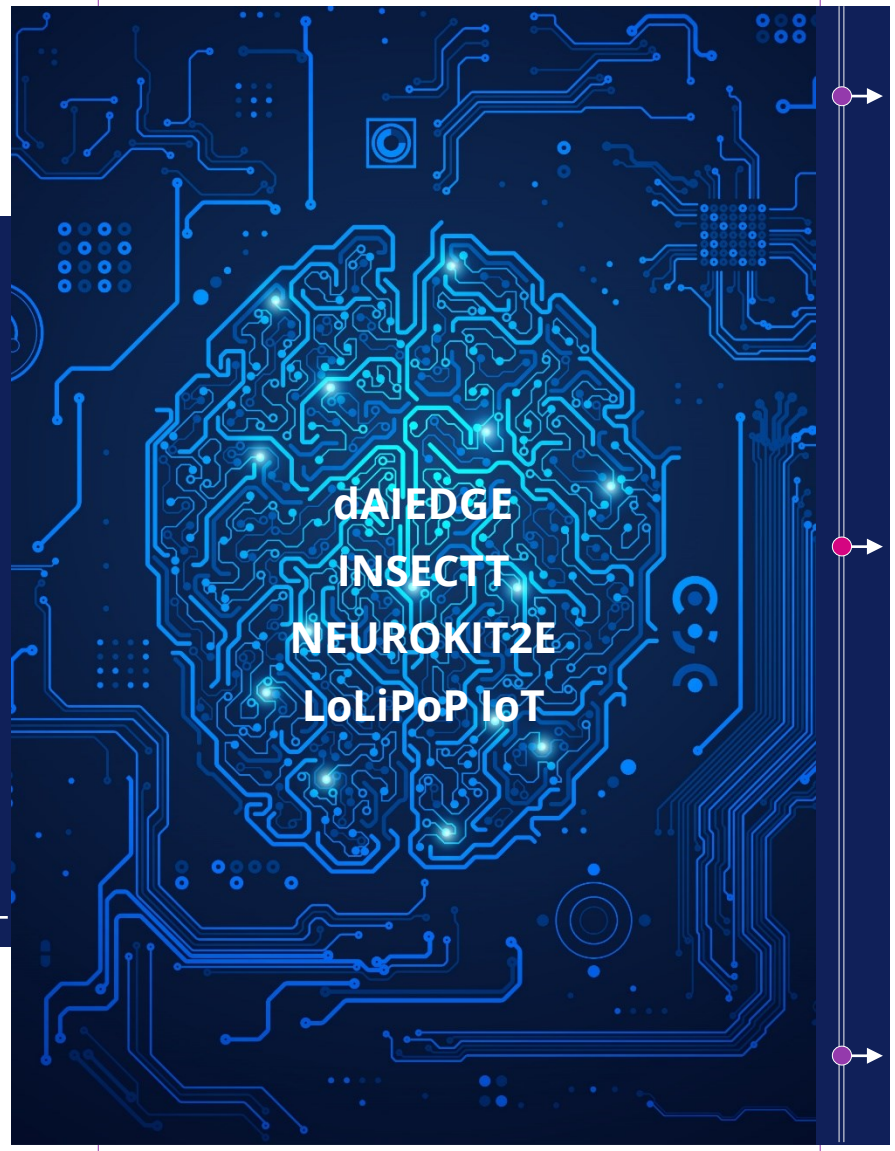
TRISTAN aims to expand, mature, and industrialize the European RISC-V ecosystem to compete with commercial options by leveraging the Open-Source community to gain productivity and quality. A European strategy for RISC-V designs will be defined, creating a repository of industrial rate building blocks for SoC designs in various application domains. www.tristan-project.eu

CLEVER (Collaborative edge-cLoud continuum and Embedded AI for a Visionary industry of thE futuRe) proposes innovations in hardware accelerators, design stack, and middleware software that revolutionize the ability of edge computing platforms to operate federated, leveraging sparse resources that are coordinated to create a powerful swarm of resources. www.cleverproject.eu

REBECCA (Reconfigurable Heterogeneous Highly Parallel Processing Platform for safe and secure AI) aims to democratize the development of edge AI systems and create a complete hardware and software stack centered around a RISC-V CPU, which offer higher performance, energy efficiency, safety, and security than existing systems. www.rebecca-chip.eu

Copyright © 2023

Event Organisers



dAIEDGE

dAIEDGE is the European Network of Excellence for distributed, trustworthy, efficient, and scalable AI at the Edge and promotes the application, development, and deployment of Artificial Intelligence (AI) on edge computing platforms.

InSecTT

INSECTT (Intelligent Secure Trustable Things) provide intelligent, secure, and trustworthy systems for industrial applications to provide comprehensive cost-efficient solutions of intelligent, end-to-end secure, trustworthy connectivity and interoperability to bring the Internet of Things and Artificial Intelligence together. www.insectt.eu

NEUROKIT2E

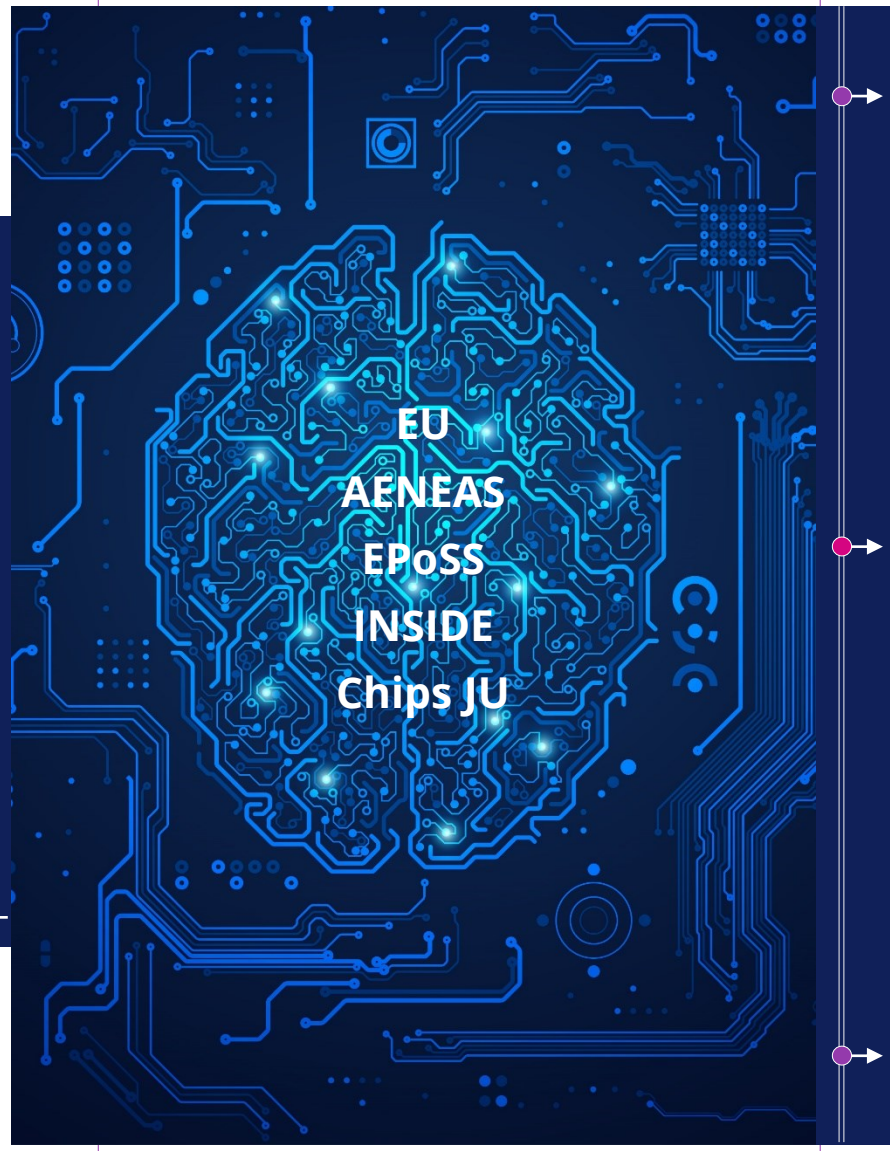
NEUROKIT2E (Open source deep learning platform dedicated to Embedded hardware and Europe) proposes a Deep Learning Platform for Embedded Hardware around an established European value chain (AI HW/SW). The solutions developed support neural network design, optimisation, and implementation on constrained HW.

LoLiPoP-IoT

The objectives of LoLiPoP IoT (Long Life Power Platforms for Internet of Things) are to develop energy harvesting-based innovative Long Life Power Platforms that enable retrofit of wireless sensor network edge devices for asset tracking, condition and performance monitoring. www.lolipop-iot.eu

Copyright © 2023

Supporting Organizations



AENEAS standing for Association for European NanoElectronics ActivitieS, is an industrial Association, established in 2006, providing unparalleled networking opportunities, policy influence & supported access to funding to all types RD&I participants in the field of micro and nanoelectronics enabled components and systems. https://aeneas-office.org/

The European Technology Platform on Smart Systems Integration is an industry-driven policy initiative, defining research, development and innovation needs as well as policy requirements related to Smart Systems Integration and integrated Micro- and Nanosystems. The main objective is to develop a vision and to set up a Strategic Research Agenda. www.smart-systems-integration.org

Inside Industry Association is the European Technology Platform for research, design and innovation on Intelligent Digital Systems and their applications. The Association is a membership organisation for the European research and innovation actors with more than 200 members and associates from all over Europe. www.inside-association.eu

Chips Joint Undertaking supports research, development, innovation, and future manufacturing capacities in the European semiconductor ecosystem. Launched as part of the Chips for Europe Initiative, it confronts semiconductor shortages and strengthens Europe's digital autonomy, engaging a significant EU, national/regional and private industry funding of nearly €11 billion. www.kdt-ju.europa.eu

Copyright © 2023