AI@EDGE

Applying Artificial Intelligence on Edge devices using Deep Learning with Embedded optimizations

VLAIO TETRA HBC.2019.2641

Project kick-off 03-03-2020
User group meeting ai-edge.be
iot-incubator.be www.eavise.be

1

Agenda

1. Project partners
2. Project goals
3. Members of the user group
4. Poll & discussion
5. Demo
6. Administration

2

Project partners

IoT Incubator

VIVES campus Brugge

https://iot-incubator.be

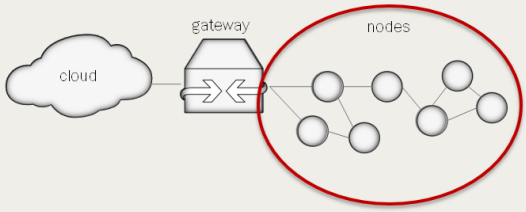
eavise
VISION - AI - EMBEDDED

KU Leuven campus De Nayer

www.eavise.be

vives KU LEUVEN AI@EDGE 3

3



IoT-Incubator.be

Research – Prototyping – Training

Piet Cordemans

Research and education | Internet of Things
VIVES University of Applied Sciences | Industrial Sciences and Technology
Spoorwegstraat 12 | 8200 Brugge | BELGIUM
piet.cordemans@vives.be | IoT-Incubator.be

vives

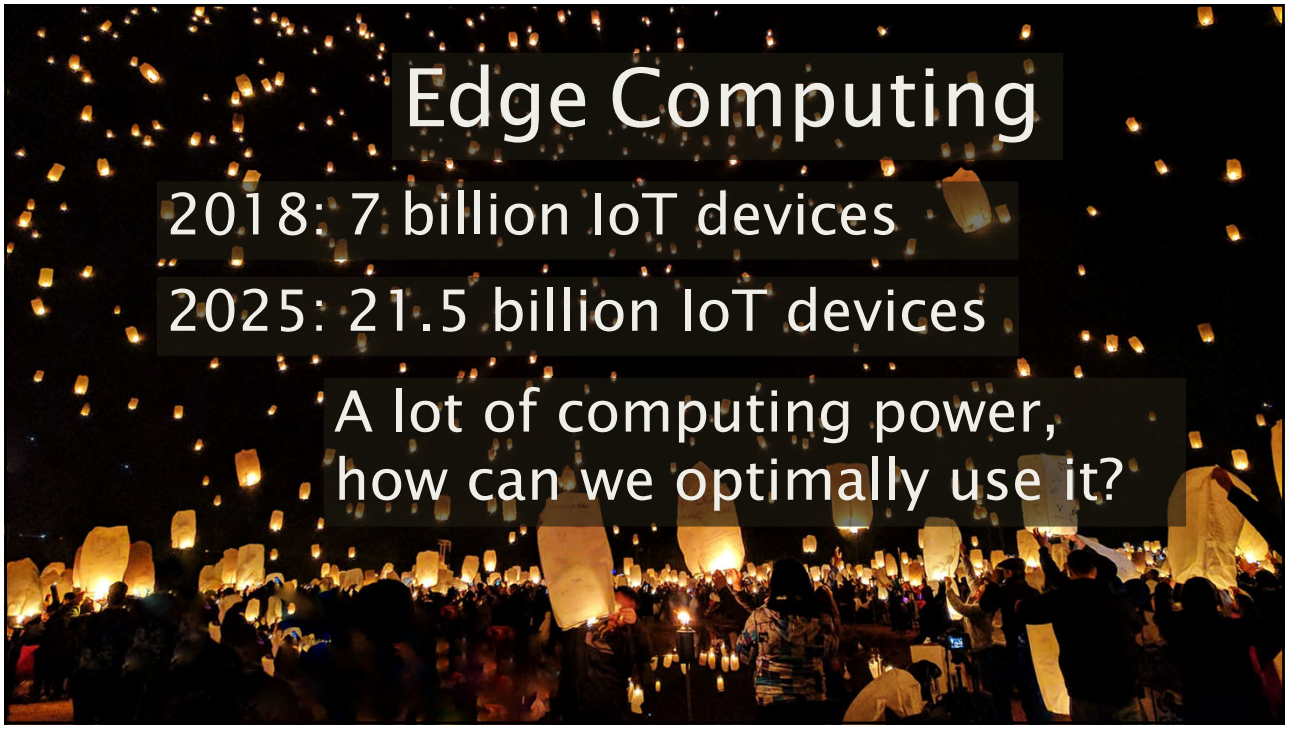
4

Research

Edge Computing
LoRaWAN
Security

Building prototypes
Feasibility
Technology transfer

5



6

Edge Computing

conserves bandwidth

better privacy

deals with network failure

autonomous systems

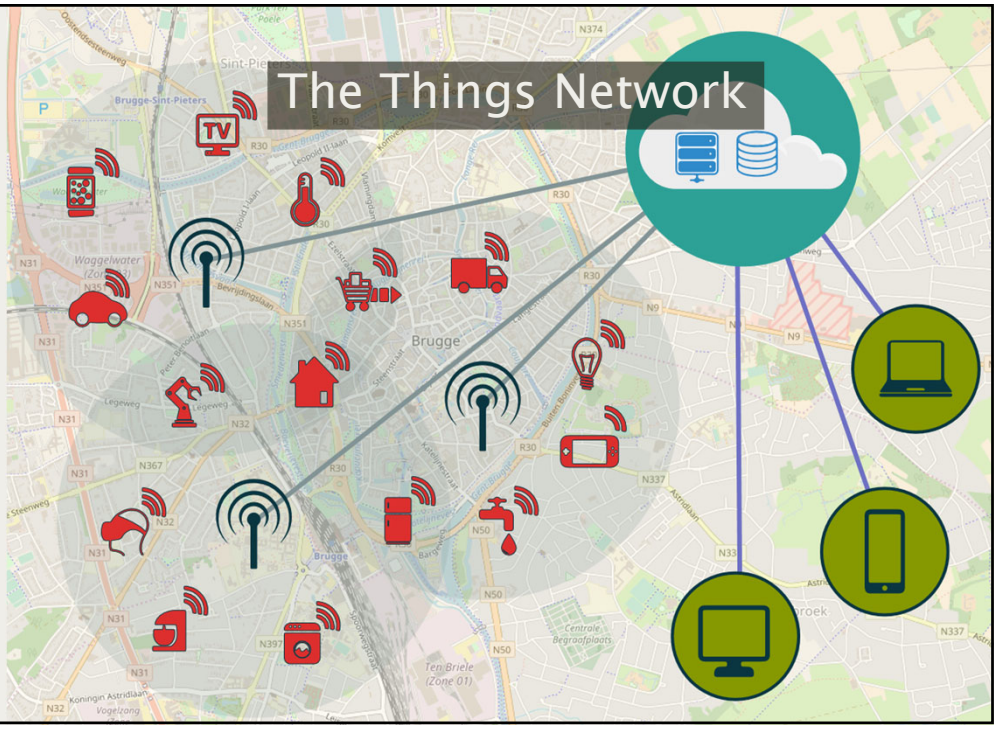
smarter applications

reduces latency

7

LoRaWAN

Low power
Low bitrate
Wide range
Reliable
Secure



8


LoRaWAN
TTN Gateway
PCBs
Dashboard
Apps
PWO project

9


Training the Human Firewall
PWO project
The greatest culprit is not technology,
but human failure.
-- Secure IT

10

Services

Custom prototyping

Training courses

LoRa Academy
Internet of Things: Mbed
Cisco Academy
- CCNA
- Cybersecurity Operations

Ethical hacking

App development
- NativeScript
- Node.js

DevOps
Wireshark

https://www.vives.be/nl/navormingen-ict

11

KU LEUVEN eavise

EAVISE:
Embedded & Artificially intelligent Vision Engineering

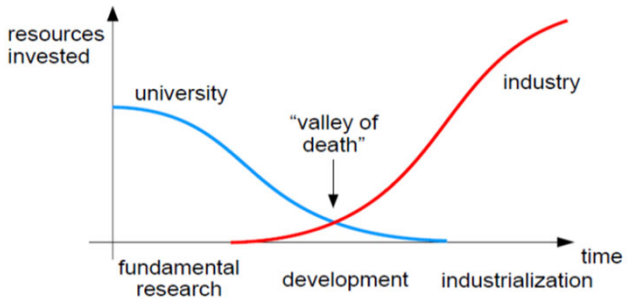
Toon Goedemé
Joost Vennekens
Patrick Vandewalle

eavise
VISION - AI - EMBEDDED

12

Valley of death

KU LEUVEN VISICS DTAI → eavise → industry



vives KU LEUVEN AI@EDGE

13

Embedded & Artificially Intelligent Vision Engineering

- Research goal:
  - **Translating** state-of-the-art **image processing** algorithms and **artificial intelligence** techniques to solutions for **industry-specific application** problems
  - Optimizing vision algorithms to **real-time performance**
  - **Increasing robustness** of experimental algorithms to industry standards
  - Implementing **advanced image processing applications** on **embedded systems**:
    FPGA, DSP, GPU, multicore CPU, cluster

vives KU LEUVEN AI@EDGE

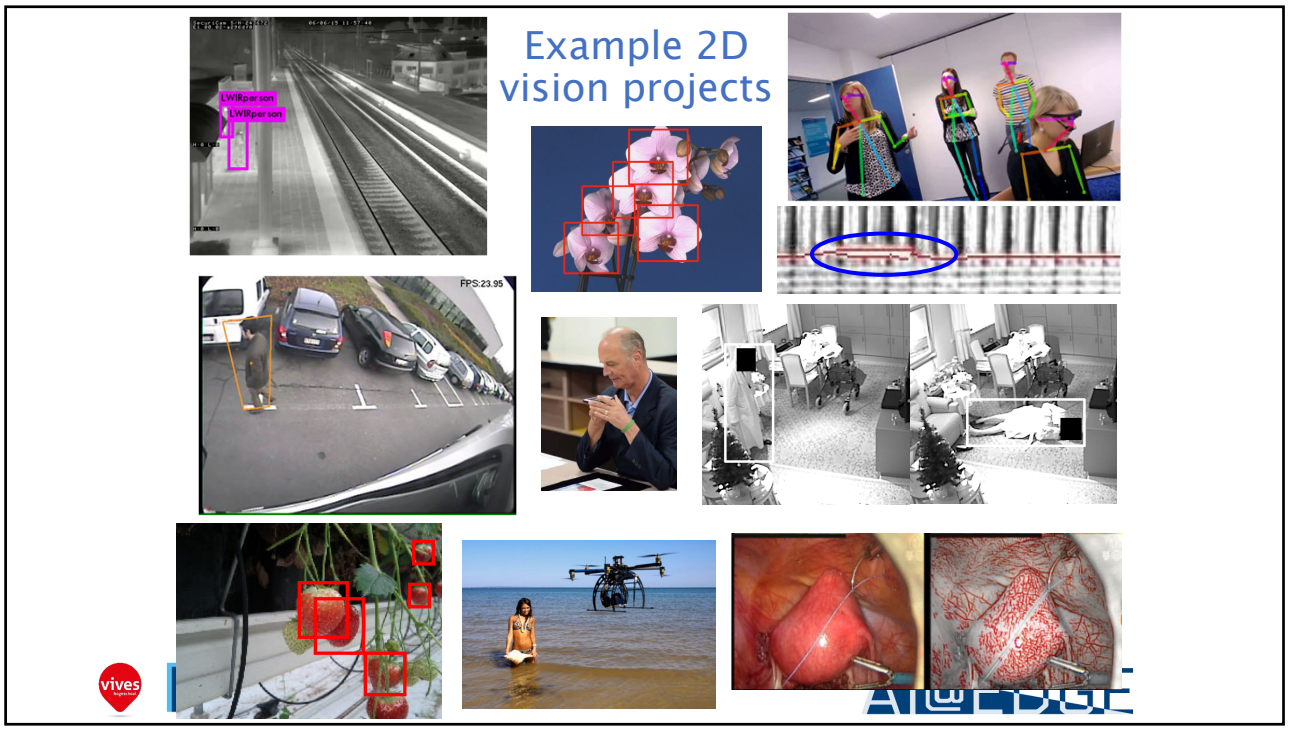
14


People@
EAVISE
Prof. Dr. Toon Goedemé
research leader
2D Computer Vision
Prof. Dr. Patrick Vandewalle
research leader
3D Computer Vision
Prof. Dr. Joost Vennekens
research leader
Dr. Kristof Van Beeck
Real-time
embedded vision
Dr. Steven Puttemans
Object Detection
Dr. Dries Hulens
Robotic UAV
Dr. Wiebe Van Ranst
GPU & Mobile
Floris De Feyter
Person Re-ID
Maarten
Vandersteegen
IR Person Detection
Uma Raman Kumar
3D deep learning
Inge Coudron
3D scene interpretation
Tanguy Ophoff
720° person detection
Timothy Callemein
2D scene interpretation
Robby Neven
On-edge learning
Wouter Sterckens
CV for recycling
Kylian Van Dessel
Timetabling
Laurent Mertens
AI Time series
Marjolein De Ryck
Time series AI
Bram Aerts
KR for industrial appl.
Simon Vandevelde
KR and time series
Chen-Chou Lo
3D Sensor Fusion
Laurens Le Jeune
FPGA DL
vives
KU LEUVEN
AI@EDGE

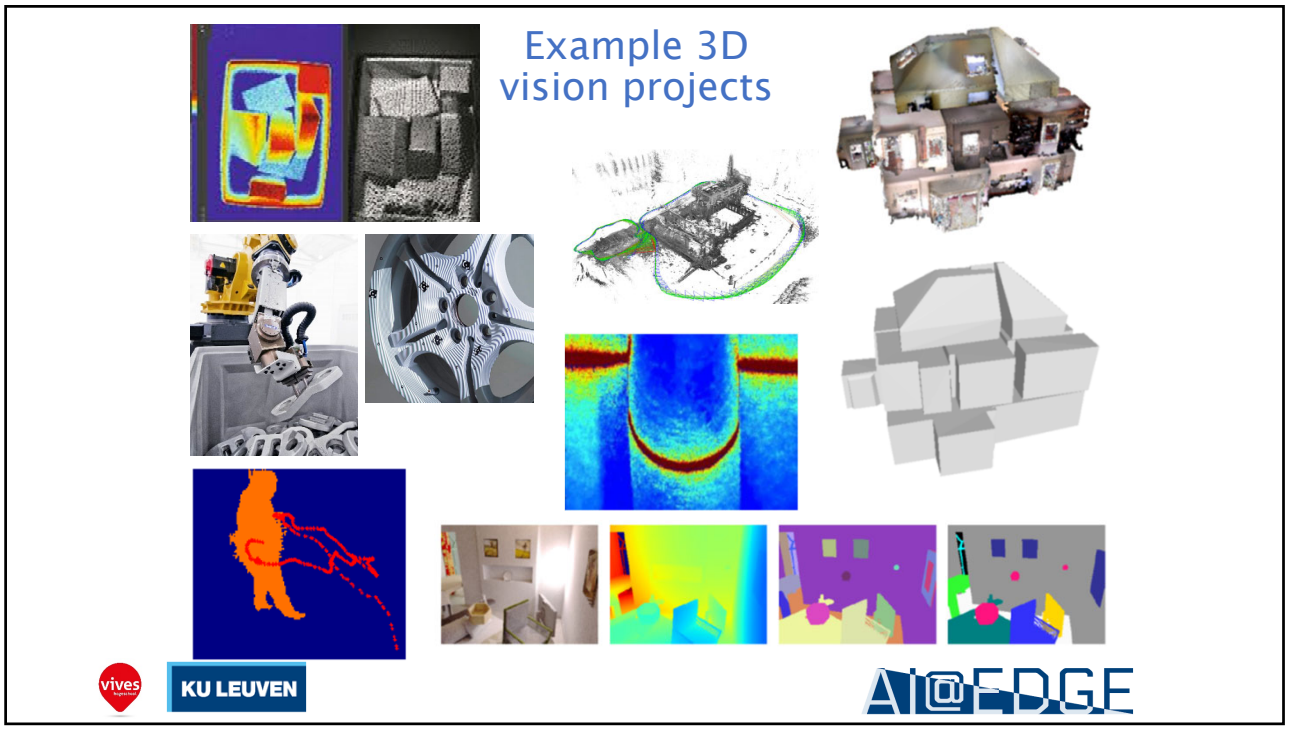
15


People@
EAVISE
Prof. Dr. Toon Goedemé
research leader
2D Computer Vision
Prof. Dr. Patrick Vandewalle
research leader
3D Computer Vision
Prof. Dr. Joost Vennekens
research leader
Dr. Kristof Van Beeck
Real-time
embedded vision
Dr. Steven Puttemans
Object Detection
Dr. Dries Hulens
Robotic UAV
Dr. Wiebe Van Ranst
GPU & Mobile
Floris De Feyter
Person Re-ID
Maarten
Vandersteegen
IR Person Detection
Uma Raman Kumar
3D deep learning
Inge Coudron
3D scene interpretation
Tanguy Ophoff
720° person detection
Timothy Callemein
2D scene interpretation
Robby Neven
On-edge learning
Wouter Sterckens
CV for recycling
Kylian Van Dessel
Timetabling
Laurent Mertens
AI Time series
Marjolein De Ryck
Time series AI
Bram Aerts
KR for industrial appl.
Simon Vandevelde
KR and time series
Chen-Chou Lo
3D Sensor Fusion
Laurens Le Jeune
FPGA DL
vives
KU LEUVEN
AI@EDGE

16


Example 2D
vision projects
LWIRperson
LWIRperson
FPS:23.95

17


Example 3D
vision projects
KU LEUVEN

18

Commercial valorisation success stories

ARIS

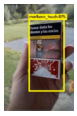
Orchid detection for grading, classification and inspection. Installed in virtually all orchid nurseries in B and NL

geronimo

Automatic cinematographic steering and selection of unmanned cameras. Used in the video production "Het Gezin" (VRT)

OCTINION
Engineering the future

Strawberry detection for automatic picking robot. Integrated in final pre-commercial robot prototype



EMAKINA

Real-time deep learning based product and brand recognition. Used in Augmented Reality app by Philip Morris.

Sensotec

App reads printed text for blind and visually impaired. Downloaded more than 50.000 times at a price of 100 USD.

VANDEN BROELE GROUP
COMMUNICATION & PUBLISHING

Face recognition and aesthetical photo quality assessment. Core engine of a commercial individualised school photo album generation service.

SATURNUS
A Sony Group Company

Digital stabilisation of endoscope images at extreme low latency. Included in commercial product NUCLeUS.

3FROG

Automatic scan completion and segmentation software of 3D scans of building interiors. Commercially exploited in real estate sales.

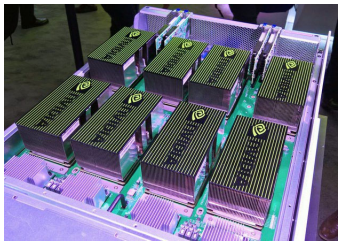
vives
KU LEUVEN

AI@EDGE

19

EAVISE Infrastructure

- Cameras
- Embedded platforms
- Desktop GPUs
- NVIDIA DGX-1 GPU server
  - 8X NVIDIA Tesla V100 32 GB/GPU
  - 512 GB DDR4 RDIMM
  - 2X 20-Core Intel® Xeon® E5-2698 v4 2.2 GHz



vives
KU LEUVEN

AI@EDGE

20

EAVISE: some statistics

- **Personnel**:
  - 3 research leaders
  - 4 postdocs
  - 14 PhD researchers
- 70+ **projects** since 2008 (~9M€, 800 industrial partners):
  - 15 IWT/AIO-tetra/cornet
  - 14 KMO-portefeuille
  - 10 IWT/AIO KMO-innovation projects
  - 12 IWT/AIO O&O projects
  - 1 KUL-GOA, 1 IWT-SB, 1 FWO-SBO, 1 Baekeland
- 300+ international **publications** since 2008
- 1 patent application

- Awards:
  - Dries Hulens & Floris De Smedt: **Best paper award** at CVPR Embedded Vision Workshop 2015
  - Wiebe Van Ranst: **Best demo award** Benelux Conference on Artificial Intelligence BNAIC 2015
  - Kristof Van Beeck: **Best poster award** at Technology Cluster ESAT Research Symposium 2015
  - Toon Goedemé: **Willy Asselman Award** for research achievements 2016
  - Wiebe Van Ranst: **Best paper award** at CGVCVIP 2016
  - Timothy Callemein: **Best presentation award** at ACIVS 2020

21

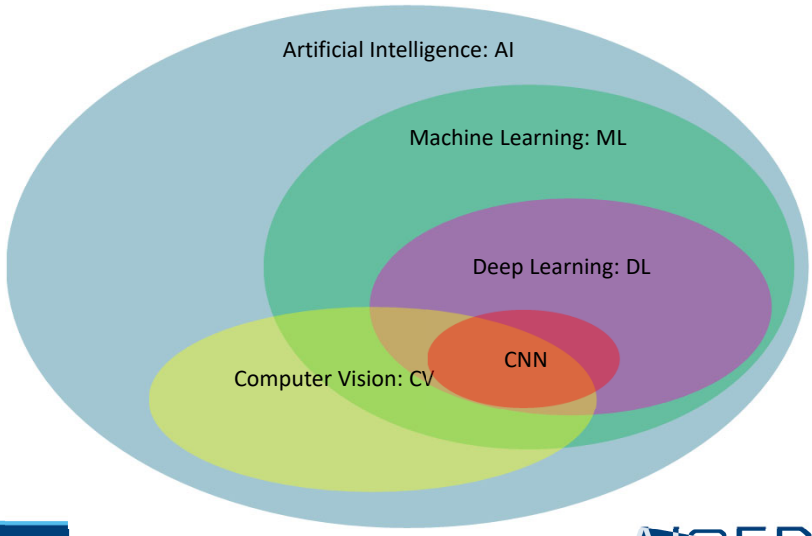
Project goals

- AI: Deep Learning Inference
  - Trained model -> target system
- @EDGE:
  - Low-cost small embedded systems
    - Microcontrollers & system processsors
  - Less power (computing & energy)
  - Autonomous systems
  - Time-critical & privacy-minded systems
    - Local decisions – Local data
    - Low latency
    - No dependence on network connectivity

22

22

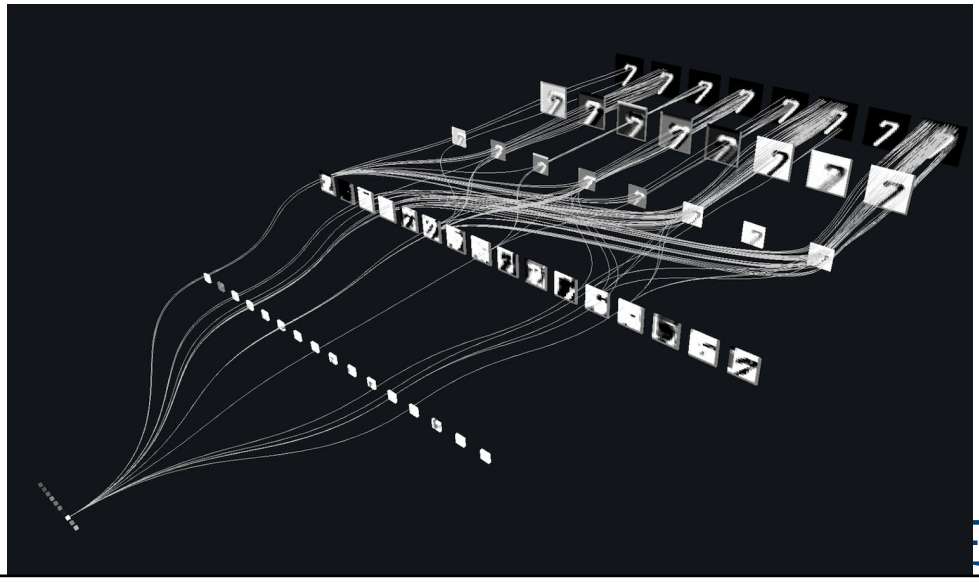
AI, Machine learning, Deep learning, …?



23

23

Deep learning – what?



24

Deep learning – what?

- Deep learning is not that new!
  - Yann Le Cun: A theoretical framework for back-propagation (1998)
  - Around 2012: breakthrough: enough **datasets** (Google, Facebook, Microsoft, Baidu,…), **architectures**, but most of all **affordable GPU hardware**

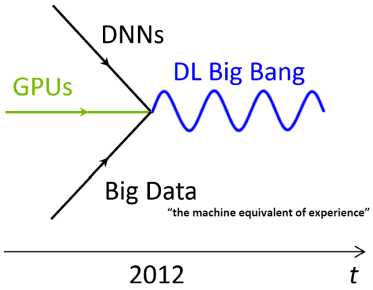


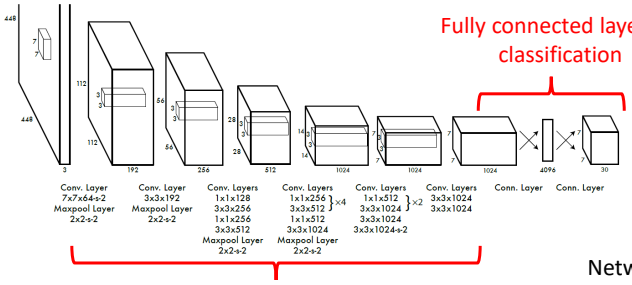
Image © NVIDIA

25

25

Deep learning – what?

- What? Step away from **manual** feature development, and let algorithm determine important features
- Feed images through a convolutional neural network (CNN), mostly consisting of convolution layers, max pooling layers and fully connected layers



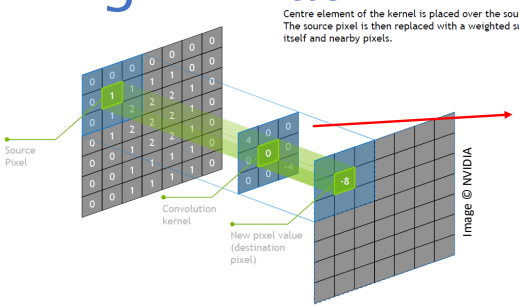
Network: YoloV1

26

26

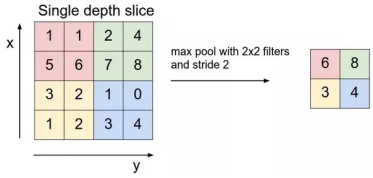
Deep learning – what?

- Convolution:



Kernel weights and interconnections are learned during training

- Max-pooling (subsampling):

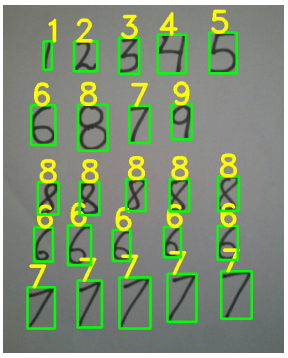


27

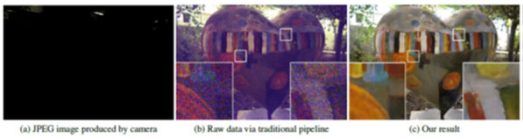
27

Power of Deep Learning?

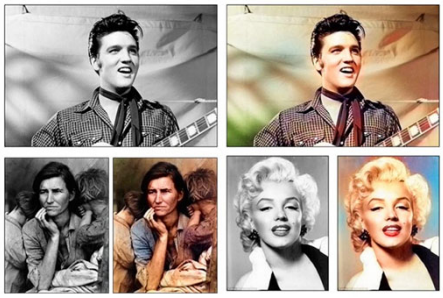
You can teach a computer whatever you want!



Handwriting recognition



Seeing in the dark



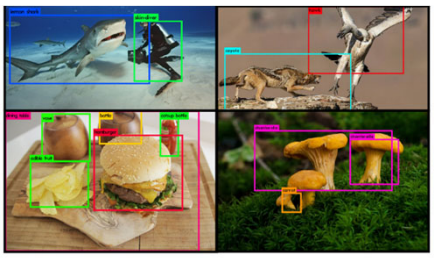
Colorization of grayscale photos

28

28

Power of Deep Learning?

You can teach a computer whatever you want!



Object detection

Image content description

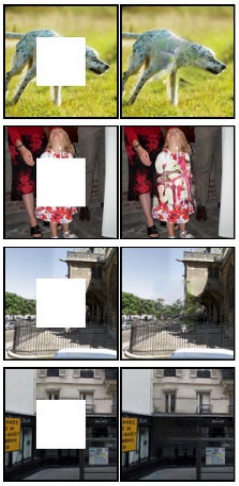
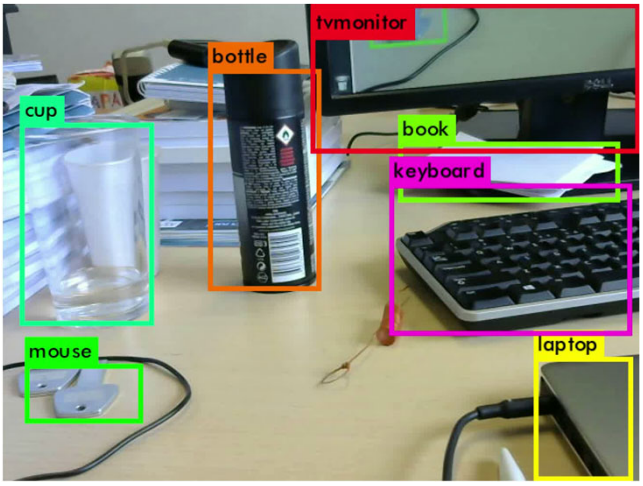


Image inpainting

29

29

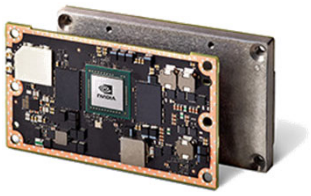
Examples in image processing



30

Embedded edge hardware?

- Example: person detection on drone with on-board processing



Jetson TX2
256 NVIDIA CUDA cores
Quad ARM A57 CPU
8 GB LPDDR4 60 GB/s memory
Power consumption: 15W
Cost: 500 USD

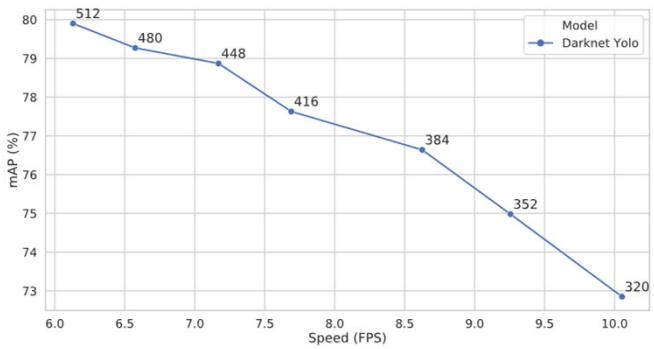
31

31

Highly optimized implementations

- Standard Darknet on Jetson TX2
- C codebase, developed by Joseph Redmon

416 x 416
Jetson TX2
Darknet
7.7 FPS

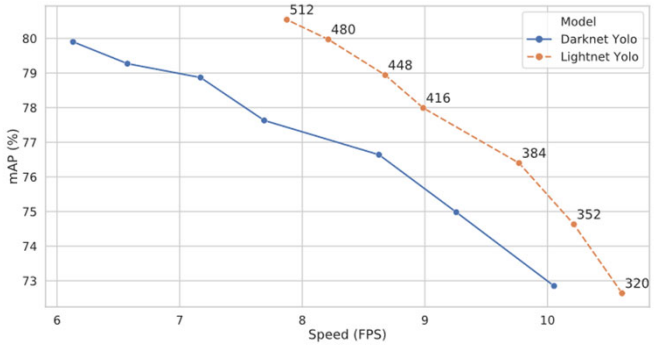


32

32

Highly optimized implementations

- LightNet: Python codebase
- Developed by EAVISE
- Built on top of PyTorch



416 x 416
Jetson TX2

Darknet
7.7 FPS

x 1.2

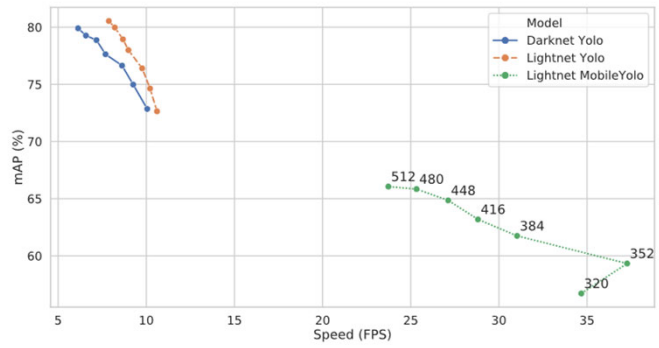
Lightnet
9.0 FPS

33

Highly optimized implementations

- MobileNet

Split up a convolution in a depthwise and pointwise part



416 x 416
Jetson TX2

Darknet
7.7 FPS

x 1.2

Lightnet
9.0 FPS

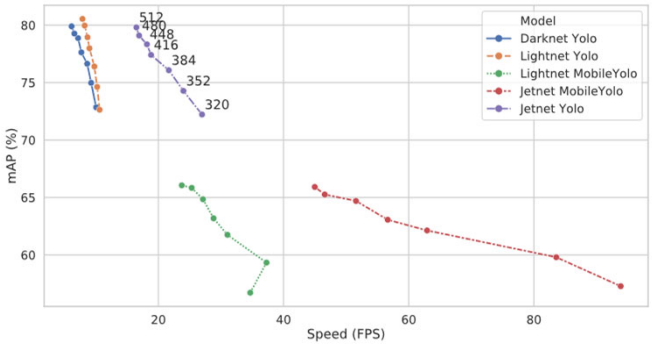
x 3.2

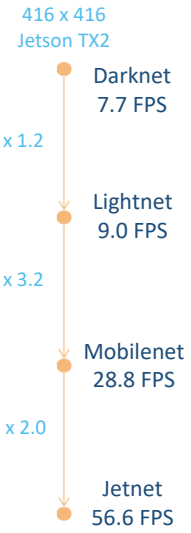
Mobilenet
28.8 FPS

34

Highly optimized implementations

- JetNet: C++ codebase
- Developed by EAVISE
- Built on top of TensorRT





35

Research questions

- What are the possibilities and applications for Deep Learning on low-cost embedded hardware?
- Which hardware is suited for these Deep Learning applications, and which trade-offs need to be made?
- What are the available software libraries and frameworks, and how do they work?
- What is the influence on the accuracy when using such embedded devices?
- What are the consequences for the power usage of the system?
- How can we improve the latency of the embedded system using local autonomous decisions?

36

36

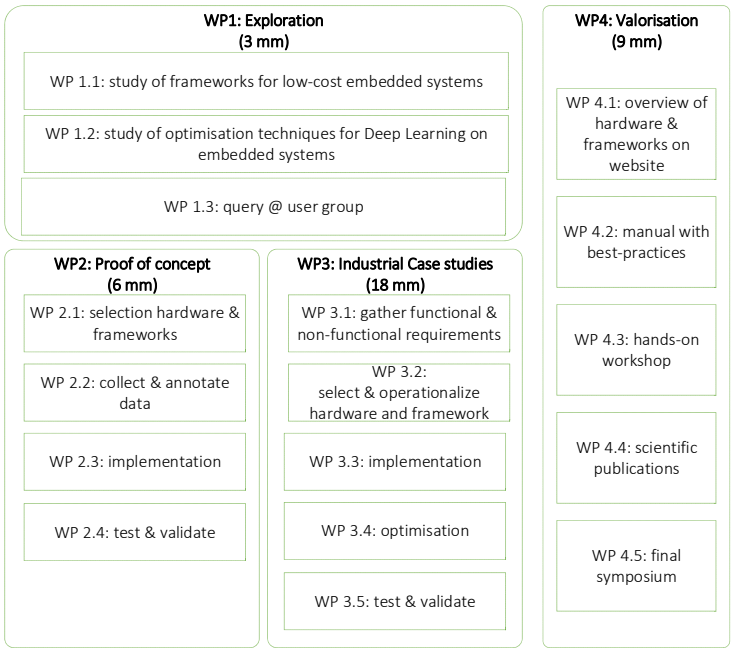
Project goals

- TETRA = Technology Transfer
  - Use cases
    - 1 Academic
    - 4 to 6 Industrial cases
  - Workshop
  - Manual

37

37

Work Packages

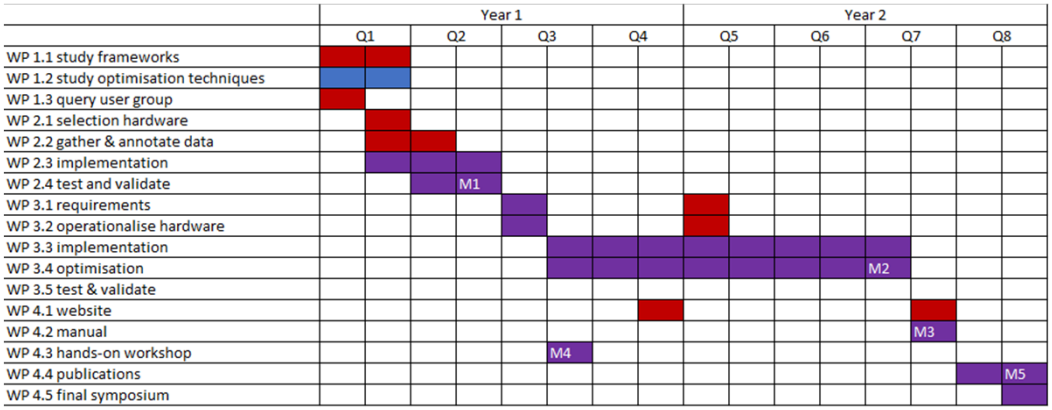


38

38

Planning

Start 01/03/2020, duration 2 years



39

39

Members of the user group

6WOLVES – Digipolis – DP Technics – DSP Valley– E.D.&A. – Edgise – Melexis Technologies – Picanol – Qmineral – ScioTeq – Sensotec – Transport & Mobility Leuven

40

40

6WOLVES

- AI & Machine Learning in sport
- Machine Learning algorithm
- Applied on a low-cost, low-power electronic system
- Exploring the capabilities of edge computing

41

41

Digipolis intro

- Digipolis
- Enterprise Architects
- BI en Advanced Analytics
- Innovatie-projecten
  - Data Science
  - Vision

42

Digipolis intro

- Altruis
  - PIO project
  - Edge AI
  - mobility
- Sluikstortherkenning
- ACO (Antwerp City Observer)
  - ANPR and more

43

DP Technics

- R&D for IoT applications
  - BlueCherry.io
- Edge computing

44

44

DSP Valley

- Technology cluster smart electronic systems

45

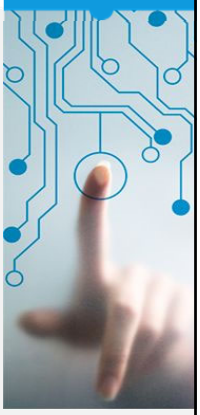
45

E.D.&A.

Electronics, Development & Assembly

The power to control

- Founded in 1981
- HQ Located in Belgium, Kalmthout (Antwerp)
- Sales office in Bonn, Germany
- 85+ employees, 38 in development
- E.D.&A. develops and produces custom-made electronic controllers for machines and appliances.
  - Industrial & Consumer market
  - Areas: HVAC, Water, Laundry, Food, Agriculture



ed&a

E.D.&A nv | Franseweg 20 -2920 Kalmthout (Belgium) | Bornheimerstrasse 127 – 53119 Bonn (Germany) | www.edna.eu

46

History

1981 – 2019

The power to control

- 1981: Foundation of E.D.&A. by Flor D'Handschotter
- 1983-1996: Development, production and sales of PLC-systems and custom-made controllers.
- 1997: Strategy change: Exclusive focus on the professional OEM-market, with custom-made controllers.
- 2009: Gert D'Handschotter takes over E.D.&A.
  E.D.&A. remains a family owned company.
- 2012: Start-up of an automatic assembly line in-house.
- 2016: E.D.&A. obtains Factory of the Future Award and ISO 9001:2015.
- 2018: E.D.&A. receives label Belgium's Best Managed Company.
- 2019: E.D.&A. obtains again Factory of the Future Award and ISO 9001:2015.

E.D.&A nv | Franseweg 20 -2920 Kalmthout (Belgium) | Bornheimerstrasse 127 – 53119 Bonn (Germany) | www.edna.eu

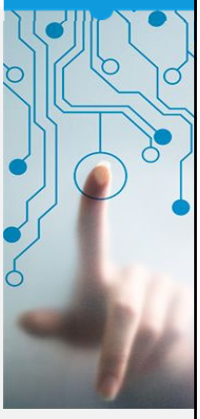
47

AI@EDGE

E.D.&A. interest in AI

The power to control

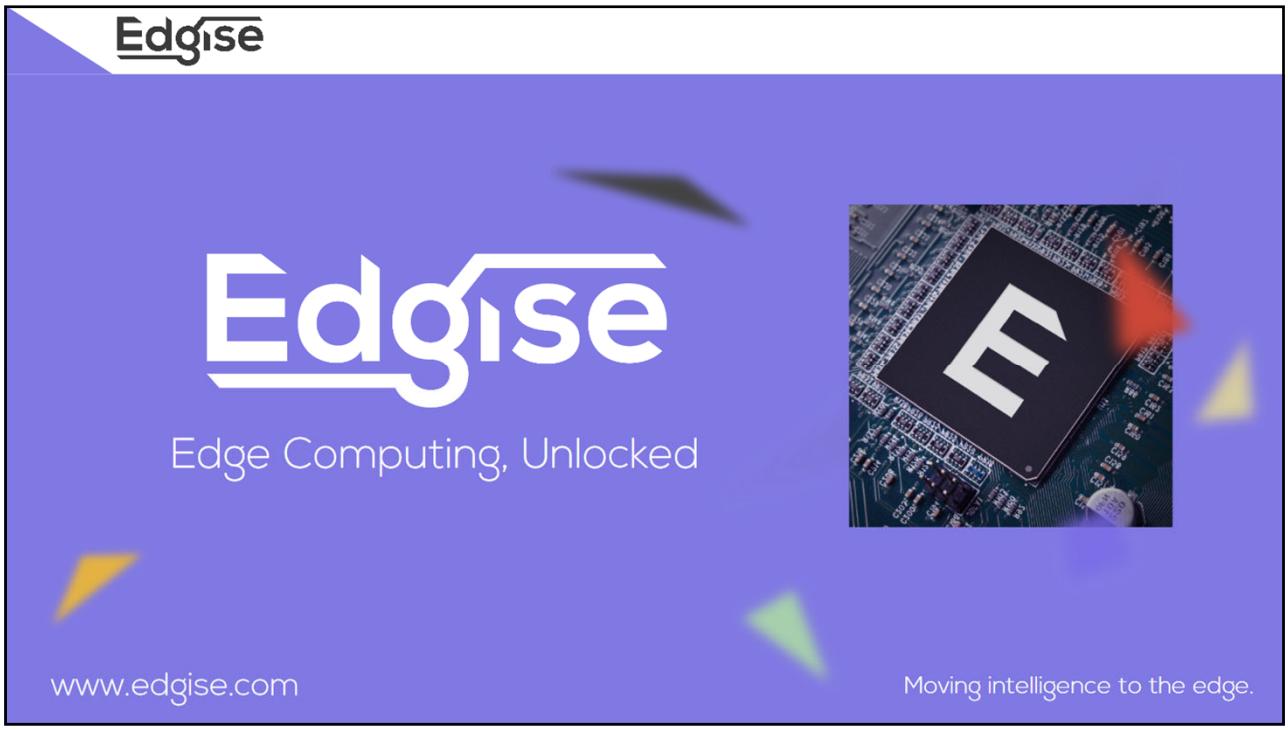
- General interest in feasibility and added value of AI/NN for replacement of algorithms / control functions that run on smaller microcontrollers or application processors
- Fall 2018 until spring 2019: B.Sc. Student and E.D.&A. engineers jointly worked on Binary Neural Network core suited to run on Cortex M4 platforms (case: handwritten number recognition).
  - Fast cycle time (< 10 ms on Cortex M4 @ 80Mhz)
- Summer 2019: Internal project for capacitive sense with Neural Network instead of capsense software library.
  - Goal: increase reliability of cap sense operation under stressed operating conditions (moisture, EMI)
  - Promising results but further research required



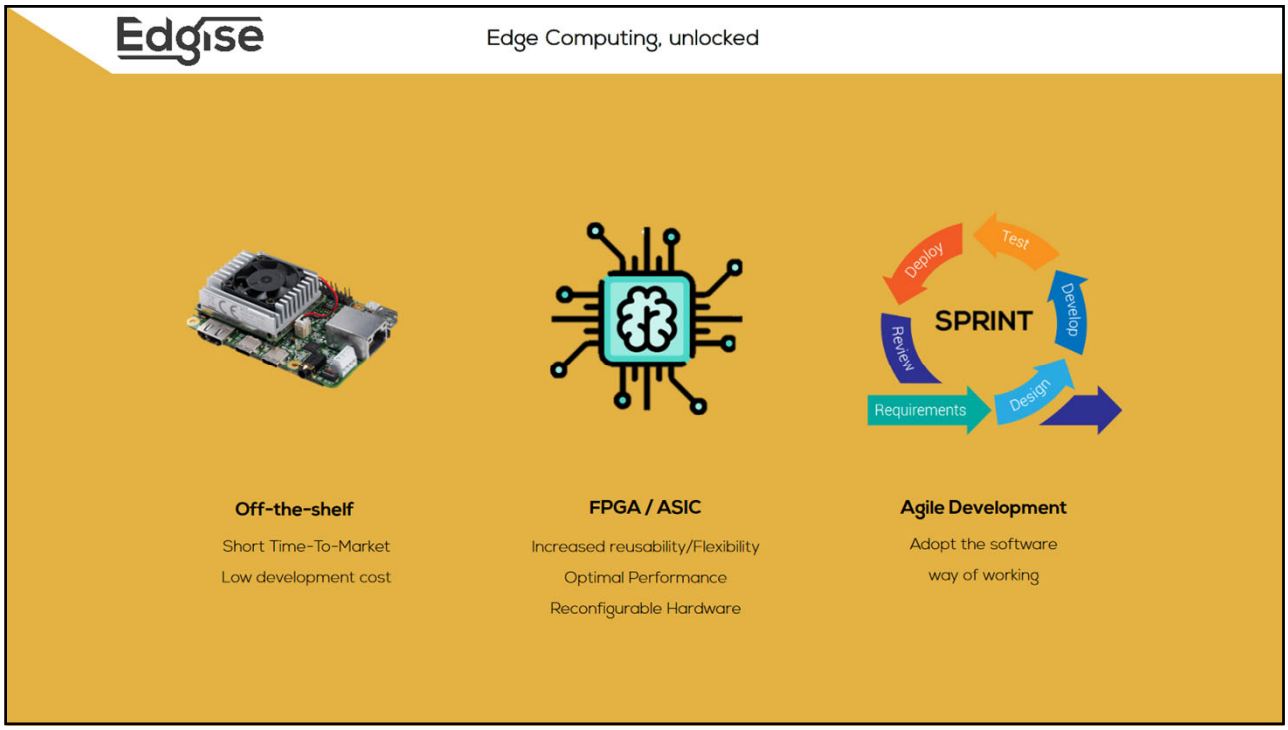
ed&a

E.D.&A nv | Franseweg 20 -2920 Kalmthout (Belgium) | Bornheimerstrasse 127 – 53119 Bonn (Germany) | www.edna.eu

48


Edgise
Edgise
Edge Computing, Unlocked
www.edgise.com
Moving intelligence to the edge.

49


Edgise
Edge Computing, unlocked
Deploy
Test
Develop
Review
SPRINT
Requirements
Design
Off-the-shelf
Short Time-To-Market
Low development cost
FPGA / ASIC
Increased reusability/Flexibility
Optimal Performance
Reconfigurable Hardware
Agile Development
Adopt the software
way of working

50

Edgise Edge Computing, unlocked

AI@TETRA for us

1. Mapping the Edge landscape
   - Frameworks
   - Tools
   - Hardware
   - AI Models
   - Best practices
   - Quantization schemes
   - Pruning possibilities
2. Use cases (Microcontroller/FPGA)
   - Low power object detection
   - Predictive maintenance
   - Sensor fusion
   - Low power Speech / context recognition

51

Melexis Technologies

- Electronic components for car industry
- Apply AI to build better sensors with more functionality

vives KU LEUVEN AI@EDGE 52

52

Picanol

- Development and production of weaving machines

53

53

Qmineral

- Mineralogy analysis
- Continuation of Start to deep learn

54

54



ScioTeq presentation & interests

AI@EDGE - 03/03/2020

Edouard Charvet – edouard.charvet@scioteq.com

55

History

1976 | 1980's | 1989 | 1990's | 1999 | 2001 | 2003 | 2005 | 2011 | 2013 | 2015 | 2019

- 1976: Chromatics was founded in Atlanta Georgia as a spin off or the ISC corporation
- 1980's: **Barco focuses on niche industrial markets with COTS CRT technology**
- 1989: Acquisition of Chromatics, in Duluth, giving access to high end markets in Rugged and ATC
- 1990's: COTS CRTs evolved into LCD technology, enabling the branch off into ATC and winning 28,000 RFDs for US Navy's AN/UYQ-70 program (started 1994)
- 1999: Introduction of first square 28" LCD display for ATC market, winning Jane's innovation of the year award in 2001
- 2001: Acquisition of Texen, in Toulouse, giving access to console and rugged computer business
- 2003: Airbus A380 auxiliary cockpit displays contract; Barco's entry in Civil Airlines
- 2005: F-35 JSF 5th generation fighter touchscreen contract
- 2011: Barco combines Rugged, Avionics ATC and Training BU's into DAT division
- 2013: Win of COMAC C919 Airlines; Fueling Barco's success in Civil Airlines
- 2015: **Esterline acquires Barco's Defense & Aerospace division (DAT)**
- 2019: **Esterline Advanced Displays rebrands as ScioTeq**

56

56

Activities

ScioTeq

Avionics Displays

Rugged Solutions

Air Traffic Control Solutions

57

57

Capabilities

ScioTeq

- Optical Stacks
  - Optics
  - LCD glass ruggedization
  - Optical Bonding & Sealing
  - Illumination technology
- Video/graphic processing
  - Graphics processing engines
  - OpenGL
  - High speed video transmission
  - Video treatment
- Computing
  - Latest multicore processing
  - Real Time Operating Systems
  - Modular Operating System Architecture middleware - MOSArt™
- Software Applications
  - PFD, NAV displays, etc.
  - Synthetic Vision System
  - Customized applications
- Extensive in-house testing
  - Environmental testing (temperature, humidity, shock, vibrations, pressure, ...)
  - EMI/EMC testing
  - Reliability (MTBF calculations, HALT)
- Support and ILS
  - Long term
  - World-wide network
  - Obsolescence management

58

58

Interests in AI@EDGE

ScioTeq

- There is a future towards introduction of AI in avionics systems
  - On-going EUROCAE working group on certification of AI systems (focused on offline supervised machine learning)
  - We see a future need for being capable of hosting such AI applications (Neural Net based) on our computing platforms.
- Avionics systems have constraints in terms of
  - Power
  - Weight
  - Size
- Which drive constraints in terms of available
  - Memory
  - Computing capabilities

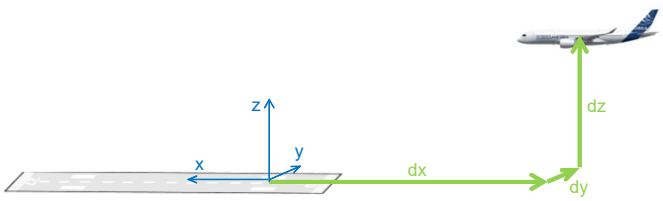
59

59

Use case proposal

ScioTeq

- Recognition of airport environment during the aircraft landing phase
  - Based on Neural Nets (architecture TBD) inferences on embedded system
  - Offline learning (not on embedded system)
  - Recognition of buildings, runways, taxiways, obstacles…
  - Real-time constraints: latency < 200ms
- Allowing to compute 6D pose of the aircraft in the runway referential
  - Pitch, roll, heading
  - Lateral deviation, height, distance to touch down point



60

60

Sensotec

- Development of assistive technology for persons with a reading or writing impairment
- Applications on language support for low-cost embedded systems

61

61

TRANSPORT & MOBILITY LEUVEN

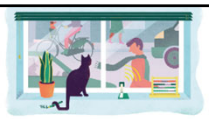
- **Applied research to support policy decisions**
- **Mission**: to help society by offering scientifically sound analyses
  - Quantitative research: modelling, statistical analyses, simulations, and prognoses.
- **Multidisciplinary team** of 28 people
  - Civil engineers, computer scientists, psychologists, economists, environmental experts, etc.
- **Topics**:
  - Traffic management, urban mobility, sustainable mobility, traffic safety, transport emissions, economics and pricing, freight transport (rail, maritime, trucks), smart mobility, bicycles, etc.

62

Dr. Péter I. Pápics (TML) - 3 March 2020 - AI@EDGE TETRA

62

TELRAAM .be TRANSPORT & MOBILITY LEUVEN MOBIEL21 ZET MENSEN IN BEWEGING waanz.in



- **Motivation**:
  - Precise traffic counts are essential for transport-related studies, but cost- and resource-intensive
  - Large, practical-question-driven demand from the public
- **Solution**: integrated application: low-cost hardware (RPI 3A+ & camera, OpenCV, python) & public online platform (telraam.net: visualisations, analysis, networks)
  - Pros: high density, continuous, multimodal, speed, cost-efficient
  - To-be-improved (tech-side): classification and object tracking
  - Active since 21 March 2019:
    75 million+ objects counted, 550 active cameras, 7 countries...
- **Target audience**: Citizens, Professionals, Policymakers
  - Flemish -> Belgian -> European (H2020) exposure WECOUNT

Dr. Péter I. Pápics (TML) - 3 March 2020 - AI@EDGE TETRA

63

63

- **Potential for AI@EDGE:**
  - Improved object tracking and classification on the edge devices using AI. (Hardware might change, but CPU architecture will stay the same.) -> Challenge: frame rate (30 FPS) & accuracy (tradeoff).

Dr. Péter I. Pápics (TML) - 3 March 2020 - AI@EDGE TETRA

64

64

Poll & Discussion

- Poll everywhere

65

65

TETRA

- Technology Transfer project
  - Co-financing: VLAIO 92,5% – 7,5% user group
- Intellectual Property

66

66

Administration

- Rules of procedure (reglement van orde)
- VLAIO User Poll
- Next user group meeting

67

67