DESIGN, AUTOMATION & TEST IN EUROPE
DATE 23
17 – 19 April 2023 · Antwerp, Belgium
The European Event for Electronic System Design & Test

This project was partially funded by the Autonomous Robotics Research Center of the Technology Innovation Institute
Technology Innovation Institute

Bio-inspired Autonomous Exploration Policies with CNN-based Object Detection on Nano-drones

Lorenzo Lamberti[1], Luca Bompani[1], Victor Javier Kartsch[1], Manuele Rusci[2], Daniele Palossi[3,4], Luca Benini[1,4]

[1]DEI - University of Bologna, Italy
[3]IDSIA - USI and SUPSI, Switzerland
[2]KU Leuven, Belgium
[4] IIS - ETH Zurich, Switzerland

ALMA MATER STUDIORUM UNIVERSITÀ DI BOLOGNA
KU LEUVEN
USI/SUPSI IDSIA Istituto Dalle Molle di studi sull'intelligenza artificiale
ETH zürich

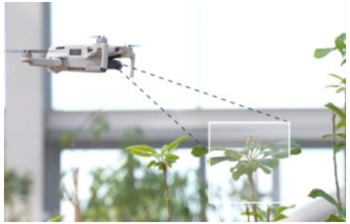
Why autonomous nano-UAVs?

Huge interest in **autonomous** (i.e., no external infrastructure) unmanned aerial vehicles (UAVs). Relevant applications in civil and industrial use cases:

Rescue missions

Surveillance, inspection

Precision agriculture*



Entertainment

* Image from Technology Innovation Institute, Abu Dhabi

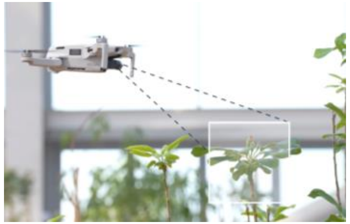
Why autonomous nano-UAVs?

Huge interest in **<u>autonomous</u>** (i.e., no external infrastructure) unmanned aerial vehicles (UAVs). Relevant applications in civil and industrial use cases:

Rescue missions

Surveillance, inspection

Precision agriculture*



Entertainment

Why autonomous <u>nano-sized</u> UAVs?

- Enhanced safety → e.g., Human Robot Interaction (HRI)
- New use cases: indoor locations, ubiquitous IoT, etc.
- Reduced costs

Standard-sized UAV

Nano-sized UAV

* Image from Technology Innovation Institute, Abu Dhabi

Autonomous nano-UAVs: challenges & goals

Nano-drones challenges:

- Small form factor (~10cm)
- Limited payload (~15g)
- Limited computing power budget (<100mW)

UAV	Standard-sized	Nano-sized
Size [⌀, weight]	~50cm / ~ few Kg	~10cm / ~50g
Tot. Power	~ 100 W	~ 5W
Processing device	High-end CPU	Low-power MCU

Our goal for autonomous nano-drones:

Autonomous nano-UAVs: challenges & goals

Nano-drones challenges:

- Small form factor (~10cm)
- Limited payload (~15g)
- Limited computing power budget (<100mW)

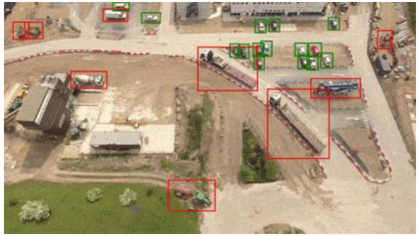
UAV	Standard-sized	Nano-sized
Size [⌀, weight]	~50cm / ~ few Kg	~10cm / ~50g
Tot. Power	~ 100 W	~ 5W
Processing device	High-end CPU	Low-power MCU

Our goal for autonomous nano-drones:

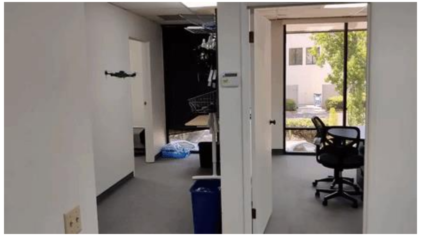
- Multi-tasking perception (as standard-sized UAVs, and biological systems [1])
- Real-time requirements

Multi-tasking biological systems [1]

Object detection



Exploration



[1] M. Giurfa and R. Menzel, “Insect visual perception: complex abilities of simple nervous systems” Current Opinion in Neurobiology, vol. 7, no. 4, pp. 505–513, 1997.

State-of-the-Art: autonomous nano-drones

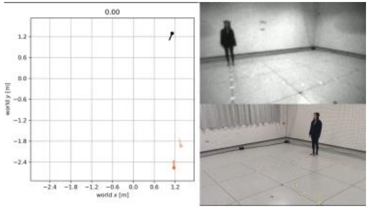
Perception on autonomous nano-drones: single tasks

Visual navigation

[2] PULP-Dronet

CNN for navigation and obstacle avoidance

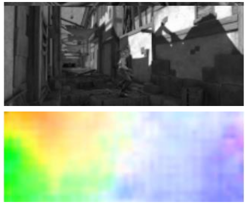
Pose estimation



[3] PULP-Frontnet

CNN for human pose estimation

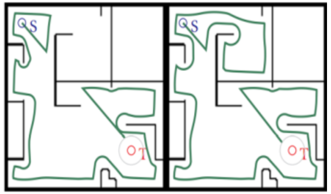
Optical flow



[4] NanoFlowNet

CNN for dense optical flow. Goal: obstacle avoidance

Exploration



[5] Bug navigation algorithms

Lightweight state machine for ranging-based exploration.

Object detection

No fully onboard execution for nano-drones

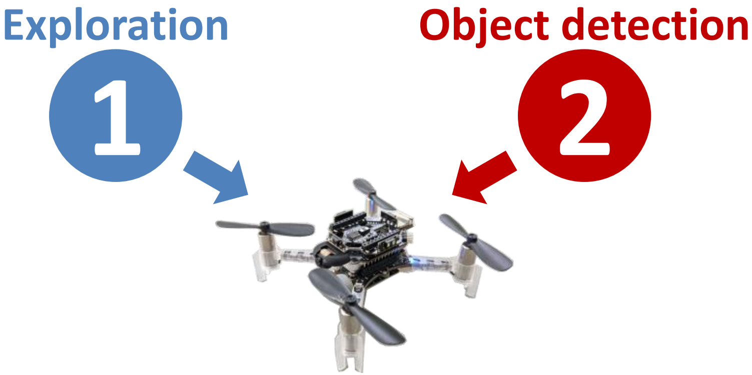
All these works deployed individual tasks on the nano-drone
→ Not targeting multi-tasking perception

[2] L. Lamberti et al., "Tiny-PULP-Dronets: Squeezing Neural Networks for Faster and Lighter Inference on Multi-Tasking Autonomous Nano-Drones"
[3] D. Palossi et al., "Fully Onboard AI-Powered Human-Drone Pose Estimation on Ultralow-Power Autonomous Flying Nano-UAVs"
[4] R. Bouwmeester et al., '' NanoFlowNet: Real-time Dense Optical Flow on a Nano Quadcopter''
[5] K. Mcguire et al., '' A comparative study of bug algorithms for robot navigation''

Our contribution

Enabling multiple perception tasks aboard an autonomous nano-drone

1. Exploration of an unknown environment;
2. Object detection.

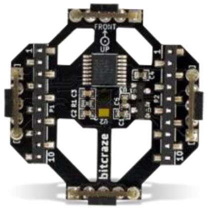


Robotic Platform: host + multi-core

Hardware

Crazyflie 2.1

+



Multi-ranger deck

+

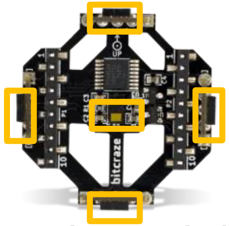
AI-Deck

=

Robotic Platform: host + multi-core

Hardware

Crazyflie 2.1



Multi-ranger deck

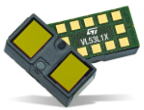
AI-Deck

Host

STM32F4 MCU
(<100 MMAC/s)

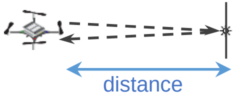
Control-based tasks
Sensor interfacing

5x ToF sensors



VL53L1x

Time-of-flight ranging sensors

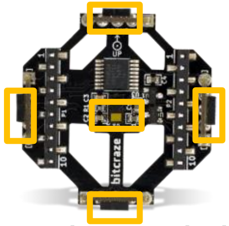


Robotic Platform: host + multi-core

Hardware

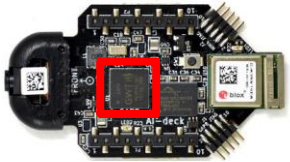
Crazyflie 2.1

+



Multi-ranger deck

+



AI-Deck

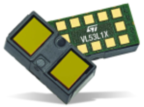
=

Host

STM32F4 MCU
(<100 MMAC/s)

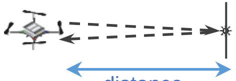
Control-based tasks
Sensor interfacing

5x ToF sensors



VL53L1x

Time-of-flight ranging sensors



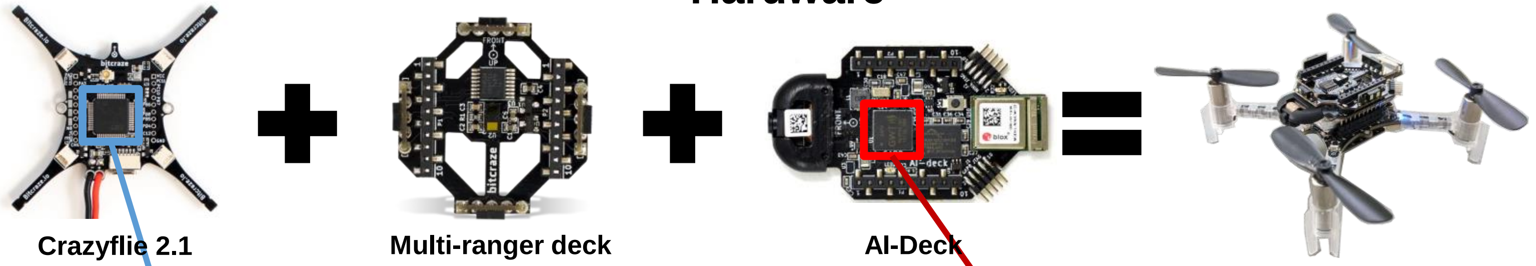
Multi-core MCU

GAP8 SoC
(~1 GMAC/s)

8 parallel ULP cores
QVGA camera

Tasks and mapping

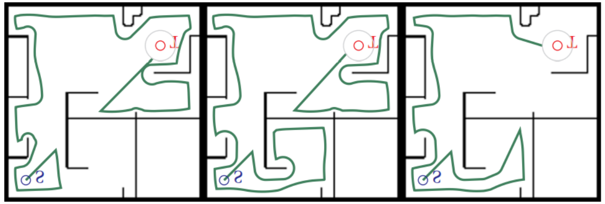
Hardware



1 Exploration

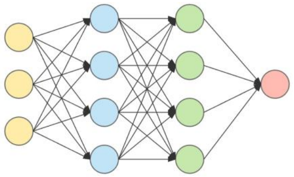
Goal: exploring unknown environment

Algorithm: bio-inspired exploration policy*

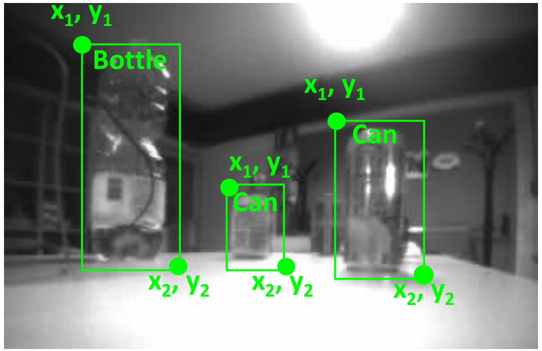


2 Object Detection

CNN detector for 2 oject classes



CNN outputs
- Bounding box
- Object class
- Confidence



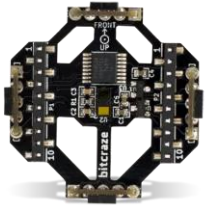
Exploration

Obstacle avoidance: ranging-based sensor

Hardware

Crazyflie 2.1

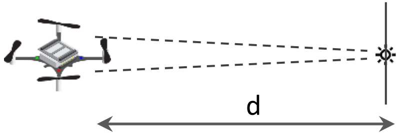
+



Multi-ranger deck

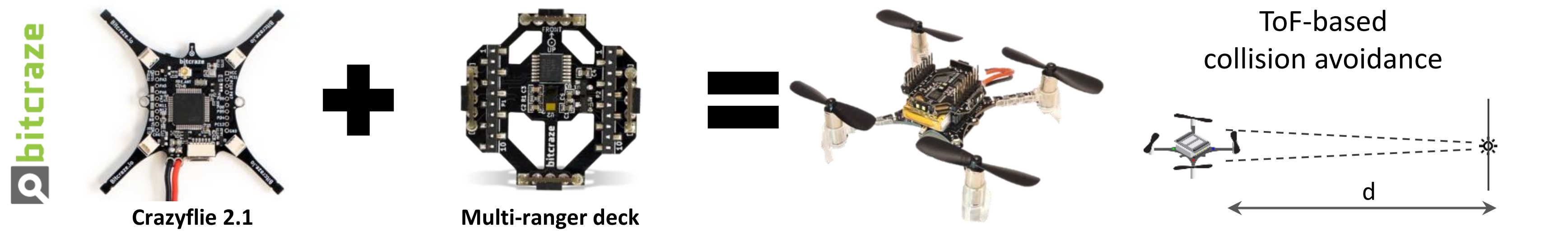
=

ToF-based
collision avoidance



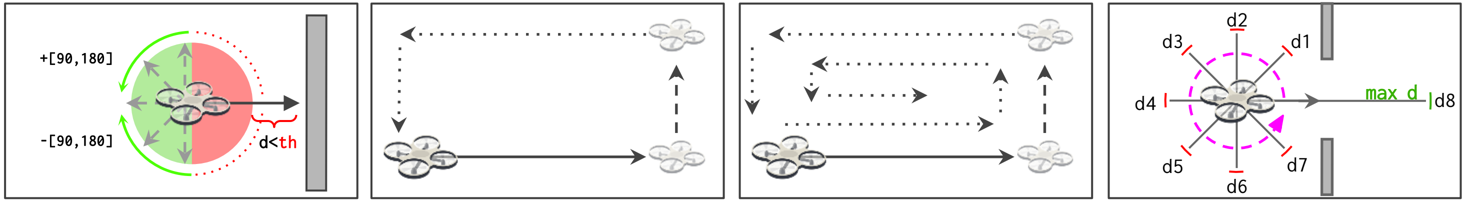
Obstacle avoidance: ranging-based sensor

Hardware



We compare 4 **exploration policies**

Pseudo-random | **Wall-following** | **Spiral** | **Rotate-and-measure**

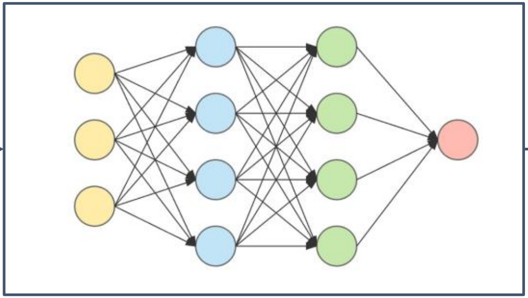


lightweight state machines

Object detection

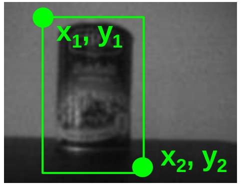
Convolutional neural network [6]:

Input

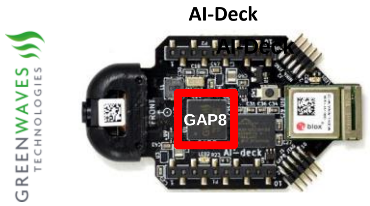


Mobilenet v2 + SSD-lite

Output



Hardware

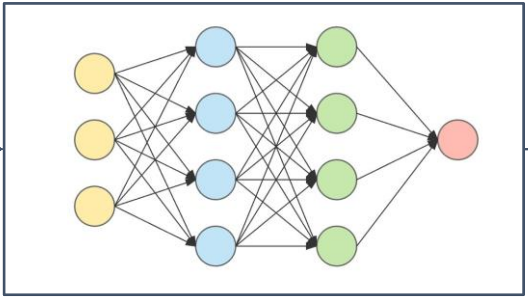


[6] Huang J *et al.* "Speed/accuracy trade-offs for modern convolutional object detectors." In *Proceedings of the IEEE CVPR*, pp. 7310-7311. 2017.

Object detection

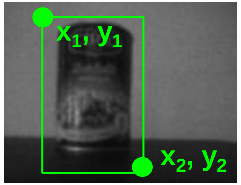
Convolutional neural network [6]:

Input



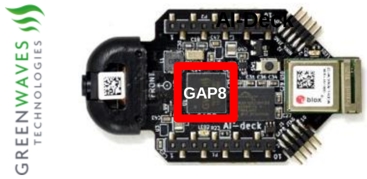
Mobilenet v2 + SSD-lite

Output



Hardware

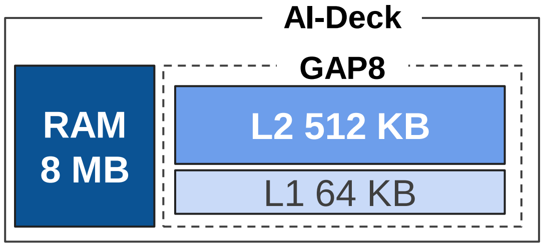
AI-Deck



CNN stats:

	CNN [6]
Parameters	**4.8M**
Operations/inference	483M MAC

Deployment: 8-bit weights and activations
Quantization aware fine-tuning

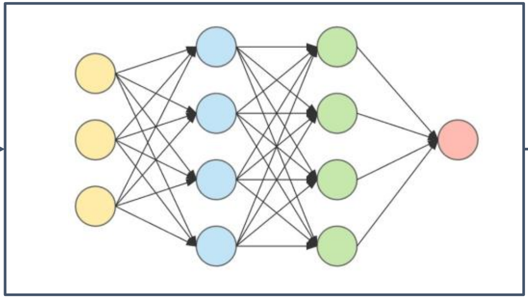


[6] Huang J *et al.* "Speed/accuracy trade-offs for modern convolutional object detectors." In *Proceedings of the IEEE CVPR*, pp. 7310-7311. 2017.

Object detection

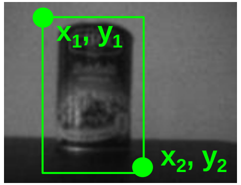
Convolutional neural network [6]:

Input



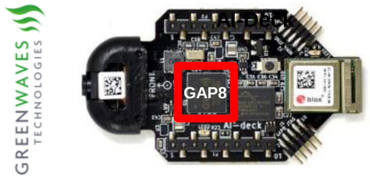
Mobilenet v2 + SSD-lite

Output



Hardware

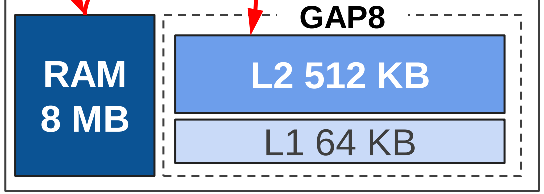
AI-Deck



CNN stats:

	CNN [6]
Parameters	**4.8M**
Operations/inference	483M MAC

Deployment: 8-bit weights and activations
Quantization aware fine-tuning

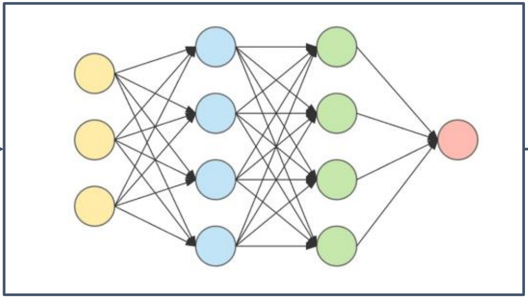


[6] Huang J *et al.* "Speed/accuracy trade-offs for modern convolutional object detectors." In *Proceedings of the IEEE CVPR*, pp. 7310-7311. 2017.

Object detection

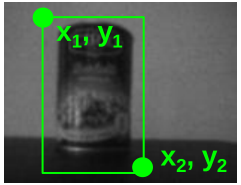
Convolutional neural network [6]:

Input

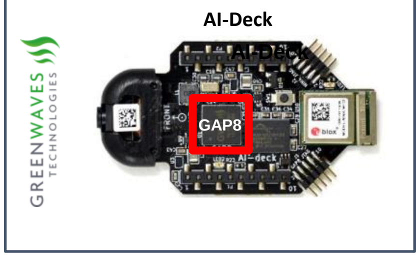


Mobilenet v2 + SSD-lite

Output



Hardware

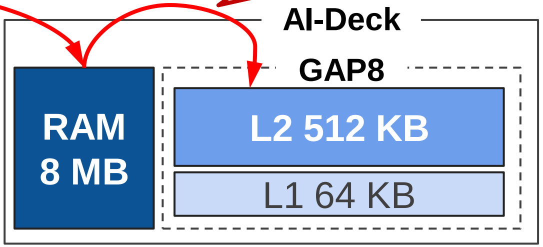


Feasible, but introduces **memory overhead**

CNN stats:

	CNN [6]
Parameters	**4.8M**
Operations/inference	483M MAC

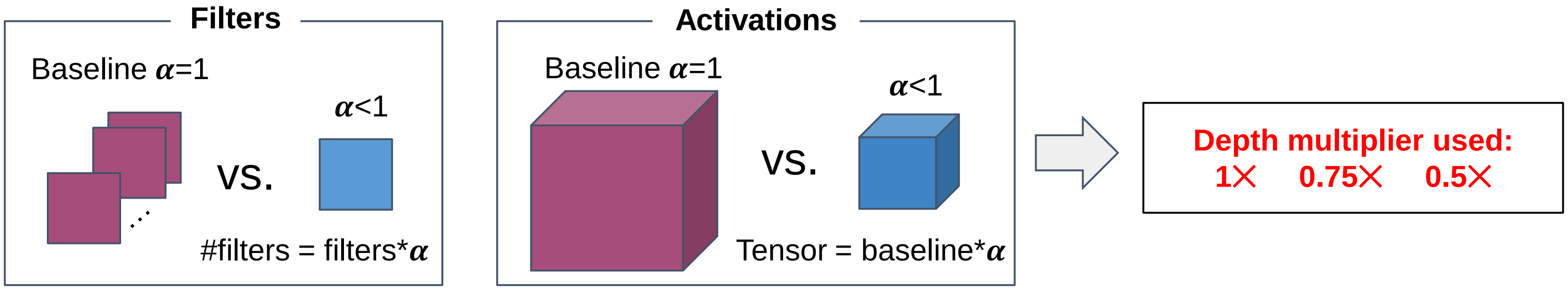
Deployment: 8-bit weights and activations
Quantization aware fine-tuning



[6] Huang J *et al.* "Speed/accuracy trade-offs for modern convolutional object detectors." In *Proceedings of the IEEE CVPR*, pp. 7310-7311. 2017.

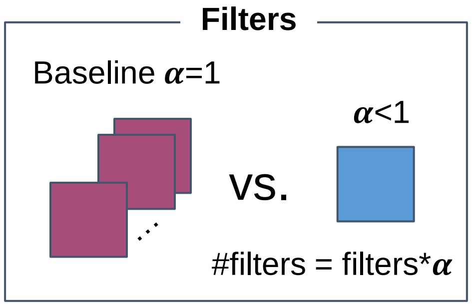
CNN optimization and training

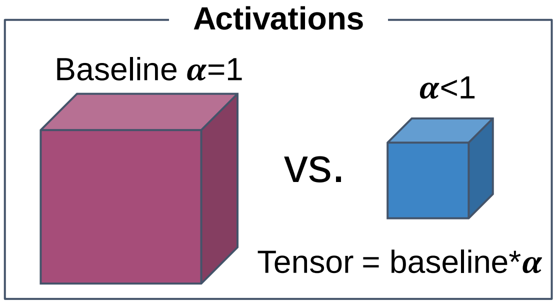
Optimization:
The depth multiplier (α) is an hyperparameter that modifies the number of filters and output channels of each layer.



CNN optimization and training

Optimization:
The depth multiplier (α) is an hyperparameter that modifies the number of filters and output channels of each layer.





Depth multiplier used:
1× 0.75× 0.5×

Training pipeline:

Training/Testing: OpenImages Dataset → Augmentation → Dataset Balancing → Himax fine-tuning

2-class training:
- bottles
- tin cans

Augmentation: flip, resize, brightness, grayscale

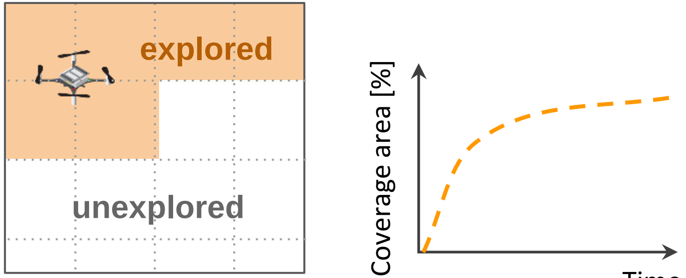
Balancing the number of images for dataset classes

Training/testing dataset: OpenImages + Himax

Evaluation metrics

Exploration policies

Metric: coverage area



Environment discretization

Object detection

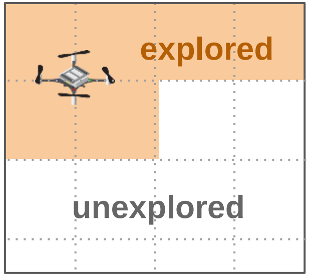
Metric: mean average precision (mAP)

In-field assessment

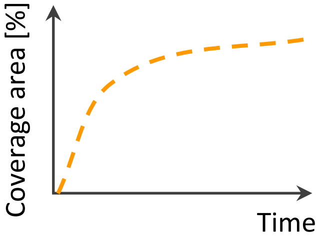
Evaluation metrics

Exploration policies

Metric: coverage area



Environment discretization



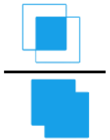
Object detection

Metric: mean average precision (mAP)

mAP

$$Precision = \frac{TP}{TP+FP}$$

$$IoU = \frac{area\ of\ Overlap}{area\ of\ Union}$$

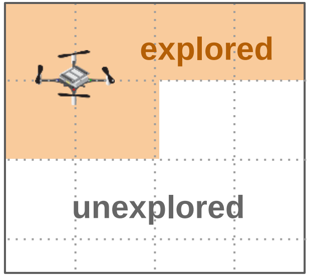


In-field assessment

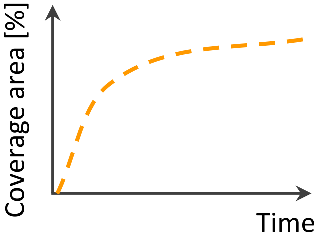
Evaluation metrics

Exploration policies

Metric: coverage area



Environment discretization



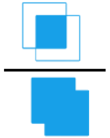
Object detection

Metric: mean average precision (mAP)

mAP

$$Precision = \frac{TP}{TP+FP}$$

$$IoU = \frac{area\ of\ Overlap}{area\ of\ Union}$$

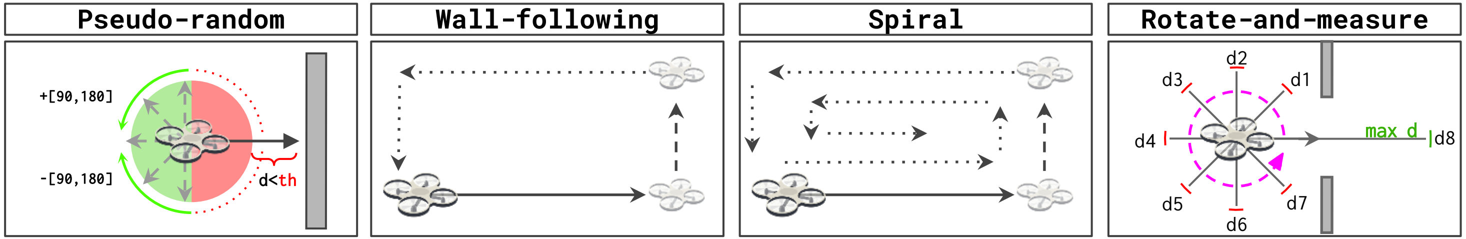


In-field assessment

$$\text{Detection Rate} = \frac{\#\text{Detected Objects}}{\#\text{Objects}}$$

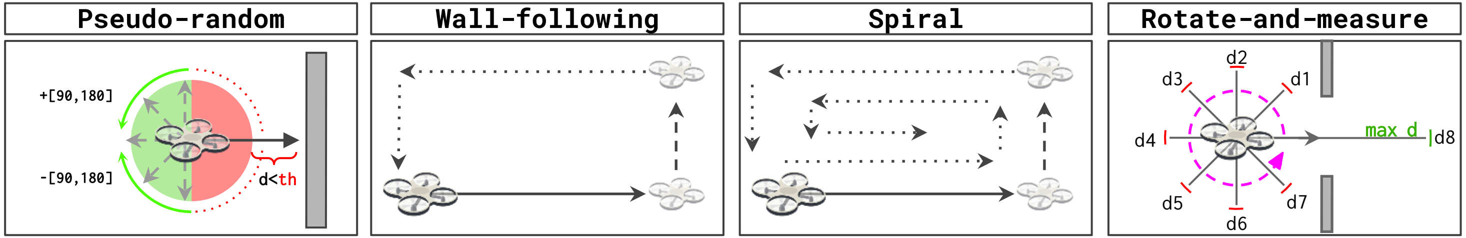
Evaluation: exploration policies

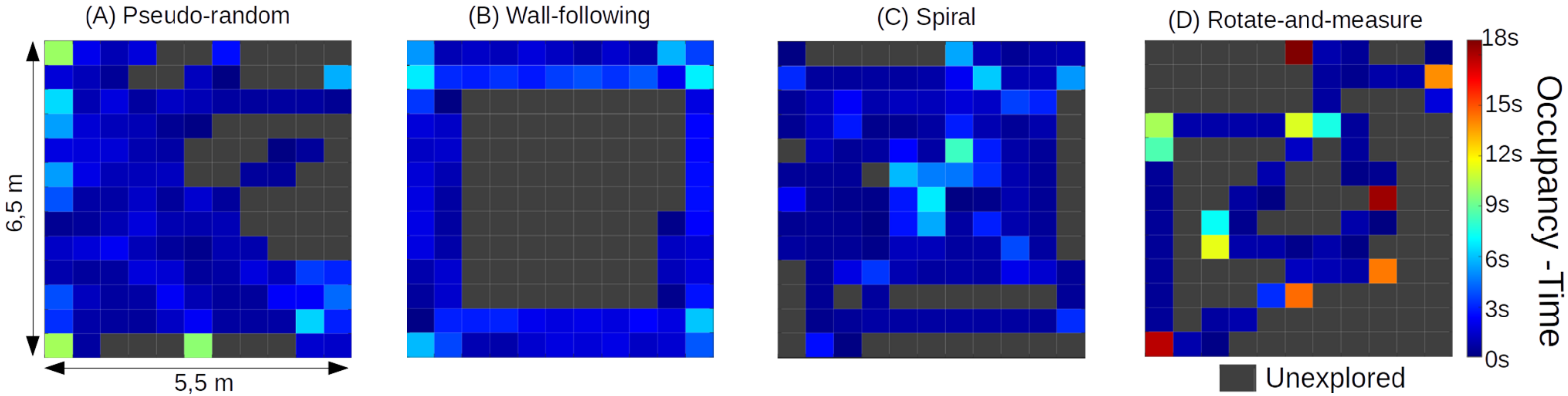
Test setup: 5min flight, 3 speeds, 4 policies. Results shown for avg. speed = 0.5 m/s



Evaluation: exploration policies

Test setup: 5min flight, 3 speeds, 4 policies. Results shown for avg. speed = 0.5 m/s

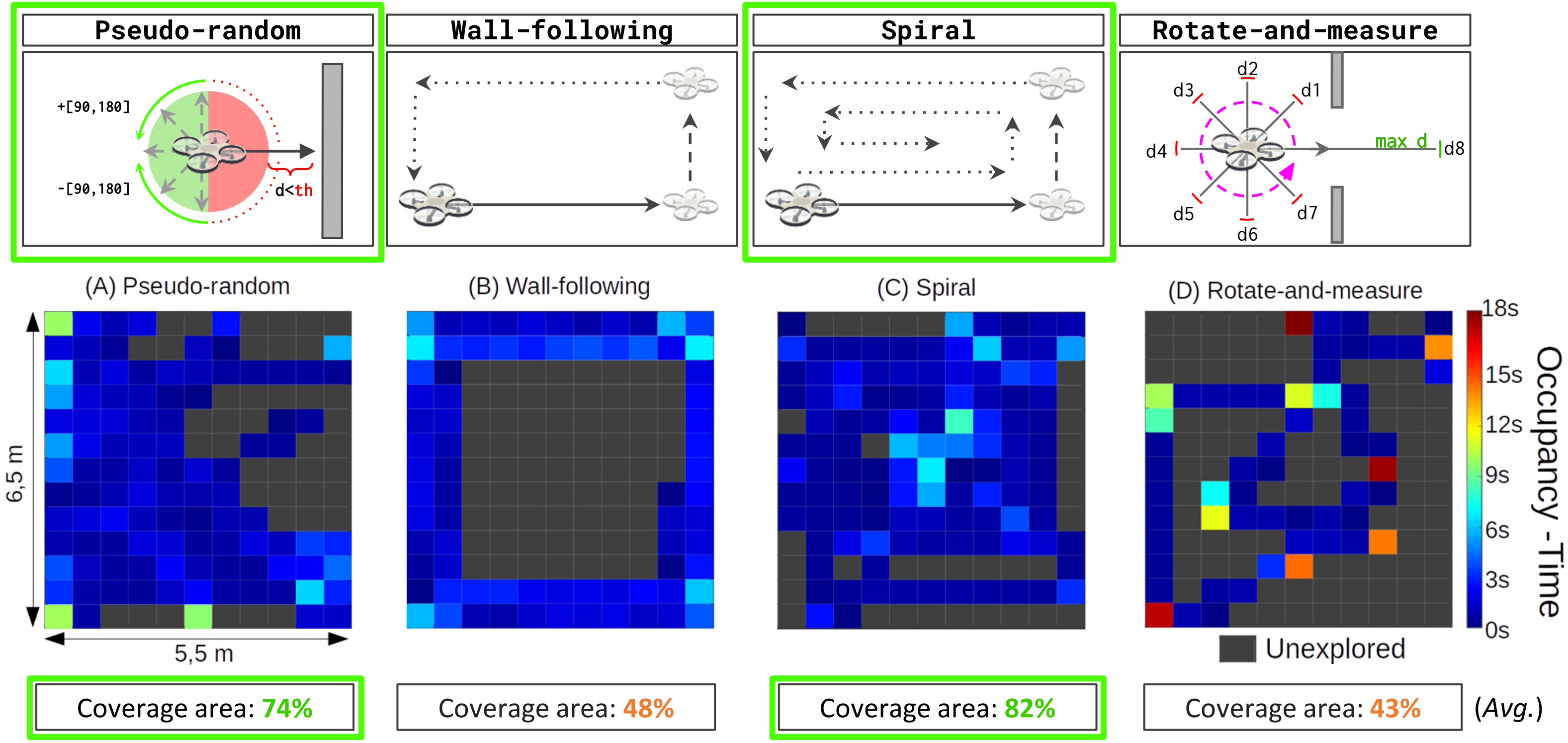




Coverage area: **74%**	Coverage area: **48%**	Coverage area: **82%**	Coverage area: **43%**	(*Avg.*)

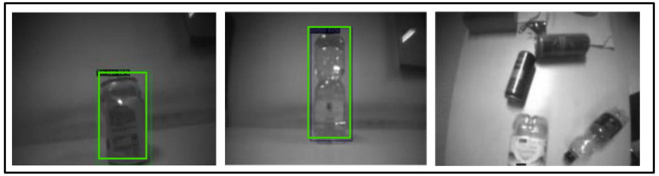
Evaluation: exploration policies

Test setup: 5min flight, 3 speeds, 4 policies. Results shown for avg. speed = 0.5 m/s



SSD evaluation

Himax dataset



Setup: Tested on Himax dataset

Tested: 3 CNN depth multipliers (1x, 0.75x, 0.5x).

SSD throughput/accuracy tradeoffs:

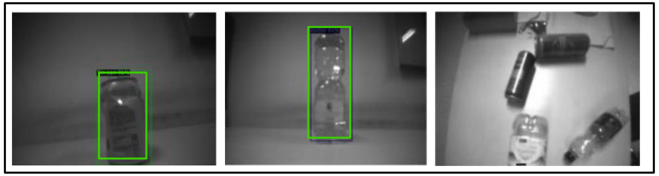
SSD	Size [MB]	MAC	mAP	Throughput [FPS]
1x	4.7	534M		
0.75x	2.7	358M		
0.5x	1.2	193M		

SSD evaluation

Setup: Tested on Himax dataset

Tested: 3 CNN depth multipliers (1x, 0.75x, 0.5x).

Himax dataset



SSD throughput/accuracy tradeoffs:

SSD	Size [MB]	MAC	mAP	Throughput [FPS]
1x	4.7	534M	50%	1.6
0.75x	2.7	358M	48%	2.3
0.5x	1.2	193M	32%	4.3

→ **most accurate** & slowest

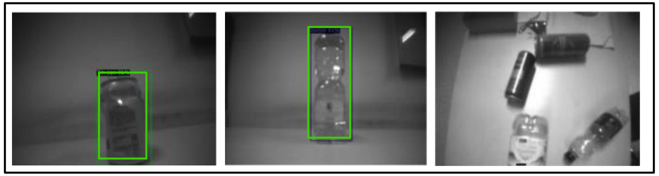
→ least accurate & **fastest**

SSD evaluation

Setup: Tested on Himax dataset

Tested: 3 CNN depth multipliers (1x, 0.75x, 0.5x).

Himax dataset



SSD throughput/accuracy tradeoffs:

SSD	Size [MB]	MAC	mAP	Throughput [FPS]
1x	4.7	534M	**50%**	**1.6**
0.75x	2.7	358M	**48%**	**2.3**
0.5x	1.2	193M	**32%**	**4.3**

→ **most accurate** & slowest

→ least accurate & **fastest**

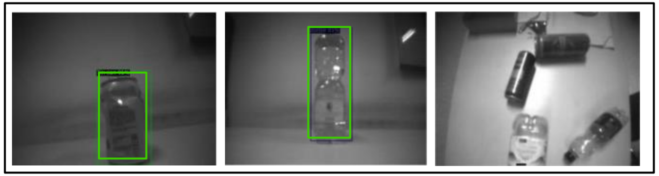
SSD with best mAP:
1x and 0.75x

SSD evaluation

Setup: Tested on Himax dataset

Tested: 3 CNN depth multipliers (1x, 0.75x, 0.5x).

Himax dataset



SSD throughput/accuracy tradeoffs:

SSD	Size [MB]	MAC	mAP	Throughput [FPS]
1x	4.7	534M	50%	1.6
0.75x	2.7	358M	48%	2.3
0.5x	1.2	193M	32%	4.3

→ **most accurate** & slowest

→ least accurate & **fastest**

SSD with best mAP:
1x and 0.75x

Power consumption

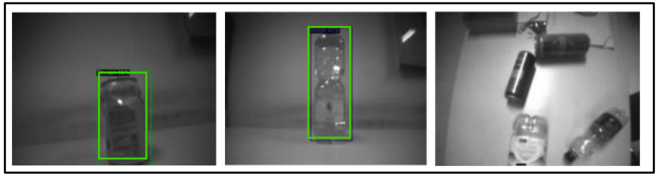
SSD	Power
1x	134 mW
0.75x	143 mW

SSD evaluation

Setup: Tested on Himax dataset

Tested: 3 CNN depth multipliers (1x, 0.75x, 0.5x).

Himax dataset



SSD throughput/accuracy tradeoffs:

SSD	Size [MB]	MAC	mAP	Throughput [FPS]
1x	4.7	534M	50%	1.6
0.75x	2.7	358M	48%	2.3
0.5x	1.2	193M	32%	4.3

→ **most accurate** & slowest

→ least accurate & **fastest**

SSD with best mAP:
1x and 0.75x

Power consumption

SSD	Power
1x	134 mW
0.75x	143 mW

Drone's power breakdown

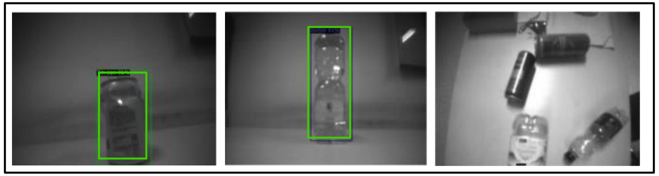
	Motors	CF elect.	AI-deck	Ranger deck	Total
Power [W]	7.32	0.277	0.134	0.286	8.02
Percentage	91.3%	3. 5%	1.7%	3.6%	100%

SSD evaluation

Setup: Tested on Himax dataset

Tested: 3 CNN depth multipliers (1x, 0.75x, 0.5x).

Himax dataset



SSD throughput/accuracy tradeoffs:

SSD	Size [MB]	MAC	mAP	Throughput [FPS]
1x	4.7	534M	50%	1.6
0.75x	2.7	358M	48%	2.3
0.5x	1.2	193M	32%	4.3

→ **most accurate** & slowest

→ least accurate & **fastest**

SSD with best mAP:
1x and 0.75x

Power consumption

SSD	Power
1x	134 mW
0.75x	143 mW

Drone's power breakdown

	Motors	CF elect.	**AI-deck**	Ranger deck	Total
Power [W]	7.32	0.277	**0.134**	0.286	8.02
Percentage	91.3%	3. 5%	**1.7%**	3.6%	100%

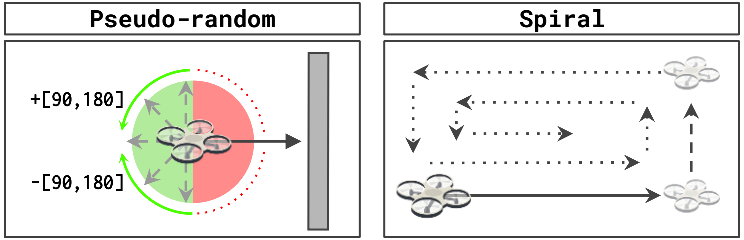
In-field evaluation: setup

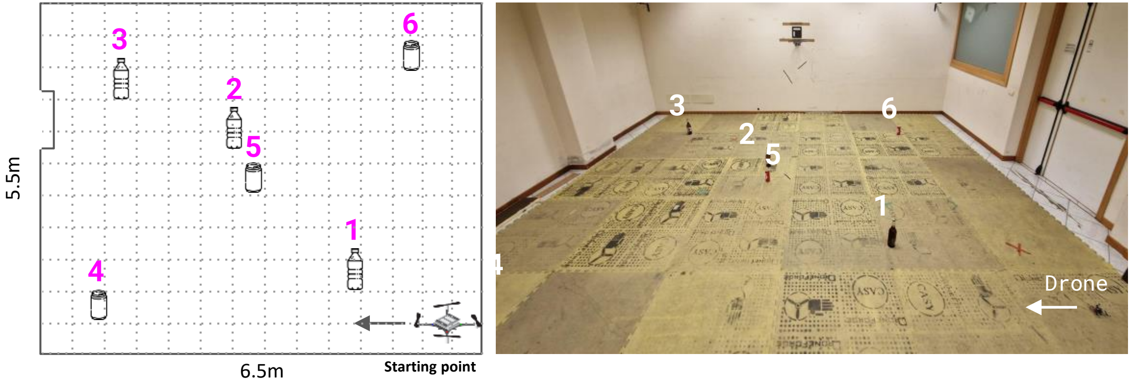
Experiment configurations:

- 2 exploration policies
- SSD models: 1x, 0.75x
- 3 mean velocities: [0.1, 0.5, 1.0] m/s
- 5 runs for each configuration

5 minutes flight

6 objects (bottles, tin cans)





Closed-loop system evaluation

Goal: best detection rate.

Comparing CNN vs. exploration policies vs. flight speed

	Detection rate		Coverage area	
Avg speed	**Pseudo-random**	**Spiral**	**Pseudo-random**	**Spiral**
0.5 m/s				
1 m/s				

Closed-loop system evaluation

Goal: best detection rate.

Comparing CNN vs. exploration policies vs. flight speed

	Detection rate		Coverage area	
Avg speed	**Pseudo-random**	**Spiral**	**Pseudo-random**	**Spiral**
0.5 m/s				
1 m/s				

Best configuration:

- CNN: SSD 1x ---> *see paper!*

Closed-loop system evaluation

Goal: best detection rate.

Comparing CNN vs. exploration policies vs. flight speed

	Detection rate		Coverage area	
Avg speed	**Pseudo-random**	**Spiral**	Pseudo-random	Spiral
0.5 m/s	**90%**	**73%**	74%	82%
1 m/s	83%	70%	80%	83%

Best configuration:

- CNN: SSD 1x ---> *see paper!*
- Exploration:

Closed-loop system evaluation

Goal: best detection rate.

Comparing CNN vs. exploration policies vs. flight speed

	Detection rate		Coverage area	
Avg speed	**Pseudo-random**	**Spiral**	**Pseudo-random**	**Spiral**
0.5 m/s	**90%**	**73%**	74%	82%
1 m/s	83%	70%	80%	83%

Best configuration:

- CNN: SSD 1x ---► *see paper!*
- Exploration: Pseudo-random

Closed-loop system evaluation

Goal: best detection rate.

Comparing CNN vs. exploration policies vs. flight speed

	Detection rate		Coverage area	
Avg speed	**Pseudo-random**	**Spiral**	**Pseudo-random**	**Spiral**
0.5 m/s	**90%**	**73%**	74%	82%
1 m/s	83%	70%	80%	83%

Best configuration:

- CNN: SSD 1x ---► *see paper!*
- Exploration: Pseudo-random
- Speed:

Closed-loop system evaluation

Goal: best detection rate.

Comparing CNN vs. exploration policies vs. flight speed

	Detection rate		Coverage area	
Avg speed	**Pseudo-random**	Spiral	Pseudo-random	Spiral
0.5 m/s	**90%**	73%	74%	82%
1 m/s	83%	70%	80%	83%

Best configuration:

- CNN: SSD 1x ---► *see paper!*
- Exploration: Pseudo-random
- Speed: 0.5 m/s

Closed-loop system evaluation

Goal: best detection rate.

Comparing CNN vs. exploration policies vs. flight speed

	Detection rate		Coverage area	
Avg speed	**Pseudo-random**	Spiral	Pseudo-random	Spiral
0.5 m/s	**90%**	73%	74%	82%
1 m/s	83%	70%	80%	83%

Best configuration:

- CNN: SSD 1x ---▸ *see paper!*
- Exploration: Pseudo-random
- Speed: 0.5 m/s

Best detection rate: **90%** (avg)

Closed-loop system evaluation

Goal: best detection rate.

Comparing CNN vs. exploration policies vs. flight speed

	Detection rate		Coverage area	
Avg speed	**Pseudo-random**	Spiral	**Pseudo-random**	Spiral
0.5 m/s	**90%**	73%	74%	82%
1 m/s	83%	70%	**80%**	83%

Flight speed

Detection rate
vs.
coverage area

Closed-loop system evaluation

Goal: best detection rate.

Comparing CNN vs. exploration policies vs. flight speed

	Detection rate		Coverage area	
Avg speed	Pseudo-random	Spiral	Pseudo-random	Spiral
0.5 m/s	**90%**	73%	74%	82%
1 m/s	83%	70%	**80%**	83%

Flight speed

The best detection rate

≠

highest coverage area!

Closed-loop system evaluation

Goal: best detection rate.

Comparing CNN vs. exploration policies vs. flight speed

	Detection rate		Coverage area	
Avg speed	**Pseudo-random**	Spiral	**Pseudo-random**	Spiral
0.5 m/s	**90%**	73%	74%	82%
1 m/s	83%	70%	**80%**	83%

Flight speed

The best detection rate

≠

highest coverage area!

Take away message

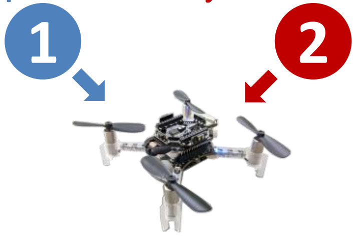
Higher flight speed → improves the coverage area.

But challenges the object detector's capability due to its limited throughput (1.6 fps).

Conclusion

Enabled multi-tasking perception on an autonomous nano drone

Tasks	
Exploration	Object detection



Best configuration	
CNN	SSD 1x
Policy	pseudo-random
Speed	0.5 m/s

Best detection rate: **90%**

This project was partially funded by the Autonomous Robotics Research Center of the Technology Innovation Institute

TII Technology Innovation Institute

Thank you for your attention !

Lorenzo Lamberti

Ph.D. student at University of Bologna, Italy
lorenzo.lamberti@unibo.it

ALMA MATER STUDIORUM UNIVERSITÀ DI BOLOGNA

KU LEUVEN

USI/SUPSI IDSIA Istituto Dalle Molle di studi sull'intelligenza artificiale

ETH zürich