DLR

esa

AGENIUM SPACE

living planet symposium | BONN 23–27 May 2022

TAKING THE PULSE OF OUR PLANET FROM SPACE

EUMETSAT ECMWF EUSPA

Distillation experiments for on board Deep Neural Networks

F. de Vieilleville [1], A. Lagrange [1], M. Verm [1], N. Lemoine [1], N. Dublé [1], R. Ruiloba Quecedo [1], R. Camarero [2], B. Le Saux [3]
[1] AGENIUM SPACE, [2] ESTEC, [3] ESRIN-PHILAB

B9.06 AI@edge and Emerging Computing Paradigms for the Future of Earth Observation
24/05/2022

Motivation

EVOLUTION OF DL

DL done on the ground

DL done on at the edge (satellite)

Pros:

- Verifiable public/private data reference databases available
- Lots of computational resources at hand (GPU, TPU, …)

Cons:

- Selective data availability, high cost
- Produced as huge models, requiring a lot computational power/memory

CHALLENGES

- Limited Computation power
- Limited energy
- Limited memory capacity
- Small models, running at low power within a limited memory consumption footprint

Pros:

- Enables autonomy
- Increases ROI of satellite
- Creates value-added insights
- Increases responsiveness
- Reduces the cost of downloads

Cons:

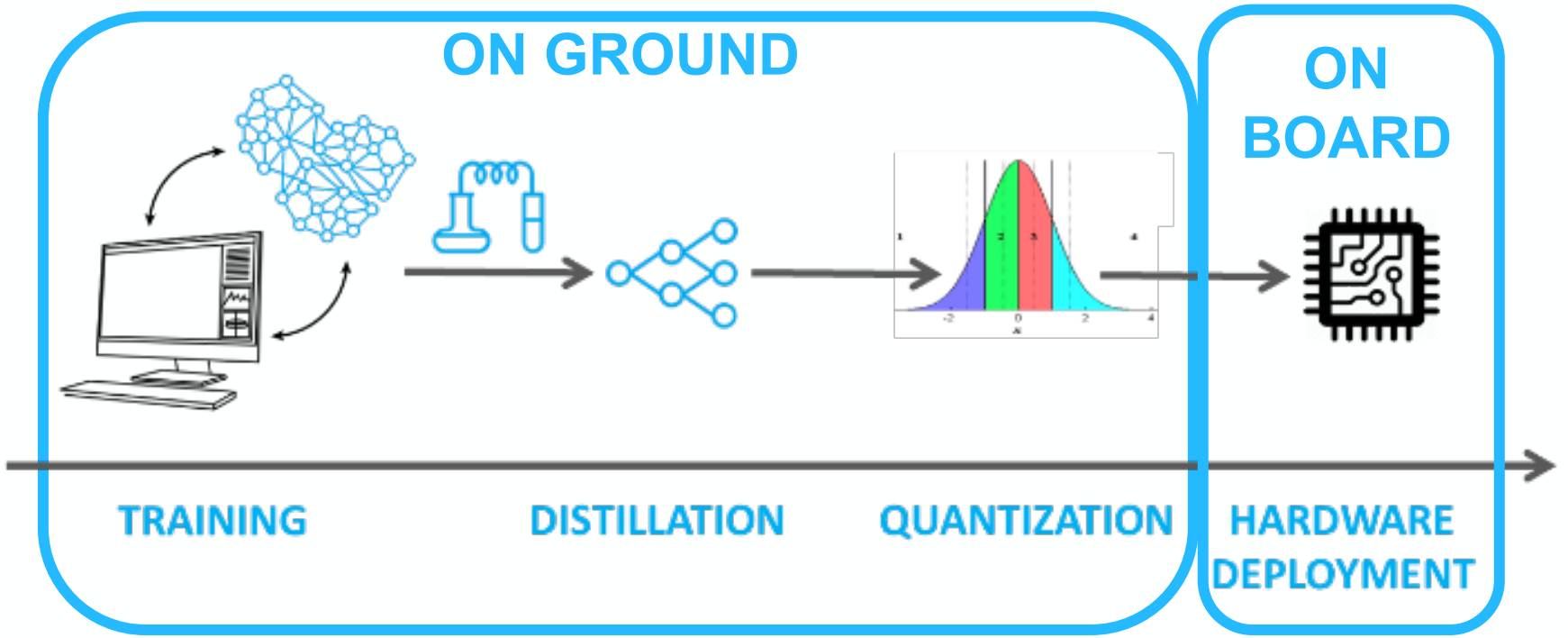
- It is constrained as hell – requires small models, running at low power within a limited memory consumption footprint

SW tools for Deep Learning HW

On board AI capabilities depend on both HW and SW tooling:

Soc FPGA/FPGA:	CONS	PROS
VHDL/Verilog/HLS	Long to develop	Ad Hoc
VITIS AI (Xilinx HW)	Black Box	Short to develop
CPU (arm)	**CONS**	**PROS**
TensorFlow	Only FP32	Fast to develop, OSS
TensorFlowLite	INT8 backend is super slow	FP16, Fast to develop, OSS
Pytorch	INT8 backend is super slow	Fast to develop, OSS
GPU (AMD G/R series)	**CONS**	**PROS**
TensorFlow	Does not work	OSS
TensorFlowLite	C++ only backend	Short to develop, OSS
VPU (Myriad)	**CONS**	**PROS**
OpenVINO		Short to develop, OSS

Proposed Generic Pipeline



Use cases

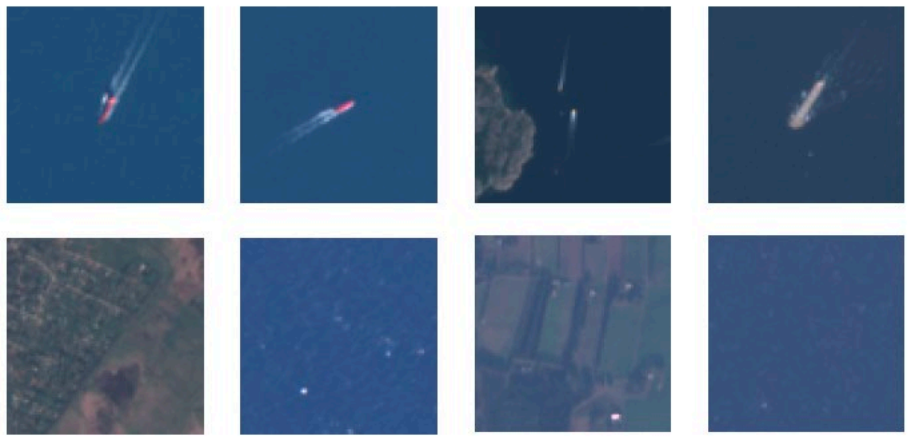
Conversion of a ready ground-DL into onboard-DL was demonstrated through the following ESA projects:

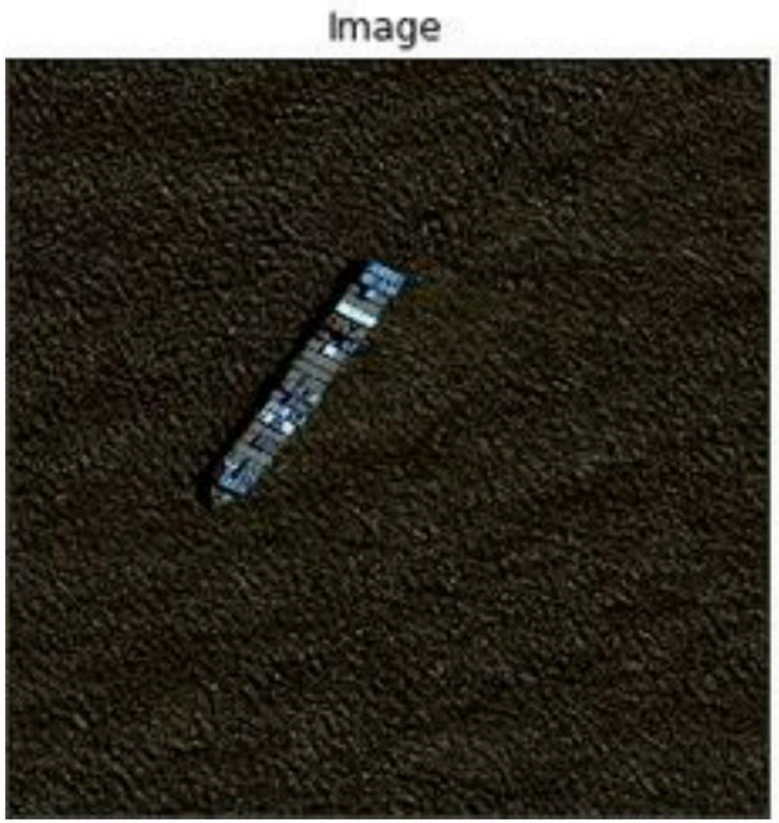
- Cortex
- Deep Cube

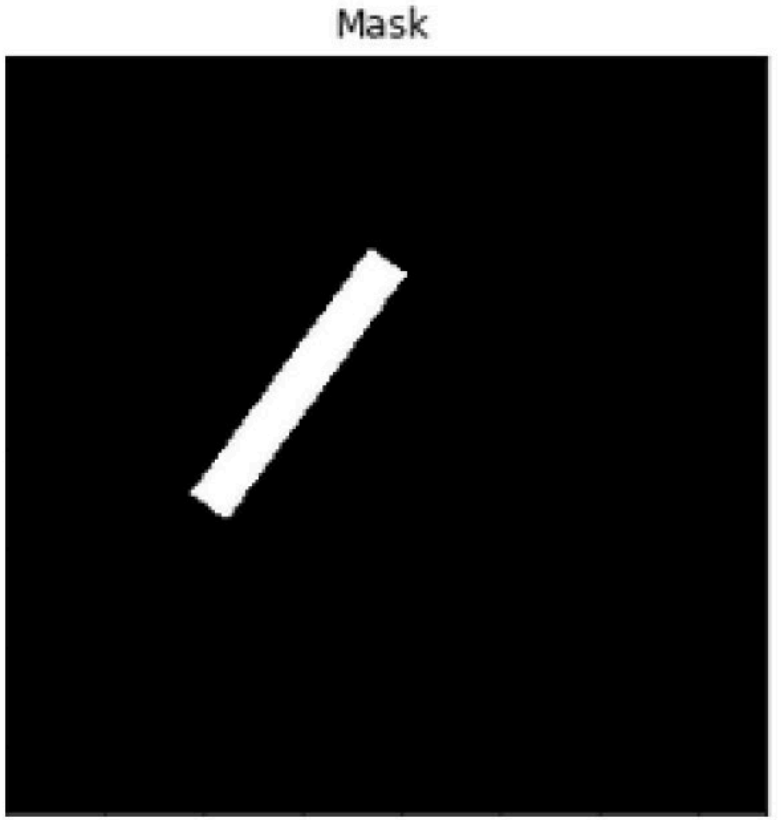
	CORTEX			DEEP CUBE			
	Boat (S2) RGB+NIR	Oil spill (S1)	Ocean Features (S1)	Boat (VHR) RGB	Clouds (S2) RGB+NIR	Clouds/snow (S2) RBG+NIR	Forest (S2) RGB+NIR
Classification	✓	✓	✓				
Segmentation					✓	✓	✓
Detection				✓			

Example of use case - boats

- Select & downlink relevant image patches with boat types

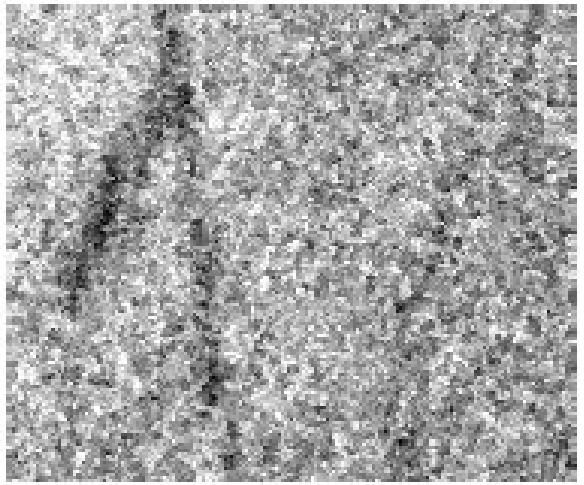
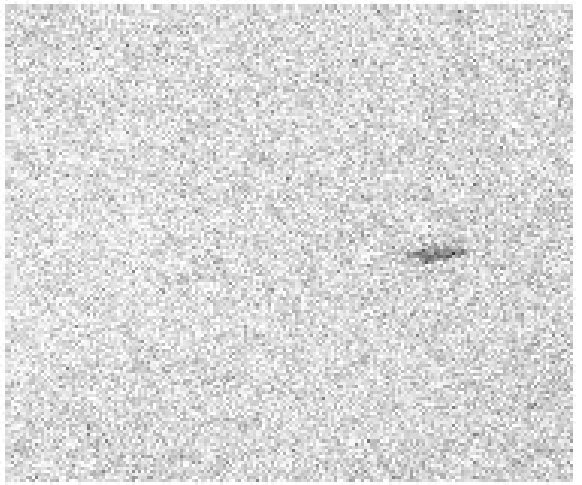
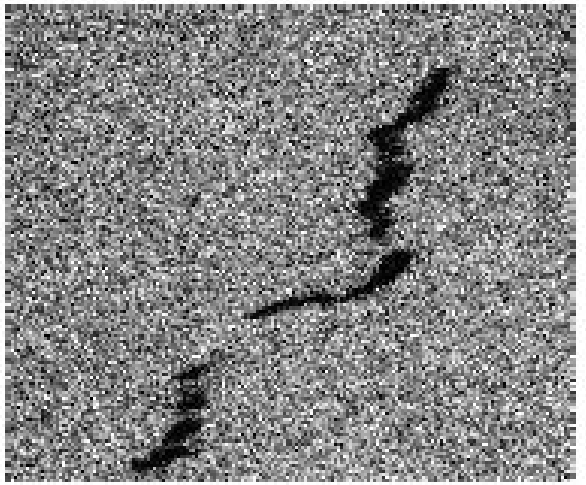
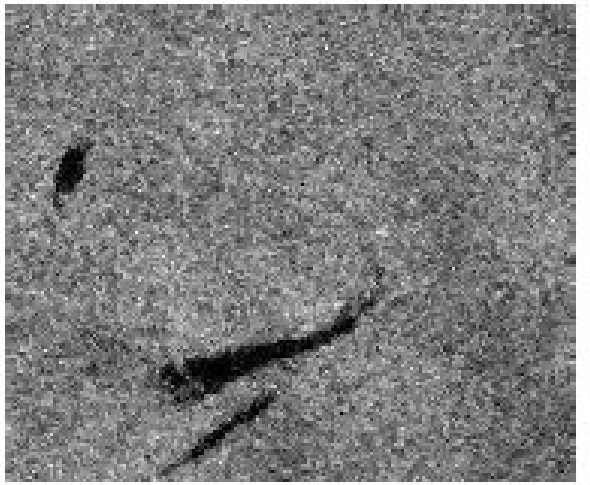






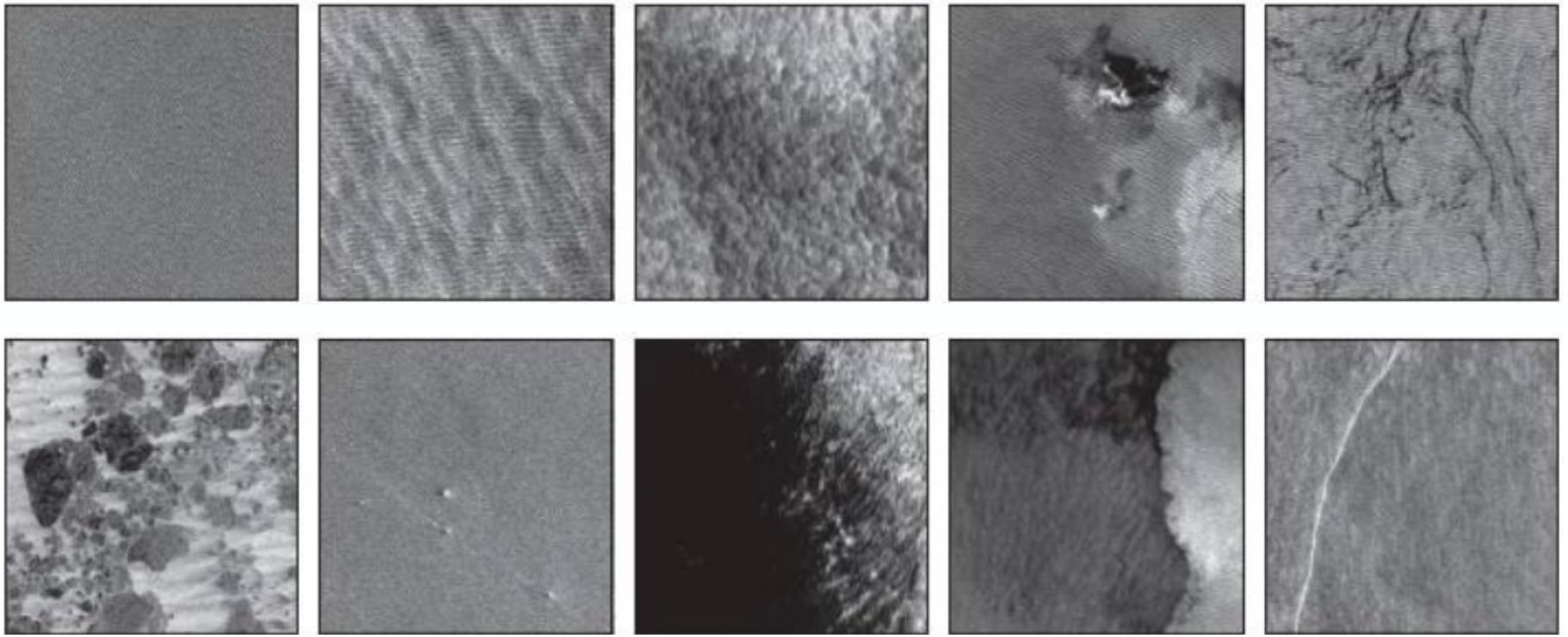
Example of use case - oil spills

- Select & downlink relevant image patches with boat types / oil spills

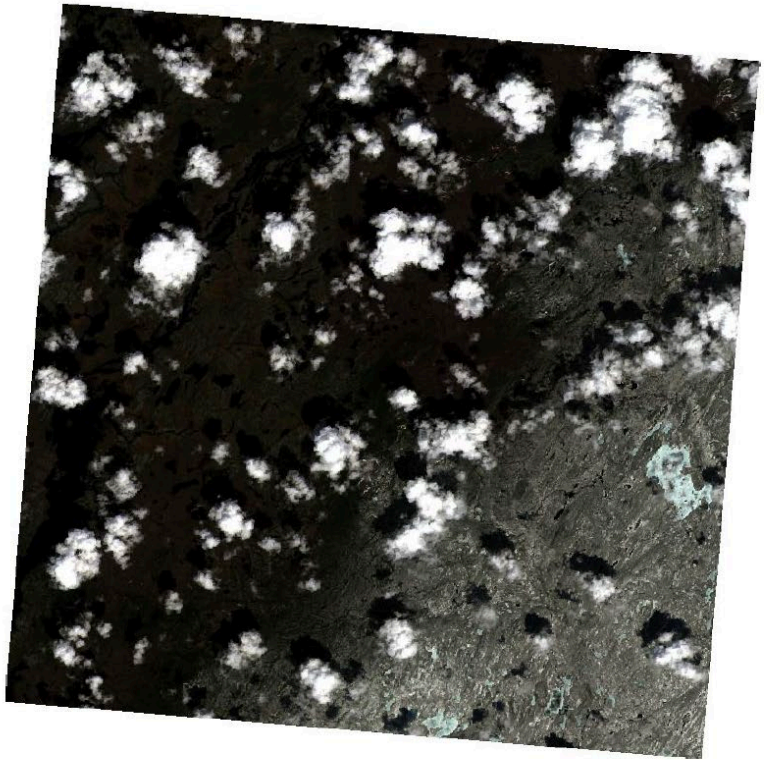
Example of use case – ocean features

- Select & downlink relevant image patches with boat types / oil spills / specific sea states

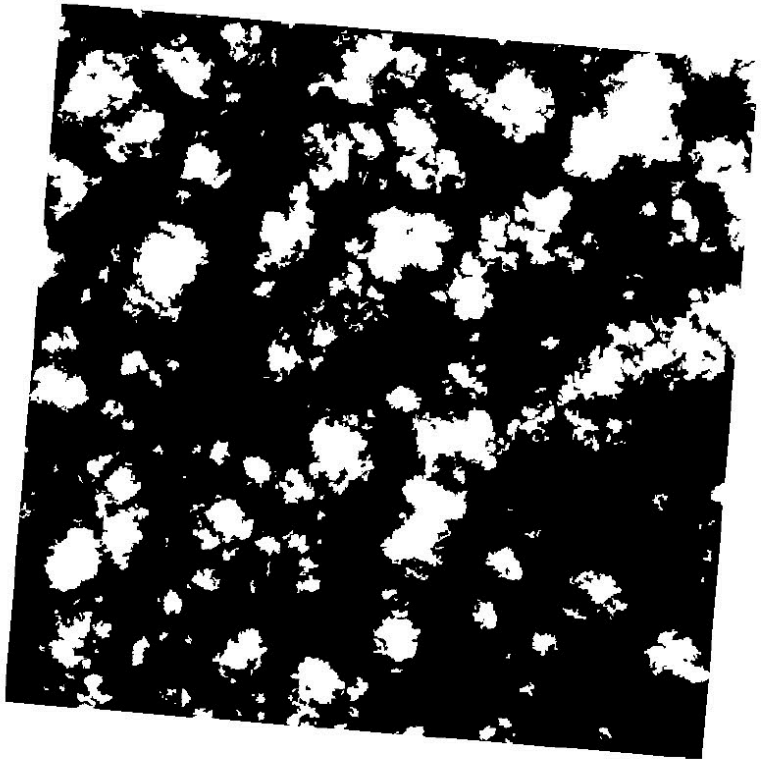


Example of use case - clouds

- Select & downlink relevant image patches with boat types
- Reduce downlink/ discards useless images/ increase sat. efficiency



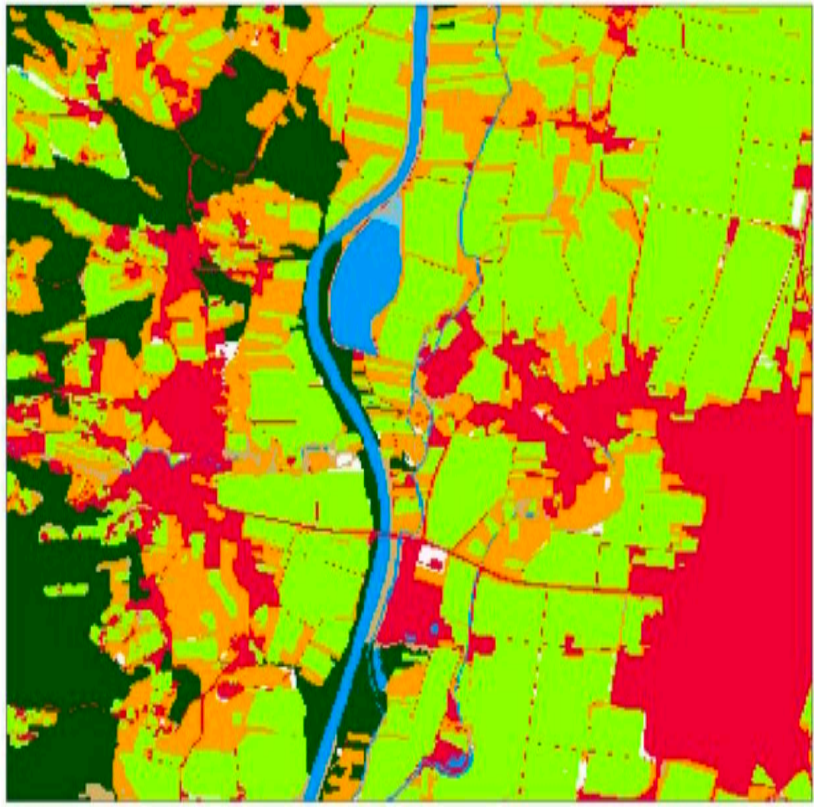
Image



Cloud ground truth

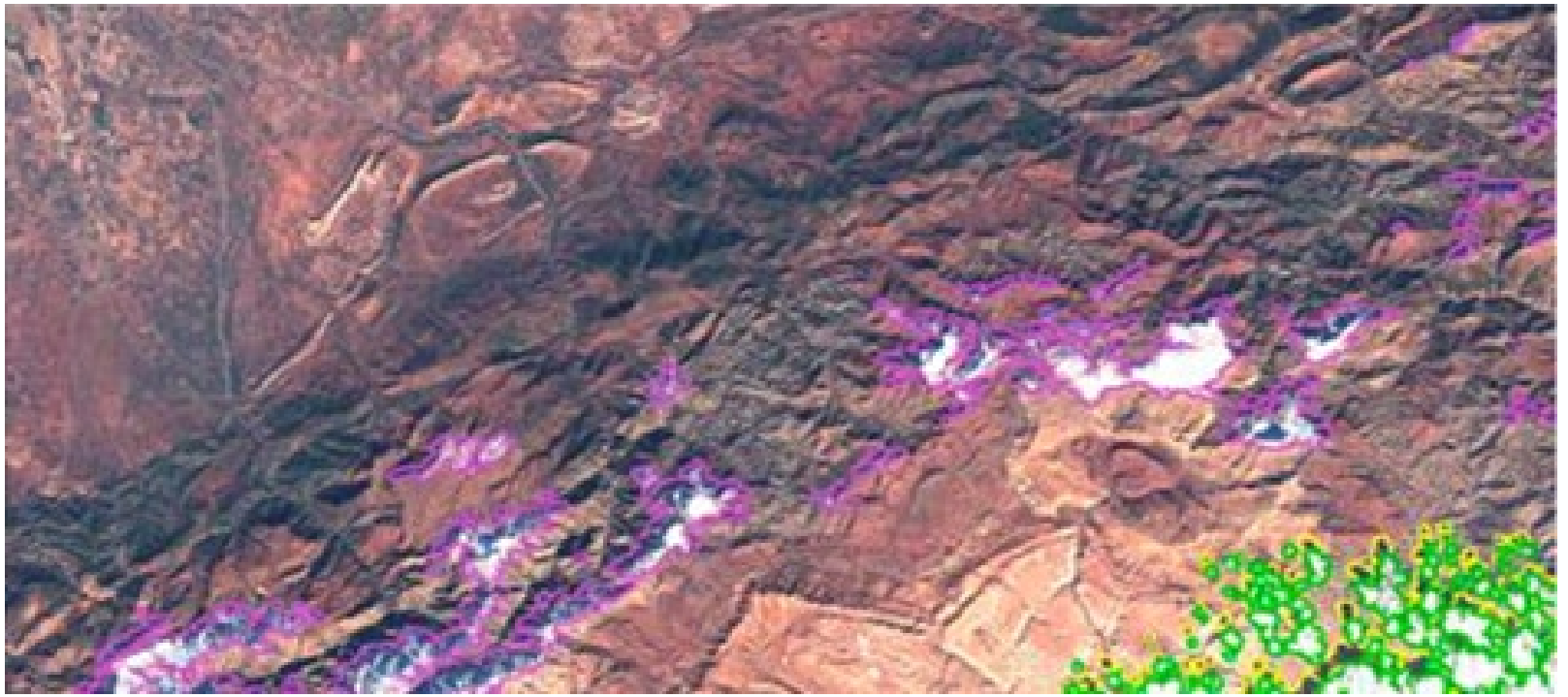
Example of use case - forest

- Select & downlink relevant image patches with boat types
- Reduce downlink/ discards useless images/ increase sat. efficiency
- Reduce downlink/ downlink only images with changes



Example of use case - clouds VS snow

- Select & downlink relevant image patches with boat types
- Reduce downlink/ discards useless images/ increase sat. efficiency
- Reduce downlink/ downlink only images with changes
- Reduce downlink/ select only images with snow



Distillation performances

All master models are around 90% F1 score.

Distillation is measured as model's parameters reduction and attrition.

	CORTEX (Distillation SW v1)			DEEP CUBE (Distillation SW v2)			
	Boat (S2)	Oil spill (S1)	Ocean Features (S1)	Boat (VHR)	Clouds (S2)	Clouds/snow (S2)	Forest (S2)
Reduction Factor	x52 135M → 2.6M	x52 135M → 2.6M	x52 135M → 2.6M	x80 24M → 300k	x60 59M → 1M	x120 59M → 0.5M	x160 16M → 100k
Attrition F1 score (Distillation V1)	<2%	<8%	<5%	NA	~13%	NA	<3%
Attrition F1 score (Distillation V2)	NA	NA	NA	~5%	~6%	<5%	NA

Tested HW

HW tested for compatibility with our pipeline from Ground to Board through 2 ESA funded projects:

- Cortex : pipeline proposition for classification on Xilinx ZU+ Soc FPGA
- Deep Cube : pipeline consolidation for other classic DL tasks on more space HW

HW performance measurements : **Throughput & Consumption**

	Xilinx Zynq SocFPGA → Z7045	Xilinx ZynqUltraScale+ SocFPGA → SDR Xiphos Q8, Leopard	Xilinx Kintex FPGA → KU060	Intel Movidius 2 VPU → Ubotica's CogniSat	AMD G-Series CPU/GPU → Unibap's DD-iX5	Intel I7-9700K → reference
Classification		✓				
Segmentation	✓	✓	✓	✓	✓	✓
Detection	✓	✓	✓	✓	✓	✓
Quantization	INT8	INT8	INT8	FP16	INT8/FP16/ FP32	FP32/FP16

HW Throughput Performances

No attrition between distilled models and quantized models (INT8 & FP16)

Cortex : VITIS AI <= 1.2 / DNNDK

Deep Cube :VITIS AI <= 1.4 / VART | TFLite/Pytorch | OpenVino

	Xilinx Zynq SocFPGA-7020 Pixels/W/s	Xilinx ZynqUltraScale+ SocFPGA - ZU9G Pixels/W/s	Xilinx Kintex FPGA-KU040 Pixels/W/s	Intel Movidius 2 VPU Pixels/W/s	AMD G-Series CPU/GPU Pixels/W/s
Classification (2.6M)	NA	570k	NA	NA	NA
Segmentation (0.1M ~ 1M)	70k to 120k	170k to 215k	430 to 640	170k to 350k	8K to 19K
Detection (0.3M ~ 0.4M)	115k to 190k	350k to 600k	3k	320k to 640k	7K to 17k

HW Consumption

Cortex : VITIS AI <= 1.2 / DNNDK

Deep Cube :VITIS AI <= 1.4 / VART | TFLite/Pytorch | OpenVino

	Xilinx Zynq SocFPGA AveW/MaxW	Xilinx ZynqUltraScale+ SocFPGA AveW/MaxW	Xilinx Kintex FPGA AveW/MaxW	Intel Myriad 2 VPU AveW/MaxW	AMD G-Series CPU/GPU AveW/MaxW	Intel i7-9700K CPU AveW/MaxW
Classification (2.6M)	NA	10.5/13.7 (3xB4096)	NA	NA	NA	NA
Segmentation (0.1M ~ 1M)	3.25/5 (1xB1152)	10/15 (3xB4096) 9/13 (2xB4096)	3.324/3.324 (1xB4096)	1/1	~10/~10	95/95
Detection (0.3M ~ 0.4M)	3.25/5 (1xB1152)	10/15 (3xB4096) 9/13 (2xB4096)	3.324/3.324 (1xB4096)	1/1	~10/~10	95/95

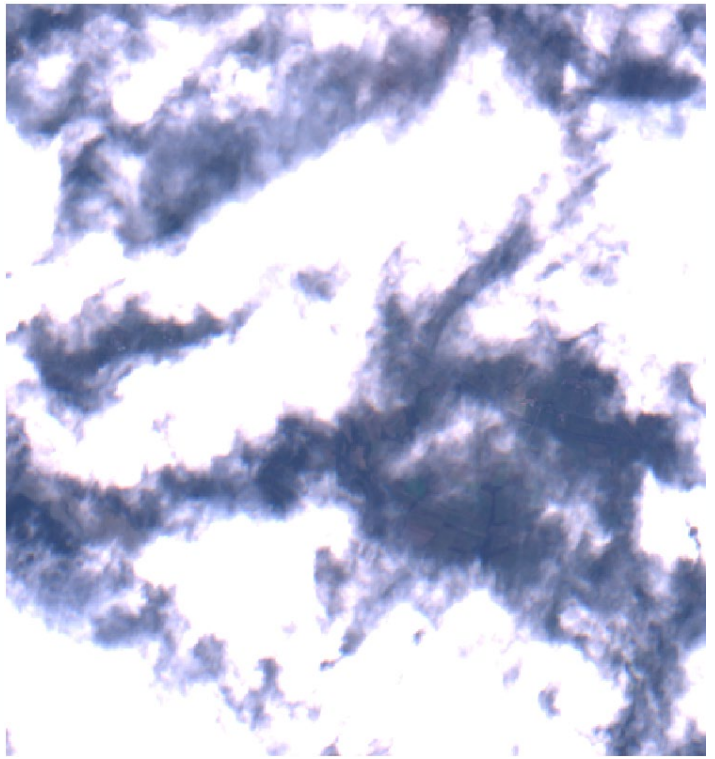
Efficiency

From our experiments, what are the best devices for DL in space ?

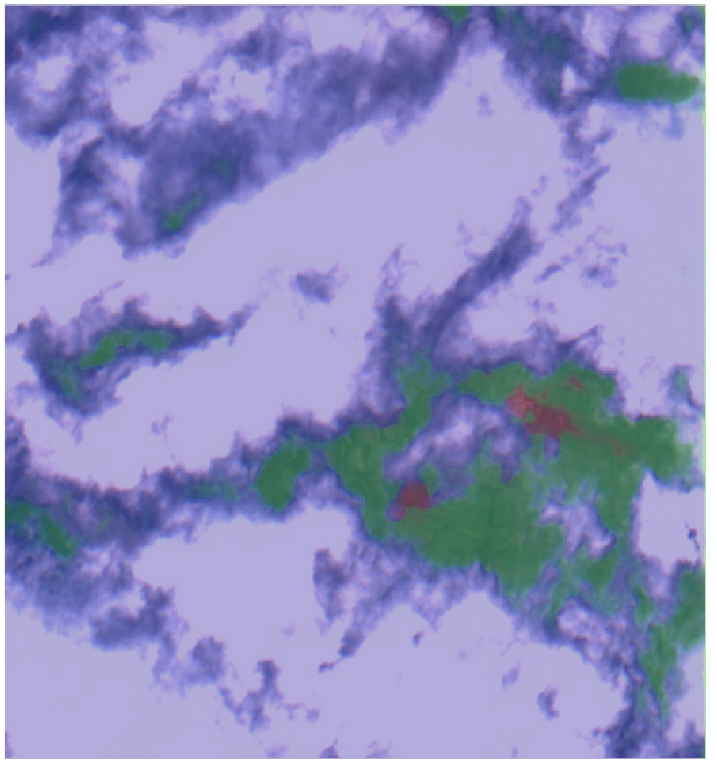
	Model parameters (M)	Intel CPU Core i7-9700K (95W)	AMD G-Series (iX5) (10W)	Xilinx HW - FPGA & SoCs (~10W)	Intel Myriad VPU 2 (1W)	Xilinx HW - FPGA (<3.5W)
Cloud Segmentation	1	1.0	0.4	**9.2**	**9.0**	0.03
Forest Segmentation	0.1	1.0	0.6	**6.5**	**9.8**	0.033
Forest & Cloud Segmentation	0.1	1.0	0.5	**5.2**	**9.6**	0.025
Snow VS Cloud Segmentation	0.5	1.0	0.4	**8.8**	**9.9**	0.023
Boat Detection	0.3	1.0	0.2	**5.6**	**6.0**	0.16

VPU and FPGA & Soc are good choices !
Still room for improvement on FPGA & Soc SW !

And now in space !



input



output

clouds

background

forest

Flight demonstration of a simplified forest segmentation algorithm (92% F1 score)
CPU on iX5 : AMD G-series @10W

And now in space !

Thanks to the WHOLE TEAM !!!

Adrien	Mathieu	Nicolas
Nimesh	Nicolas	Faisal
Théo	Andis	François

https://agenium.com/
contact.space@agenium.com
1, avenue de l'Europe
31400 TOULOUSE
FRANCE
t : +33 (0)5 61 41 03 98
m :+33 (0)6 46 78 63 34