EDHPC 2025 – October 13th 2025

esa

Explaining raw data complexity to improve satellite onboard processing

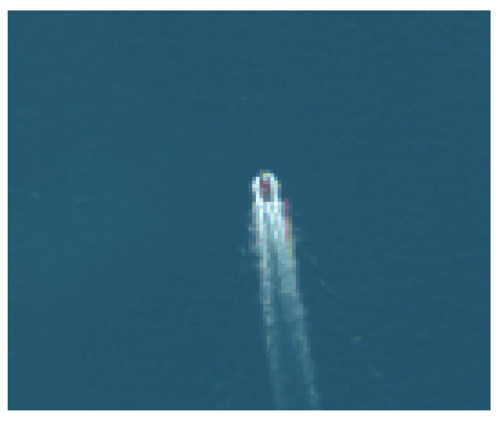
Adrien Dorise[1,2], Marjorie Bellizzi[1], Adrien Girard[1], Benjamin Francesconi[1], Stéphane May[2]

[1]IRT Saint-Exupéry – [2]CNES

CONTEXT

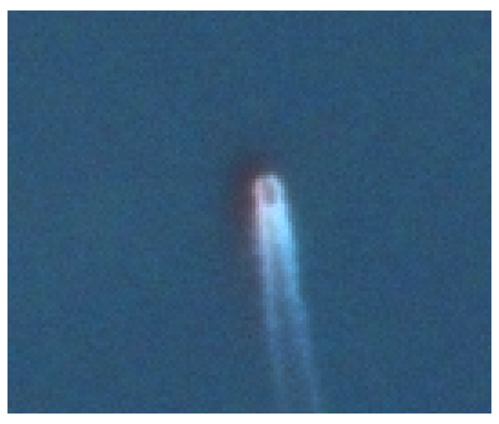
- Onboard AI for Earth observation is now feasible, but sensor outputs **are noisy, deregistered and prone to artefacts**.
- Current AI models rely on **corrected Level-1 (L1) imagery**. This gap limits reliable, real-time detection, as expensive on-board preprocessing is required.
- Training models on raw data is complicated due to the scarcity of available datasets. The ESA Φ-lab has made some progress by proposing Sentinel-2 and Venµs datasets, along with a coarse-coregistration technique **[1]**.
- Our work focuses on understanding **how raw data impacts deep learning models** by creating a controlled setting that generates raw-like and reconstructed L1 imagery from **high-resolution products**.

IMAGES SIMULATION

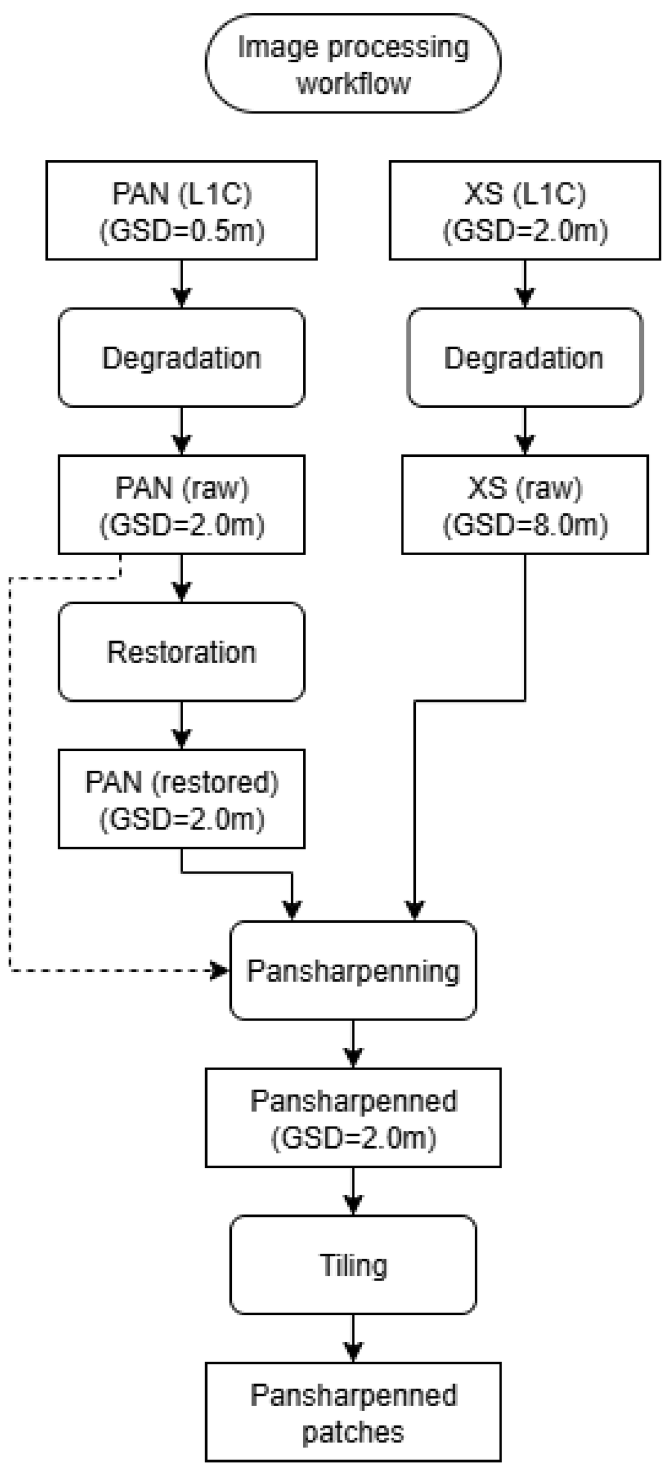


Reference image

L1 simulated



Raw simulated



To simulate a high-resolution product from panchromatic (PAN) and multispectral (XS) products, we recreated the classic processing chain.

1. The PAN and XS images are degraded to simulate the sensor characteristics in terms of SNR, MTF. Deregistration is added to the XS product.
2. The raw PAN is restored using an embedded neural network **[2]**. **This step is skipped when simulating a raw image**.
3. The Brovey method is used to create a pansharpened product.
4. The pansharpened product is divided into patches offsetly centred on the object to detect. **This serves as input for the detection model**.

Experiments are performed to validate the simulation workflow.

EXPERIMENTS

- Three datasets are created in this work:
  - **Reference RGB** images that derive from the original XS products.
  - **Simulated raw** images that represents sensor outputs.
  - **Simulated L1** images that representes raw images after restoration.

Set	GSD src / target	MTF @Nyquist	(Lum., SNR) ref. ($W/m^2/sr/\mu m$, dB)	Band offsets
L1	0.5 m / 2.0 m	0.25	(25, 80); (100, 170)	None
Raw	0.5 m / 2.0 m	0.05	(25, 50); (100, 110)	[1,4] px

Simulated dataset parameters

- Single-stage detectors **YOLOv11n** and **YOLOX-S** are trained with identical settings on each dataset.
- Evaluation relies on mAP@50/90, F1 vs. confidence, IoU for localisation, and Explainability AI (XAI) saliency maps **[3]**.

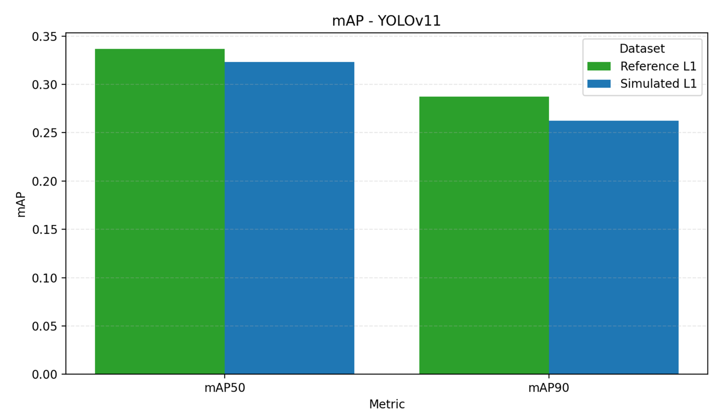
Patch size	Resolution	# train	# val	# classes
256 x 256 px	2.0 m	2716	306	6

Datasets characteristics

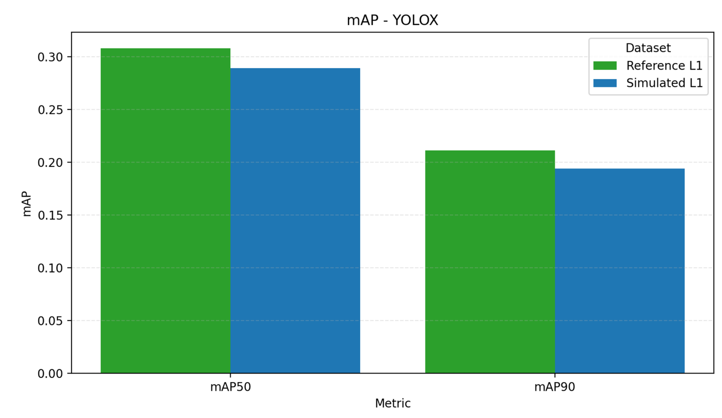
RESULTS

Set	Model	Conf. levels	mAP50	mAP90	F1	IoU on TP
ref. L1	YOLOv11n	0.001	0.3366	0.2873	0.1019	0.8707
		0.1			0.3323	0.8916
		0.5			0.3559	0.9239
	YOLOX-S	0.001	0.3081	0.2115	0.3229	0.7882
		0.1			0.3996	0.8246
		0.5			0.3805	0.8279
$\tilde{L1}$	YOLOv11n	0.001	0.3230	0.2624	0.1667	0.8324
		0.1			0.2727	0.8566
		0.5			0.2359	0.8862
	YOLOX-S	0.001	0.2895	0.1942	0.3463	0.7876
		0.1			0.3590	0.8096
		0.5			0.3754	0.8182
Raw	YOLOv11n	0.001	0.3182	0.2584	0.1495	0.8352
		0.1			0.2879	0.8675
		0.5			0.2059	0.9128
	YOLOX-S	0.001	0.2667	0.1876	0.3387	0.7898
		0.1			0.3671	0.8101
		0.5			0.2418	0.8860

Simulation workflow validation: The simulation workflow is evaluated by comparing the mAP scores of the models trained on the reference dataset vs. the simulated L1 dataset. Although a slight performance drop is observed, particularly with YOLOv11, the dataset preserves most of the relevant information and can be used for reliable experiments. Qualitative examples on simulated L1 datasets confirm that the models can detect vessels across these images.

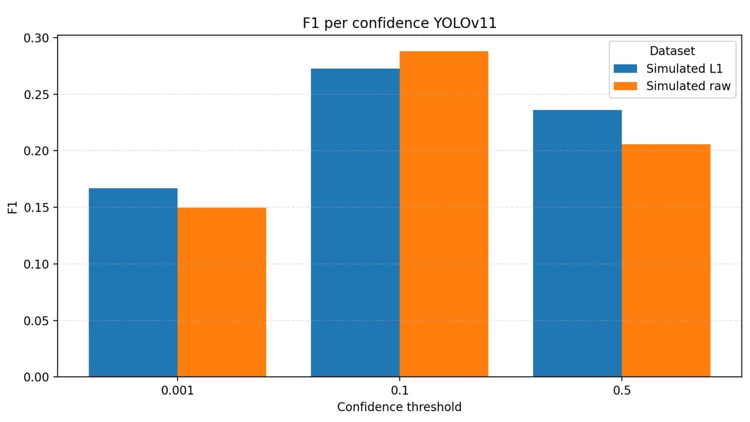


YOLOv11 mAP per dataset

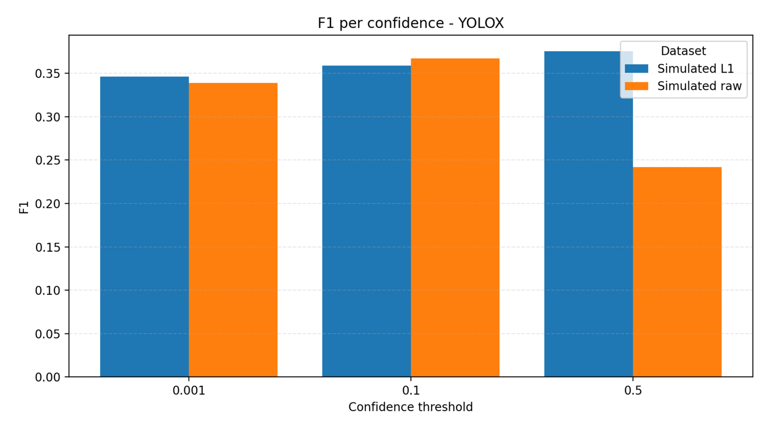


YOLOX mAP per dataset

Raw data effects: F1 curves for simulated L1 vs. simulated raw are similar at low to medium confidence, indicating comparable detectability. At higher thresholds, raw-trained models decline faster. This suggests that processing raw data is especially challenging when handling edge cases and identifying complex objects.



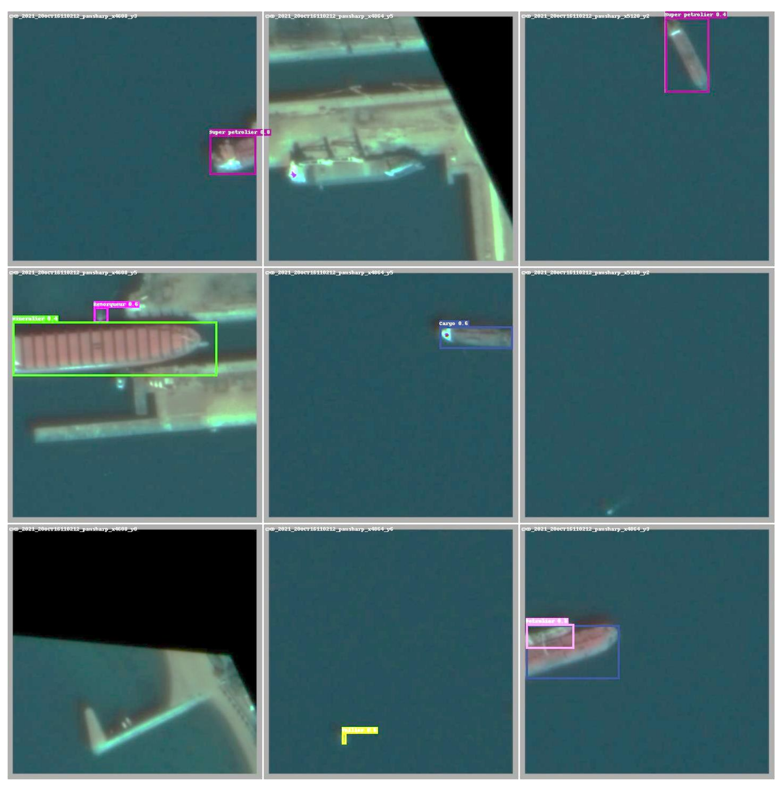
YOLOv11 F1 scores



YOLOX F1 scores

Raw data insights: The explainability analysis suggests models trained on raw data struggle to clearly identify the boundaries of the object, resulting in less reliable predictions. Since accurate boundary localisation is essential for robust predictions, this explains part of the drop in robustness.

Saliency maps for reference (left) and raw patches (right)



YOLOv11 predictions on simulated L1 patches

CONCLUSION

- Our simulation workflow effectively **produces raw and L1** images that can be used in assessing AI detection models.
- The raw images **impact the robustness of the predictions.**
- Feature attribution maps indicate that models trained on raw data **struggle to define object boundaries**. XAI can effectively assess these challenges.

REFERENCES

[1] R. Del Prete, et al. "Maritime monitoring via onboard processing of raw multispectral imagery by deep learning" IGARSS 2024

[2] B. Lim et al. "Enhanced deep residual networks for single image super-resolution", CoRR, 2017

[3] T. Fel et al. "Xplique: A deep learning explainability toolbox", 2022

Contact
Adrien DORISE
adrien.dorise@cnes.fr