Source detection and characterization using YOLO-CIANNA in simulated radio data

David Cornu

LUX, *Observatoire de Paris, PSL*

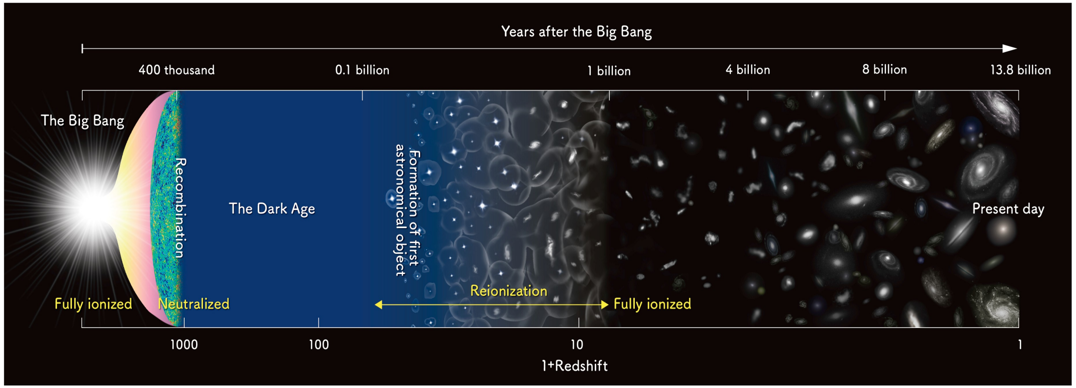
CIANNA workshop 2025, Paris

LUX Observatoire de Paris | PSL PR[AI]RIE Paris School CIANNA Workshop 2025

The Square Kilometer Array

Future largest radio-telescope

Evolution of the universe and astrophysical objects



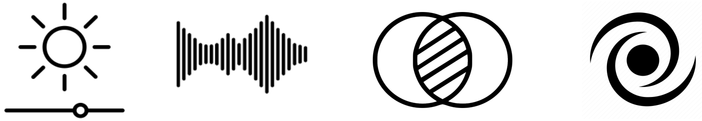
Construction phase started ! ETA : ~2028

Big Data

~700 PB / year (stored)

~1.5 millions 500 GB HDD / year

Complex data



Contrast ***Noise*** ***Confusion*** ***Morphology***

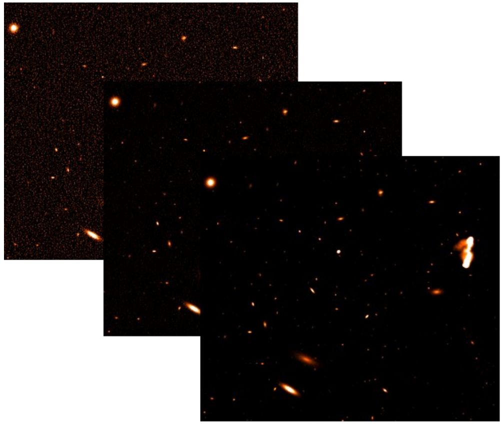
→ Requires new innovative methods

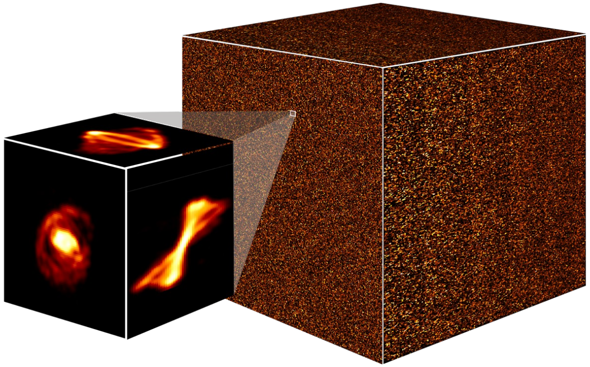
Science Data Challenges (SDCs)

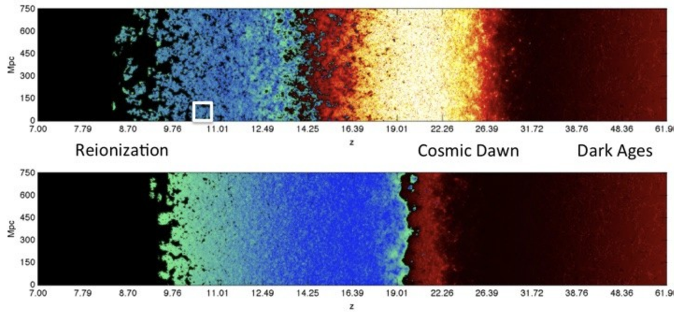
Simulated datasets that should resemble typical SKA data products

Objective: prepare astronomers, stimulate the creation of new data analysis pipeline

Source detection and characterization







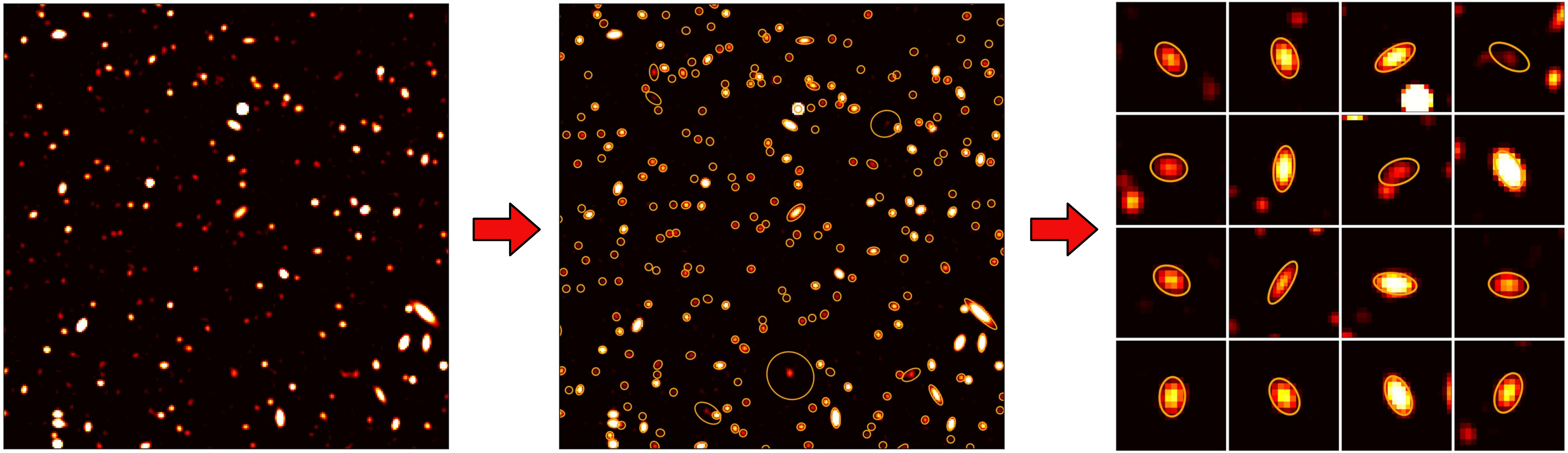
SDC1: Continuum 2D images
3 integration times x 3 bands
Each image = 4 GB
From Dec 2018 to April 2019

SDC2: Hyperspectral cube of HI emission
Full cube = 1 TB
From Feb 2021 to July 2021

SDC3: 21 cm emission
Visibility and Image
Full size ~ 7 TB
EoR Focused, 2023-2025

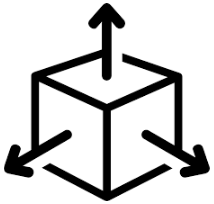
Astronomical source detection and characterization

Survey (images) **Catalogs (lists)**



Automated detection tools :

PyBDSF, SFIND, SourceExtractor, BLOBCAT, Aegan, Selavy, SoFiA, ...



Problem :

1 – Classical method scalability

2 – Robustness and completeness

Object detection in computer vision

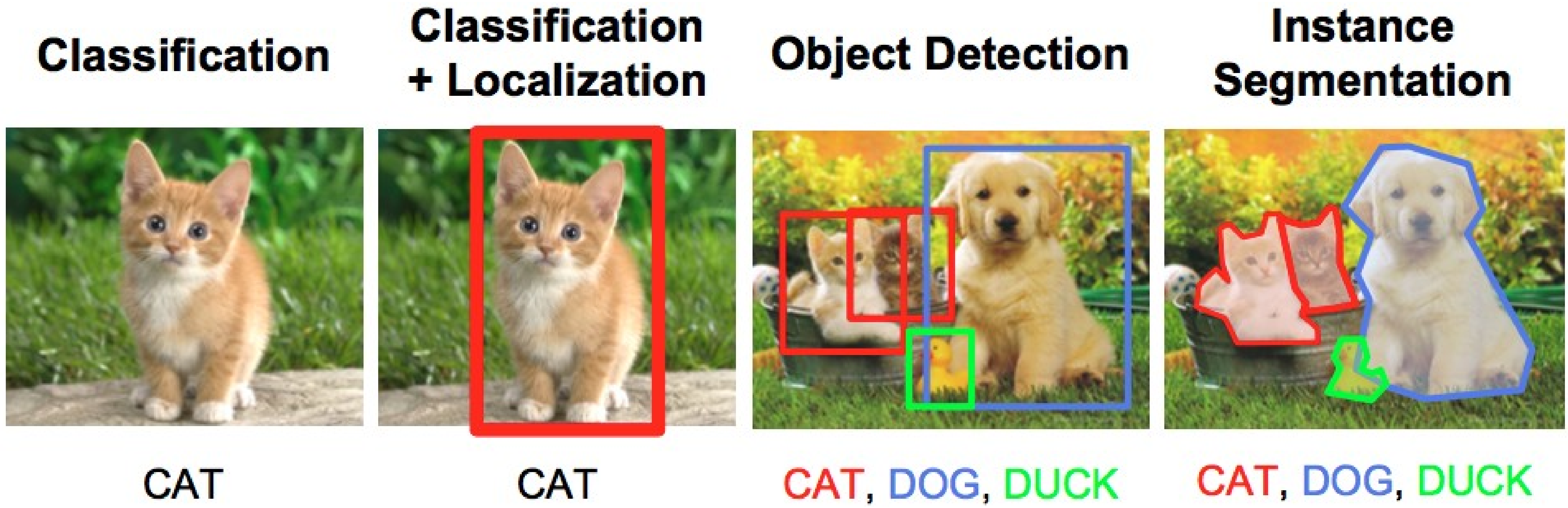
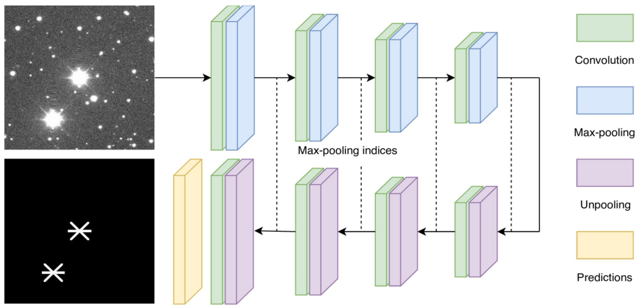


**Image from Stanford Deep Learning course cs224*

Advances methods for object detection

U-Nets

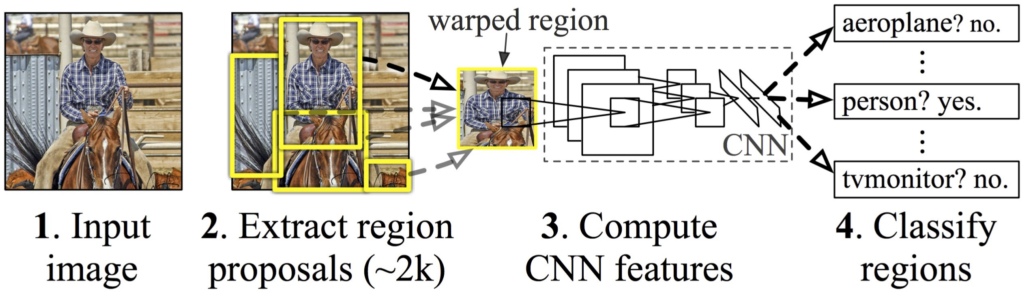
Pros: produce segmentation maps, access to latent space, ...



Region-based

Methods : R-CNN (Fast and Faster), SPP-net, Mask R-CNN, ...

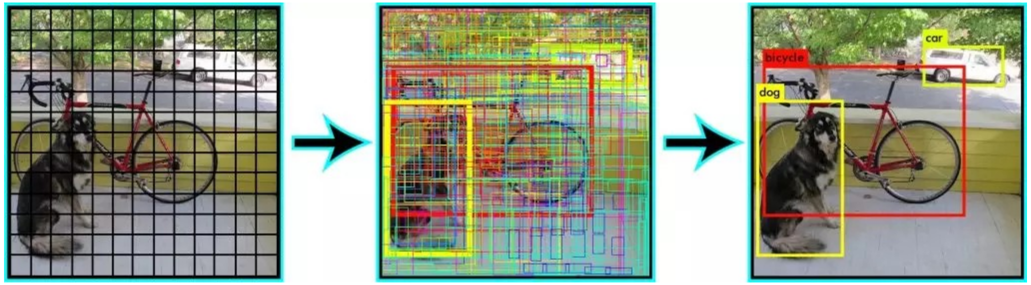
Pros: Best accuracy, ...



Regression-based

Methods : SSD (Single Shot Detector), YOLO (You Only Look Once), ...

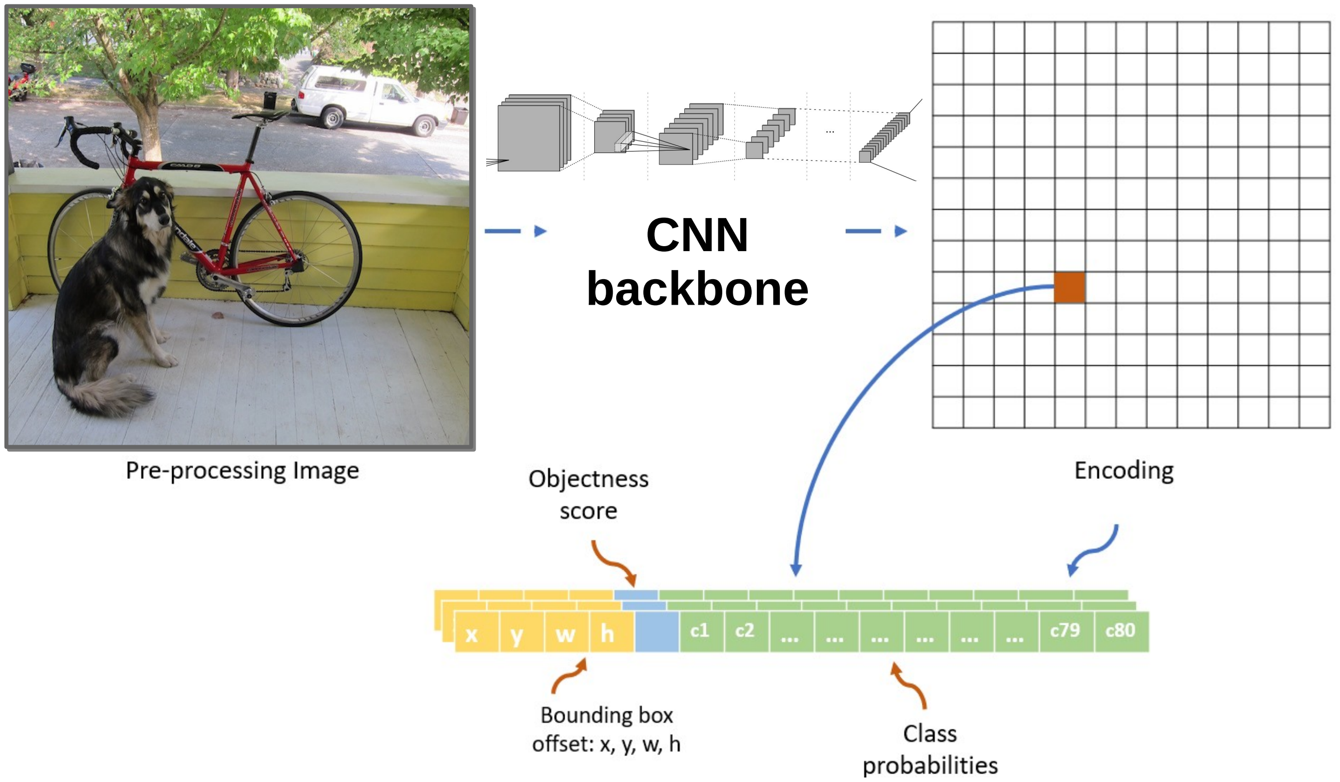
Pros: Very Fast, straightforward architecture,...

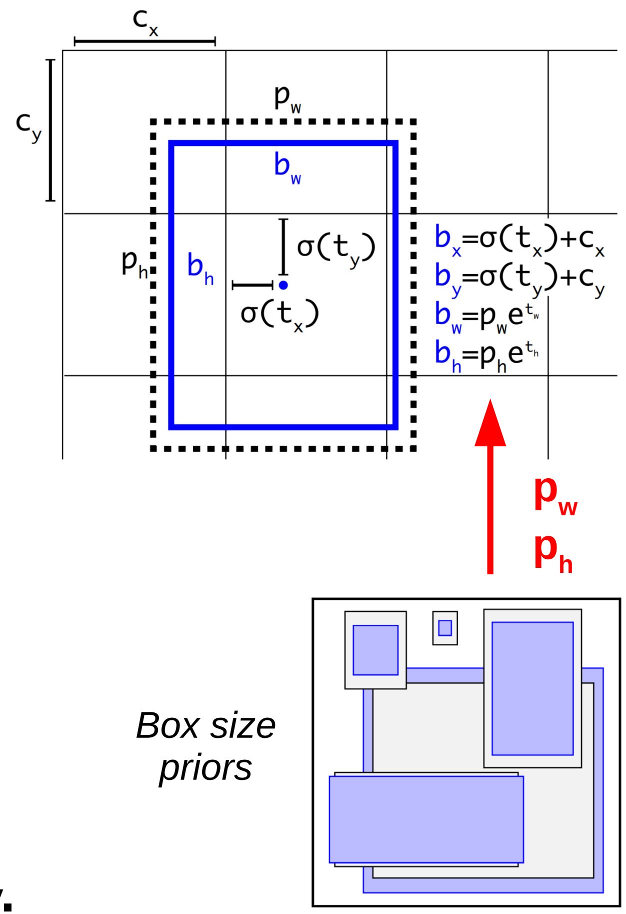


The You Only Look Once (YOLO) detector

Originally introduced in Redmon et al. 2015 (V1), 2016 (V2), 2018 (V3)

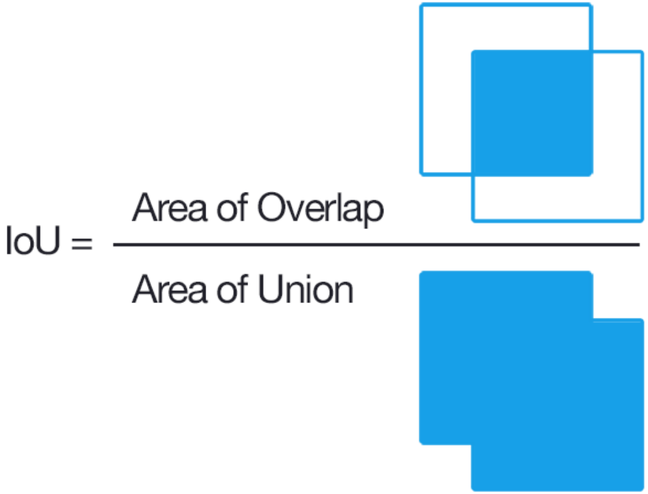
**Images from blog post and Redmon papers*

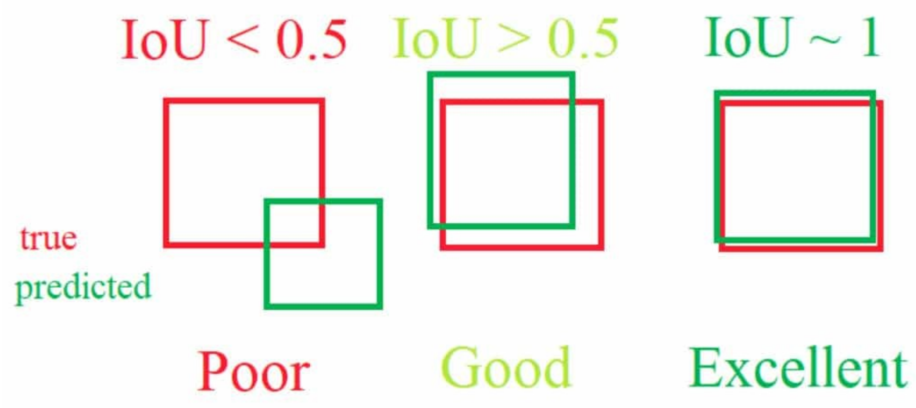




The last layer is conv. = the boxes « share » weights spatially.
The output is a 3D cube encoding all possible boxes on the output grid.

How to estimate box « correctness » ?

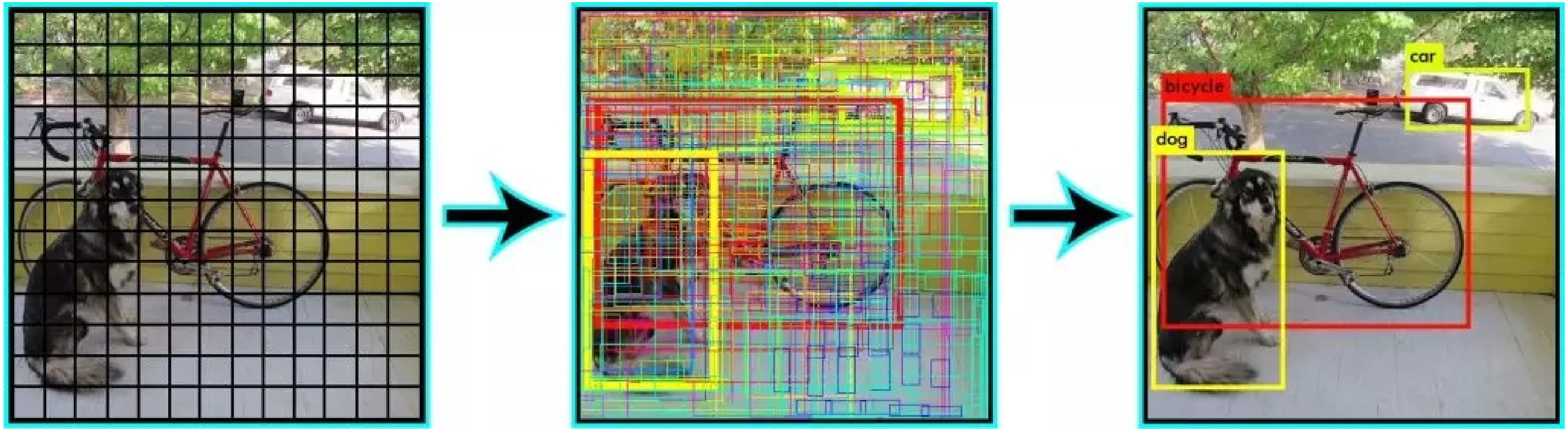




For each box the network will predict an "objectness" score defined as:

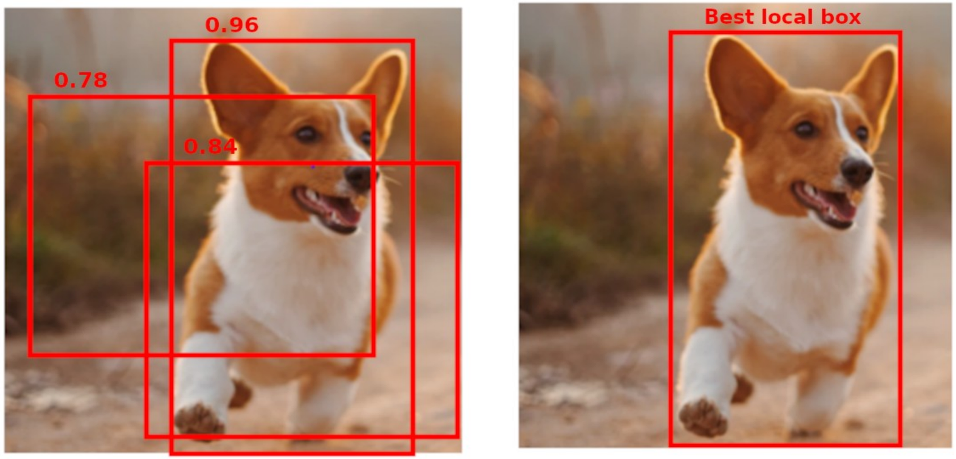
$$Pr(\text{object}) * IOU(b, \text{object}) = \sigma(t_o)$$

Post Processing: filtering predictions



1) Most probable boxes are kept using a threshold in **objectness**

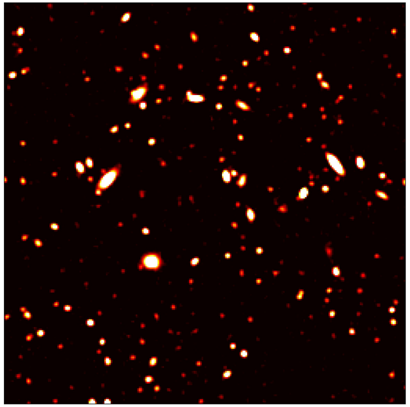
2) Removes **multiple detections** of single objects using **Non Max Supression**



Using DL methods for astronomical data

Astronomical images and objects do not fulfill classical implicit biases

≠



- **No sharp edges**
- **Source blending**
- **Disjoint objects**
- **Extreme density** ***(high or low)***
- **Large dynamic**
- **Noise**

YOLO → YOLO CIANNA

Redmon et al. 2016

Cornu et al. 2024

Complete rework for astronomical applications

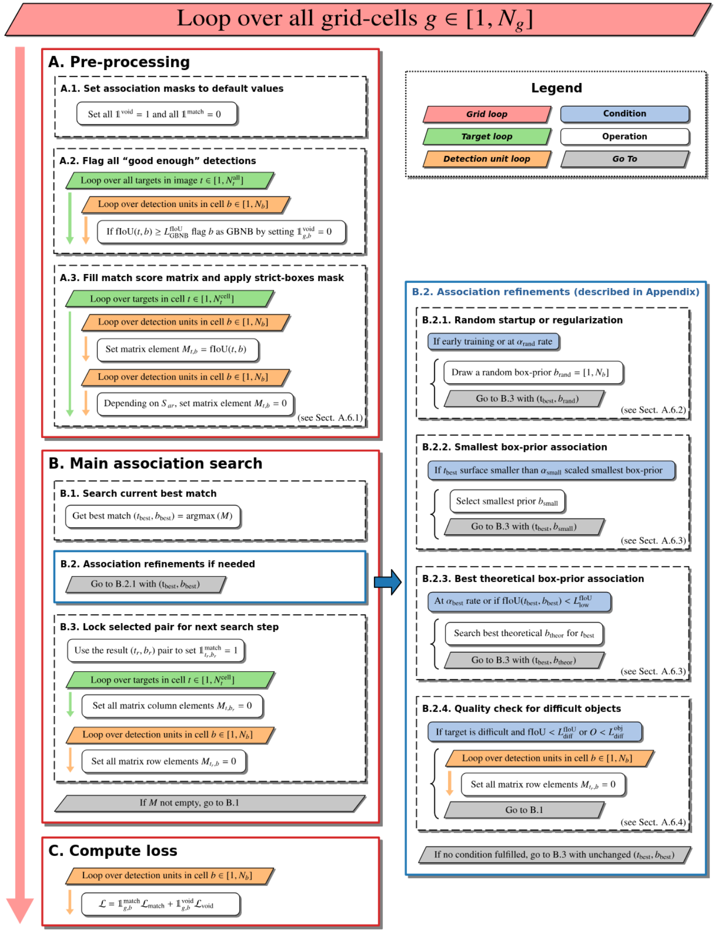
- ***Sub-pixel positioning***
- ***Accurate flux quantization***
- ***Resilient to high density***
- ***On the fly characterization***
- ***3D detection*** ***(end-to-end 3D conv.)***

How to train a YOLO-CIANNA detector ?

Well, its long to explain → Cornu et al. 2024 (42 pages paper, 11 of method)!

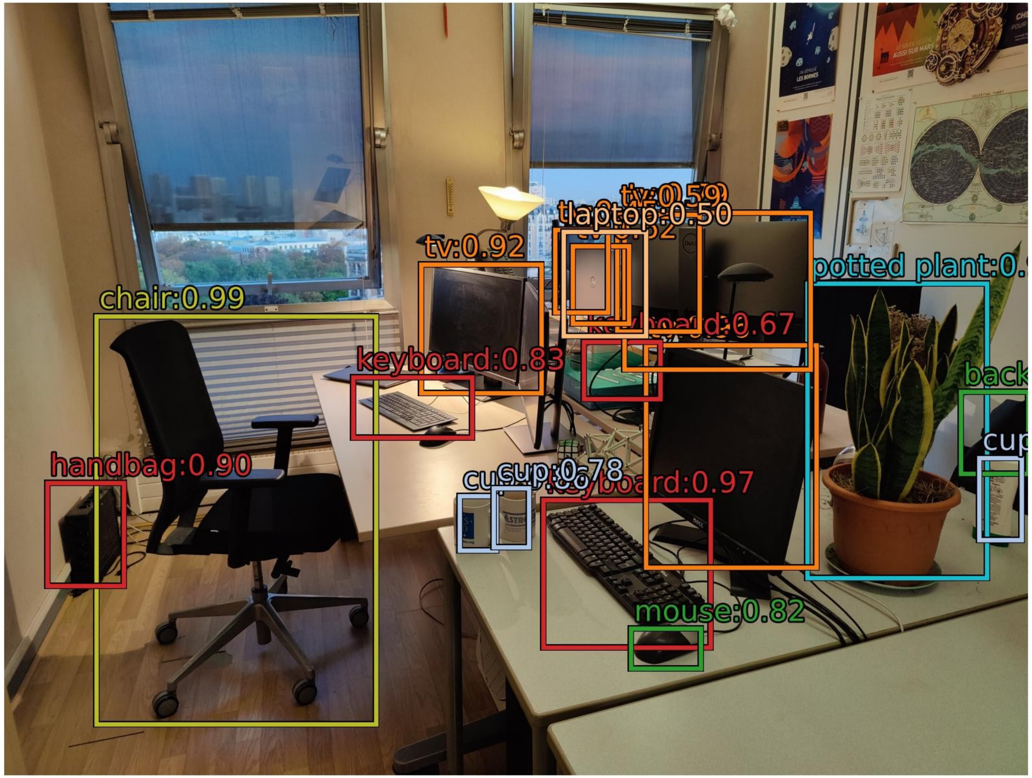
→ *Don't worry we've made the hard work for you.*

$$\mathcal{L} = \sum_{i=0}^{N_g} \sum_{j=0}^{N_b} \mathbb{1}_{ij}^{\text{match}} \Bigg(\lambda_{\text{pos}} \left[(o_{ij}^{x} - \hat{o}_{ij}^{x})^2 + (o_{ij}^{y} - \hat{o}_{ij}^{y})^2 \right]$$
$$+ \lambda_{\text{size}} \left[(o_{ij}^{w} - \hat{o}_{ij}^{w})^2 + (o_{ij}^{h} - \hat{o}_{ij}^{h})^2 \right]$$
$$+ \lambda_{\text{class}} \quad \mathbb{1}_{ij}^{C} \sum_{k}^{N_C} \left(-\hat{C}_{ij}^{k} \log(C_{ij}^{k}) \right)$$
$$+ \lambda_{\text{param}} \quad \mathbb{1}_{ij}^{p} \sum_{k}^{N_p} \gamma^{k} \left(p_{ij}^{k} - \hat{p}_{ij}^{k} \right)^2$$
$$+ \lambda_{\text{prob}} \quad \mathbb{1}_{ij}^{P} \left(P_{ij} - 1 \right)^2$$
$$+ \lambda_{\text{obj}} \quad \mathbb{1}_{ij}^{O} \left(O_{ij} - \text{fIoU}_{ij} \right)^2 \Bigg)$$
$$+ \sum_{i=0}^{N_g} \sum_{j=0}^{N_b} \mathbb{1}_{ij}^{\text{void}} \lambda_{\text{void}}^{j} \Big(\lambda_{\text{prob}} \left(P_{ij} - 0 \right)^2 + \lambda_{\text{obj}} \left(O_{ij} - 0 \right)^2 \Big).$$



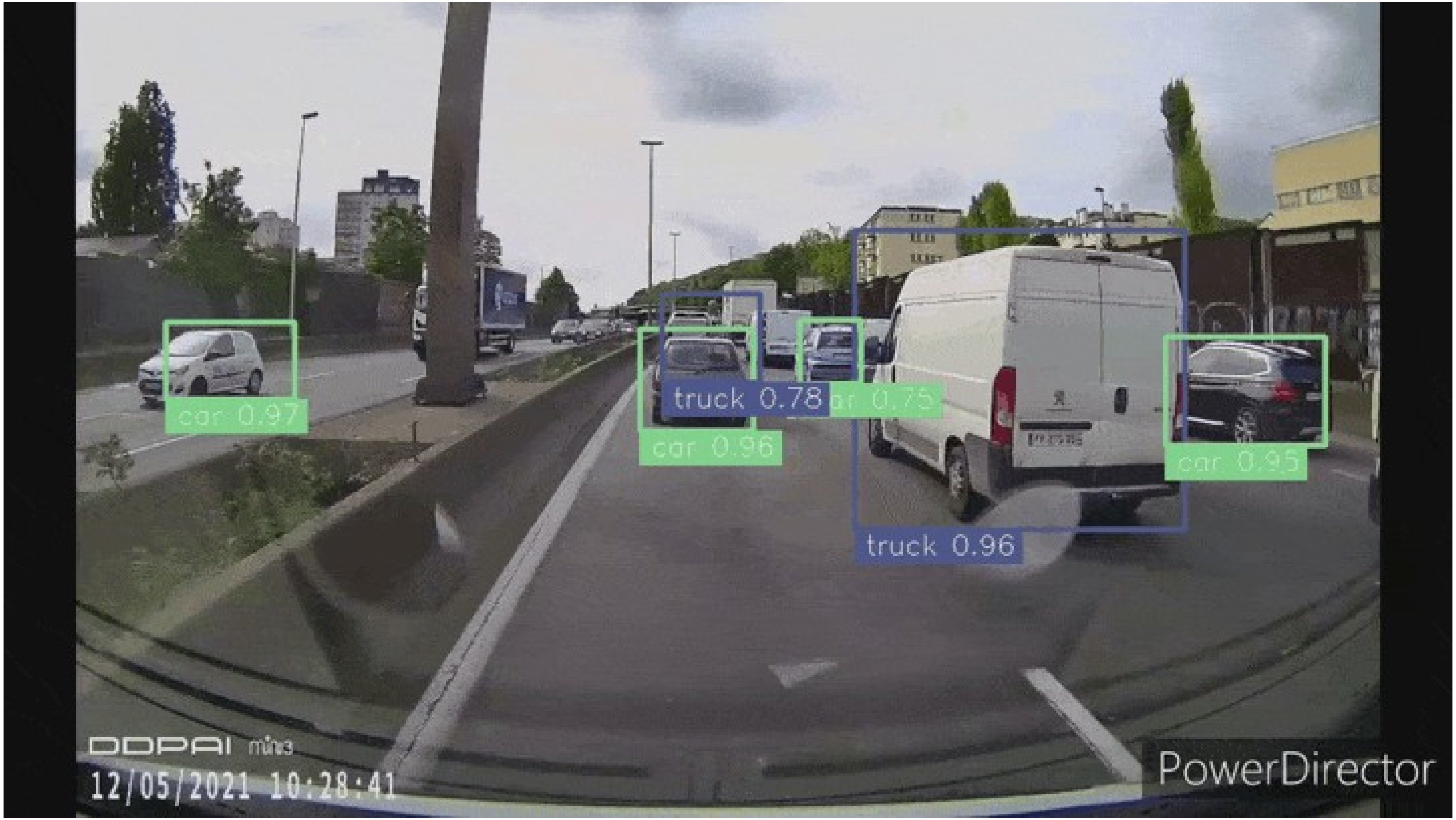
Examples of YOLO-CIANNA detections

Model pre-trained on ImageNet and then trained on the COCO datasets independently.



Examples of YOLO detections

Result using a YOLO detector with a darknet-19 “like” backbone trained on COCO.
The network run at 690 ips at a 416p resolution on an RTX 4090.



Application to SKAO SDC1

Data:

Simulated continuum image:

- 5.5 square degree area (pixel size 0.6'')
- 560 MHz, 1000h integration time
- **4GB image (32,768 pixel square)**

The challenge:

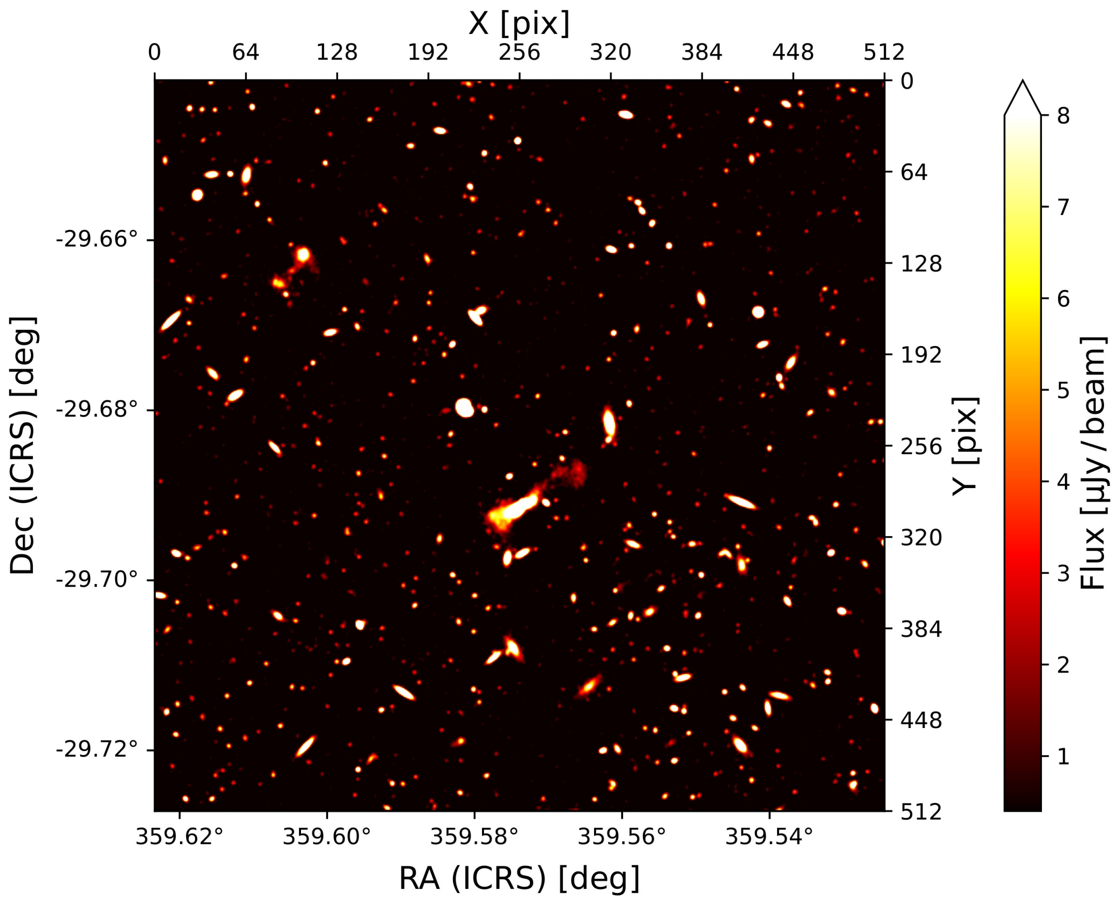
1. Find the sources (RA, Dec)
2. Characterize each source:
   → (Flux, Bmaj, Bmin, PA, ...)

Training labels provided for a subpart of the image (5% of the surface, ~34 000 sources).

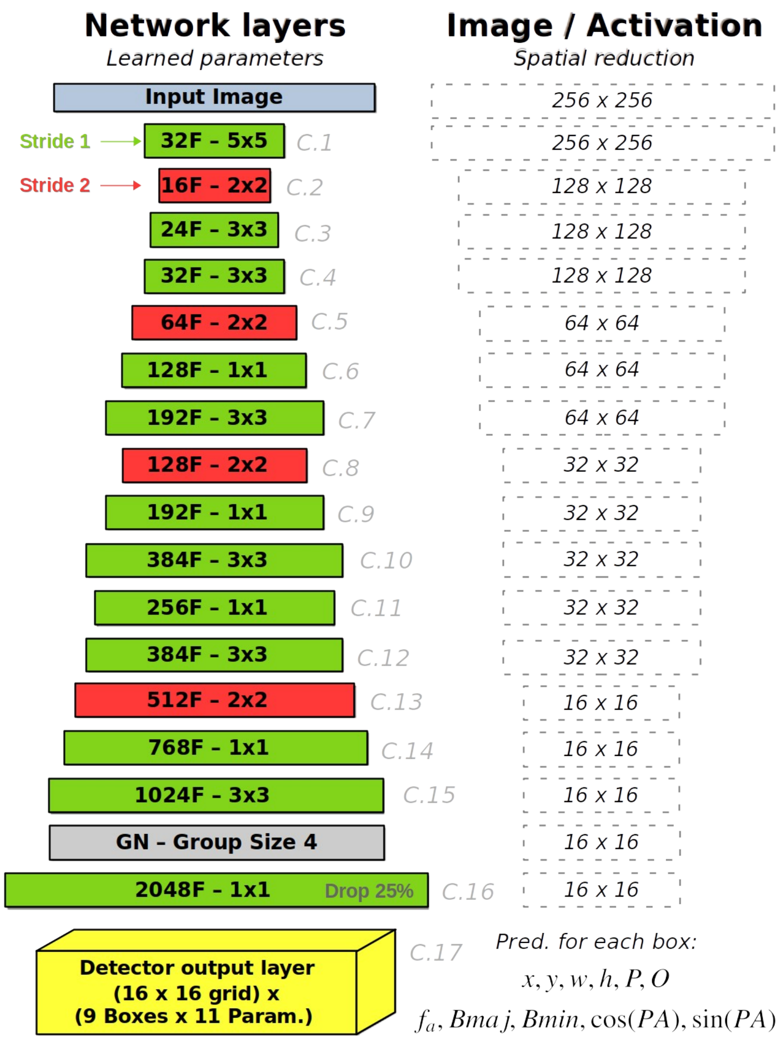
→ SDC1 summary paper, Bonaldi et al. 2021

SKA SDC1 took place early 2020. Data from the challenge are still freely accessible on the dedicated web-page.

Example 512² sub-field



Custom SDC1 model



Architecture:

- **17 conv. layers** → **~13 Million parameters (~50MB)**
  → +8% in score compared to the classical darknet19 backbone
- **9 box priors** ranging from 10 to 32 pixels
- **5 additional parameters**: **Flux, Bmaj, Bmin, cos(PA), sin(PA)**

Training the network using IANNA

- **256x256** cutouts are randomly selected in the training area (54 MB)
- ~ **34000 sources** in the selected training catalog
- Data are augmented based on cutout position and flips

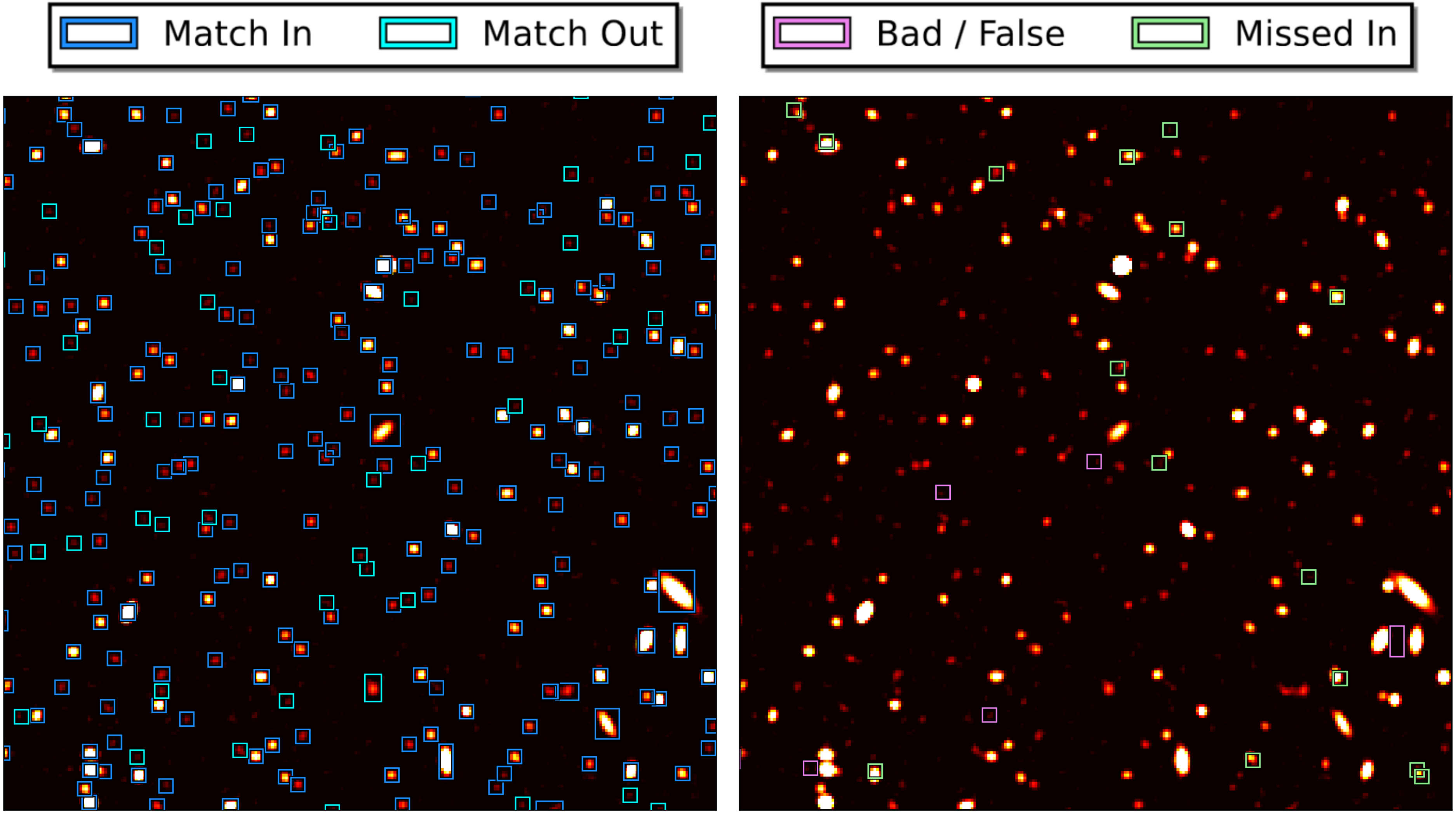
Using a single RTX 4090 GPU, training time is **~ 4 hours**

Inference:

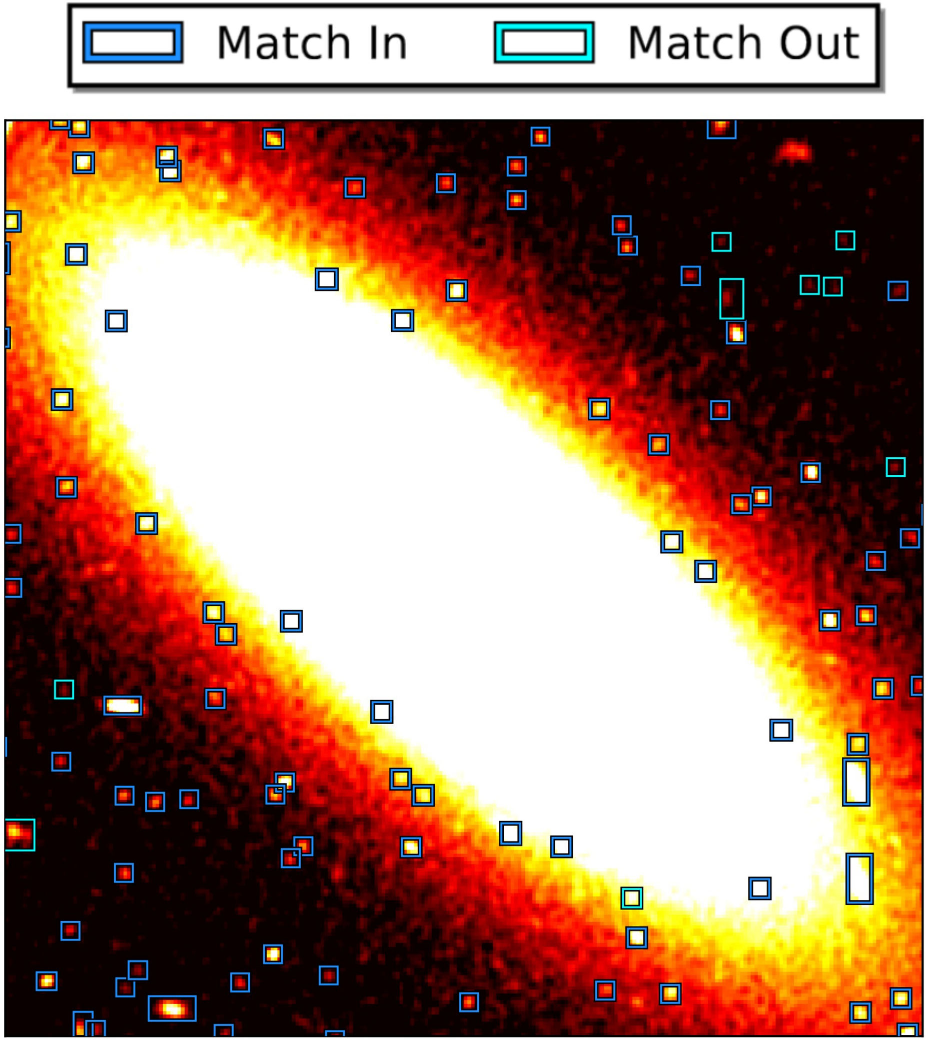
- The full SKA SDC1 image is split in 512x512 regions
  with an overlap of 32 pixels, → **~4500 images**
- Overlapping regions are filtered with a dedicated secondary NMS

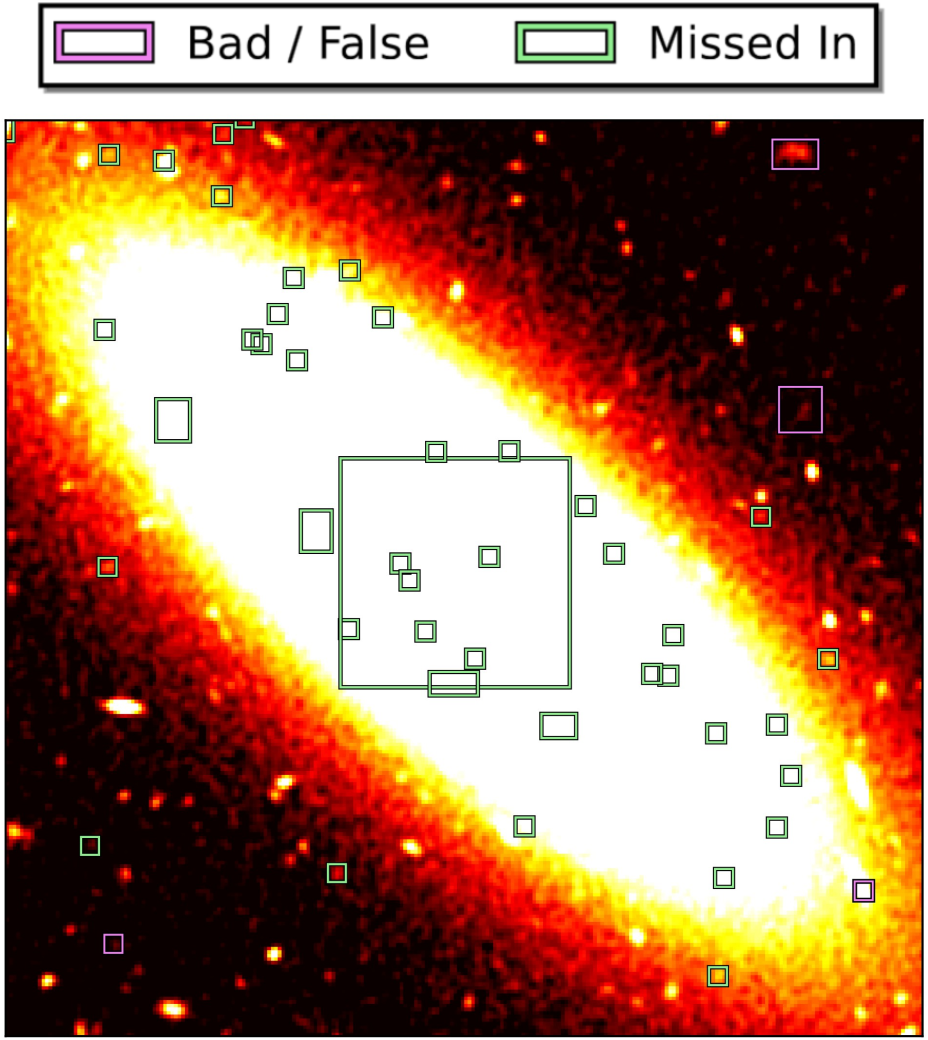
The full inference in FP16-TC takes ~8 sec → 130 Mpix/s

Detection example fields



Detection example fields





Results comparison

Based on Bonaldi et al. 2021 + submitted catalogs

$$s_i = \frac{1}{7}\sum_{j}^{7} s_i^j \qquad M_s = \sum_{i}^{N_{match}} s_i - N_{false}$$

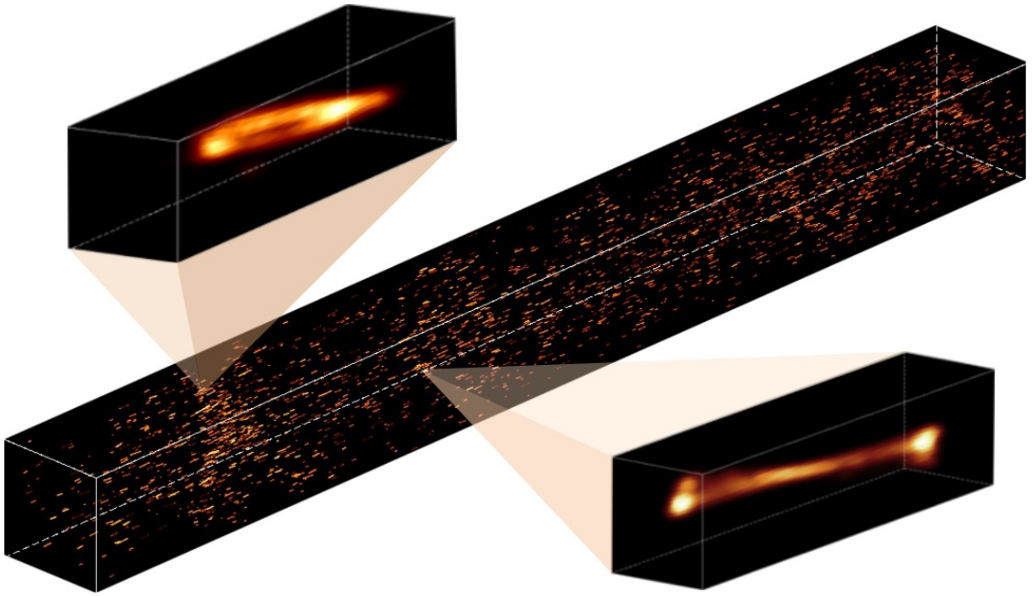
Average per-source score

	Team (method)	M_s (Score)	N_{det}	N_{match}	N_{false}	$N_{bad} \in N_{false}$	Purity	$\bar{s}$	
	Post-challenge results								
ML (CNN)	MINERVA (YOLO-CIANNA)	**480450**	724480	680000	44480	16839	93.86%	0.7719	*After challenge ending*
	↪ *purity-focus thresholds*	418434	541542	536412	5130	2506	**99.06%**	0.7896	
ML (CNN)	JLRAT2 (JSFM2)	**298201**	502146	484212	17934	2274	96.43%	0.6529	
	Original challenge results								
Classical	Engage-SKA (PROFOUND)	**200939**	421992	418384	3608	2677	99.15%	0.4889	*Original leaderboard*
Classical	Shanghai (multiple methods)	**158841**	292646	291553	1093	698	99.63%	0.5486	
ML (CNN)	ICRAR (CLARAN)	**142784**	279898	259806	20092	6875	92.82%	0.6269	
	7 other teams, one other ML attempt		...	...	...	...	...	...	

→ SDC1 is still a very interesting dataset for source detection pipeline development !

YOLO-CIANNA SDC1 → Cornu et al. 2024

SCIENCE DATA CHALLENGE 2



Data: a 3D cube of simulated HI emission

- 20 square degree area
- 950 to 1150 MHz frequency (30KHz res; z = 0.235–0.495)
- **2000h integration time**
- **Near 1 TB cube (5851 x 5851 x 6668)**

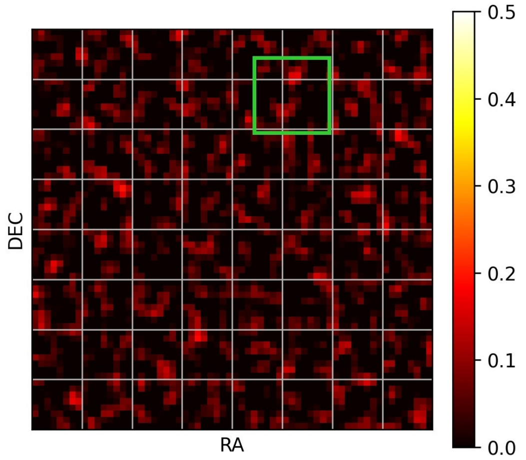
The challenge:

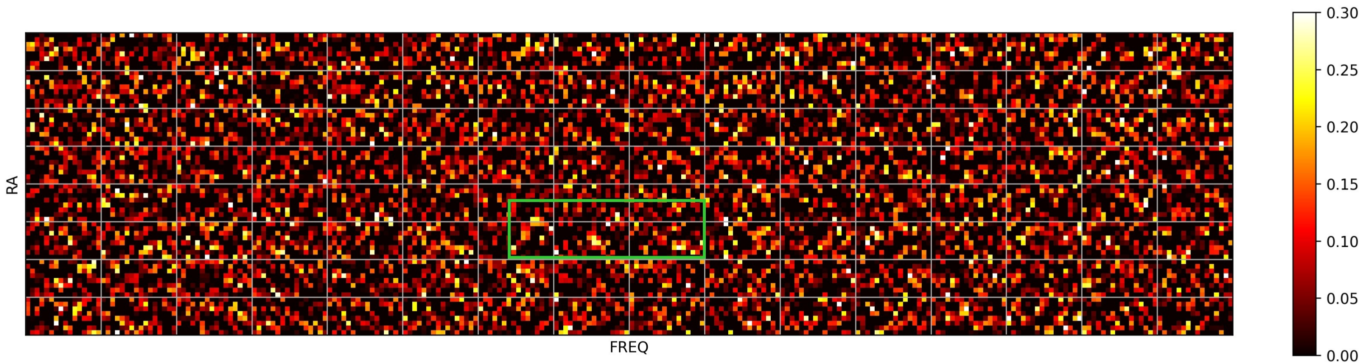
1. Find the sources (RA, Dec, Freq)
2. Characterize each source:
   → Flux, HI size, line width, PA, Inclination

→ SDC2 summary paper, Hartley et al. 2023

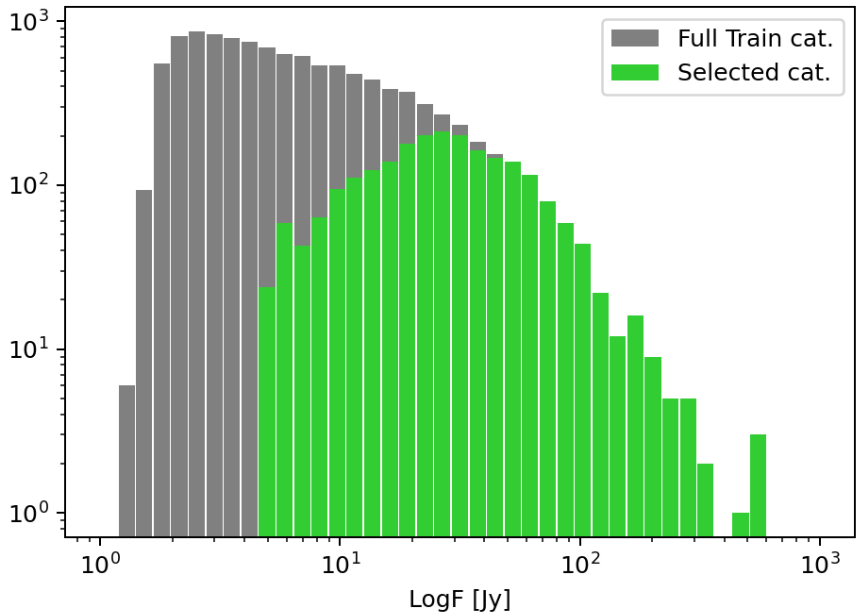
SKA SDC2 took place in 2021. Data from the challenge are still freely accessible on the dedicated web-page.

Selection function difficulties



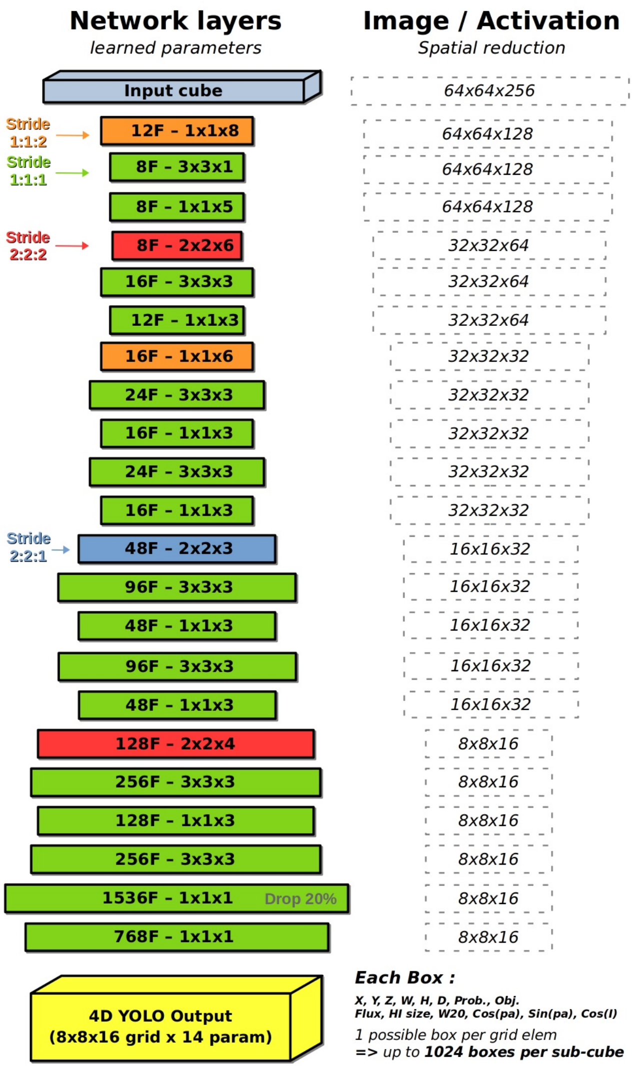


Selection function based on brightness or SNR are not sufficient to fully represent the noisy 3D information.



→ "Bootstrap learning" (~active learning):
After a first training, **un-selected true sources** with **high predicted objectness** can be re-injected in the training sample.

Model modifications for SDC2



YOLO parameters:

- **Generalized to 3D detection**
- **23 layers ~ 4 Million parameters**
- 1 single box prior per grid element
- Predict 6 additional source parameters
- No class prediction

Training the network using IANNA

- **64x64x256** cubes are randomly selected in the training area (40 GB)
- Around **2000 sources** in the selected train catalog
- Data are augmented using shifting and flips

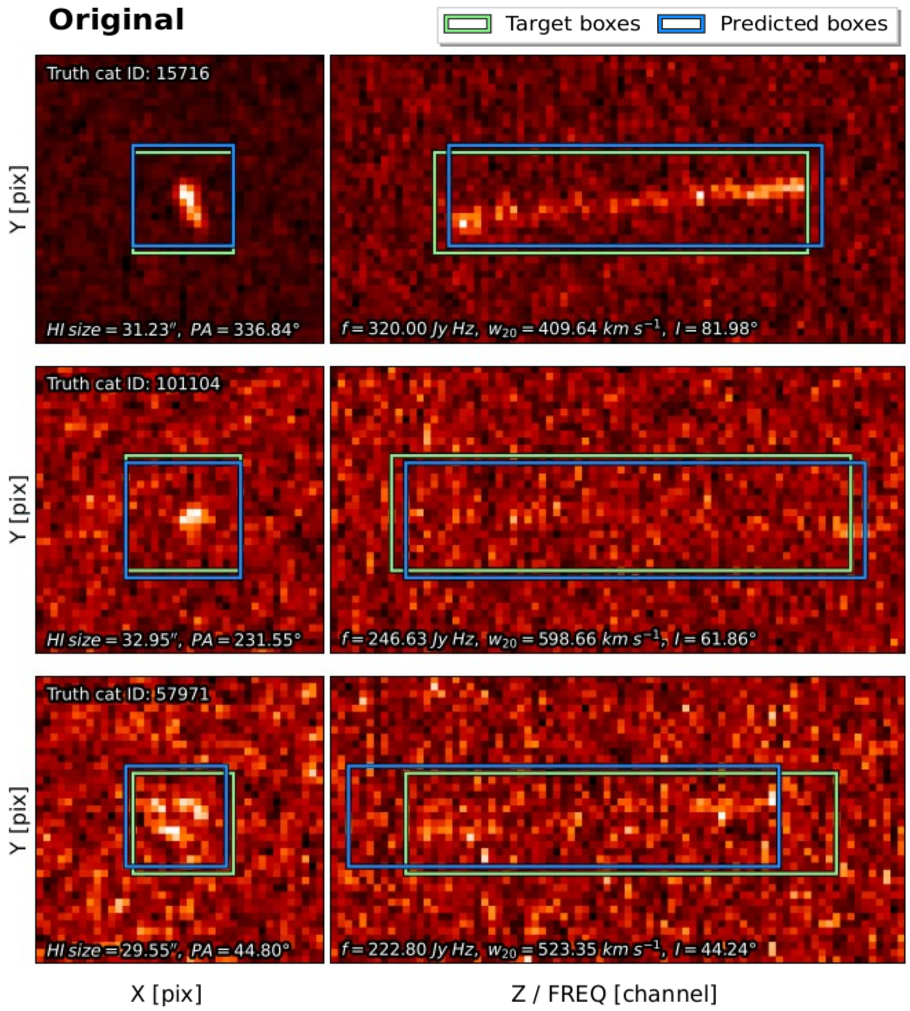
Using a single RTX 4090 GPU
→ Training time up to **12 hours** (already good results after 6-8h).

Inference:

- The full SKA SDC2 1 TB cube is split in regions with large overlaps
- Box in overlapping regions are filtered with a dedicated NMS

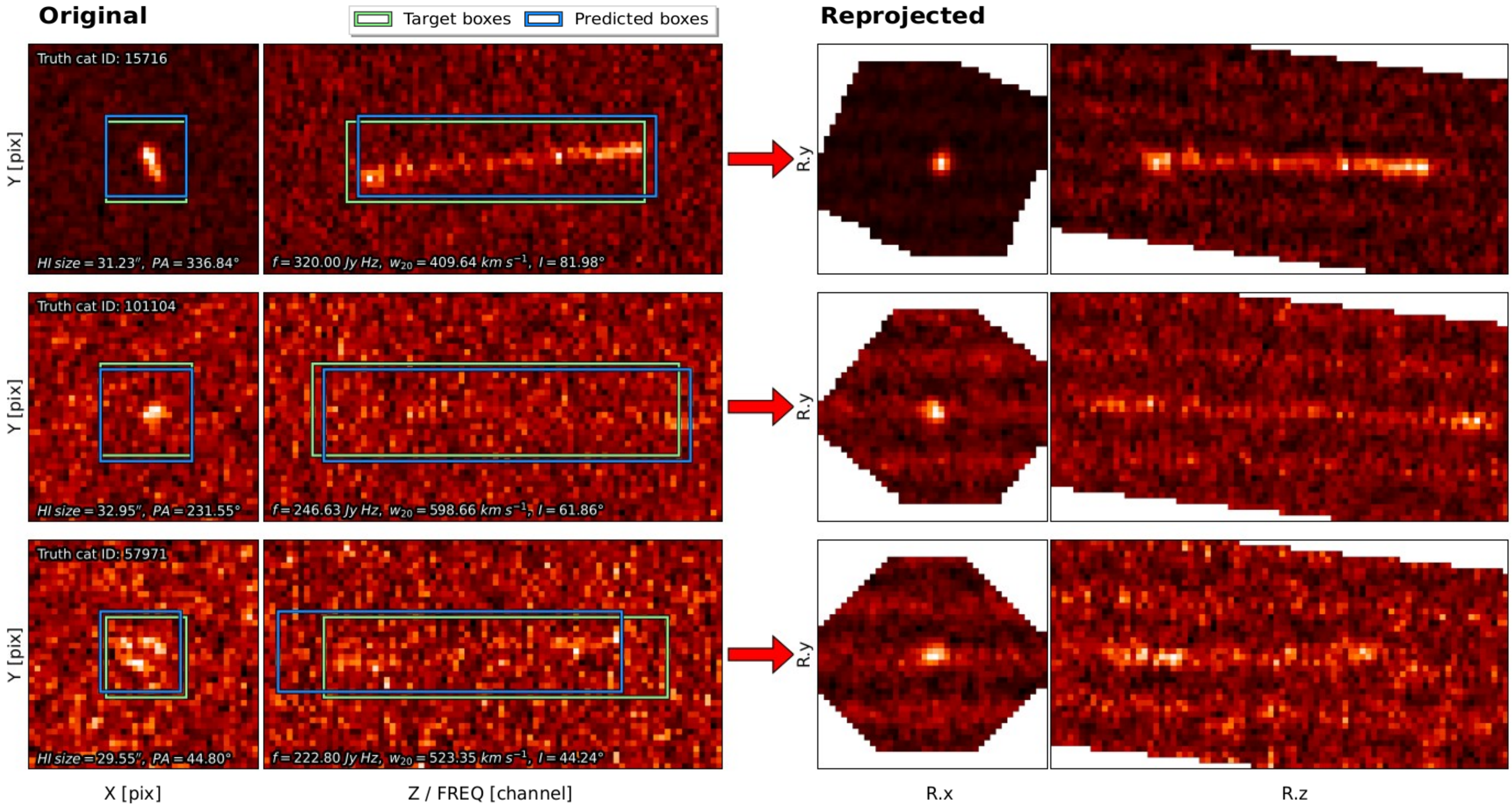
The full cube prediction takes ~1 hour (vastly dominated by data loading time) **using a single RTX 4090** ***(raw 260 ips)***

Detection examples



Images are based on 40x40x120 cutouts centered on a source. Signal is averaged over the source dimension in the projected axis.

Detection examples



Images are based on 40x40x120 cutouts centered on a source. Signal is averaged over the source dimension in the projected axis.

LEADERBOARD

Results from Hartley et al. 2023

	Team name	Score	N_d	N_m	Accuracy
ML	MINERVA New*	23482	34441	31709	83
ML	MINERVA	23254	32652	30841	81
ML + SoFiA	FORSKA-Sweden	22489	33294	31507	77
SoFiA	Team SoFiA	16822	24923	23486	78
SoFiA	NAOC-Tianlai	14416	29151	26020	67
SoFiA	HI-FRIENDS	13903	21903	20828	72
Wavelets + ML	EPFL	8515	19116	16742	65
SoFiA	Spardha	5615	18000	13513	75
SoFiA	Starmech	2096	27799	17560	70
ML	JLRAT	1080	2100	1918	66
Wavelets + ML	Coin	-2	29	17	60
ML	HIRAXers	-2	2	0	-
Other	SHAO	-471	471	0	-

***Minerva:** *YOLO-only score, obtained after the challenge end*

***MINERVA:** YOLO and CHADHOC combination*

***FORSKA**: U-Net segmentation, parameters using SOFIA (Håkansson et al. 2023)*

***EPFL:** Denoising with 3D wavelet filtering, identification with jointed likelihood, Parameters with several CNNs*

***JLRAT:** Region proposal CNN detection, classical for parameters*

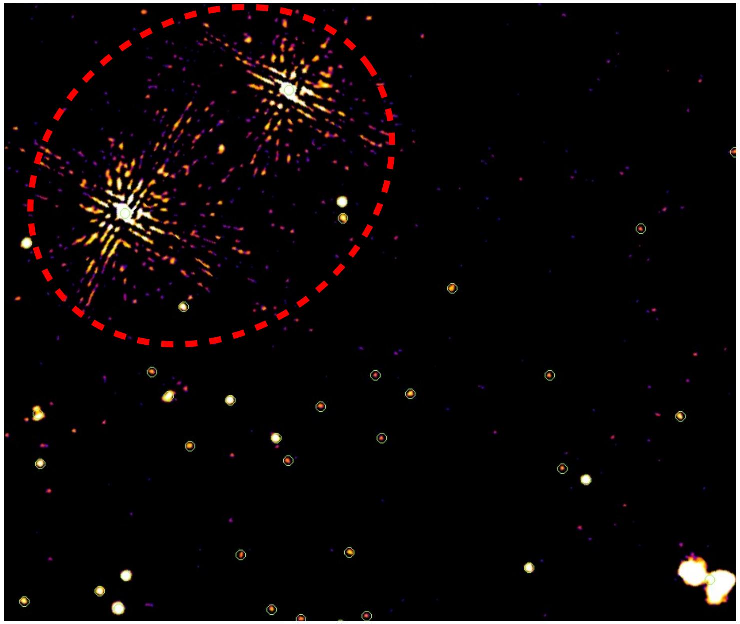
***Coin:** Multiple CNNs for detection and dedicated CNNs for parameters*

***HIRAXers:** Multiple CNNs for both detection and for parameters*

YOLO-CIANNA 3D SDC2 → Cornu et al. 2025

Challenges of switching to real data

Example on the LoTSS survey (LOFAR)



Difficulties : Artifacts / Noise / Resolution / Sizes / Morphology

How to define the training sample?

- **Use costly observations on few sources**
  Pros: Very robust labels **Cons:** few examples & imbalance
- **Use classical detection methods!**
  Pros: Easy to use, large samples **Cons:** possible bias
- **Use simulations (e.g SKA SDCs models)**
  Pros: infinite examples **Cons:** bias, instrument model required
- **Use Citizen Science (e.g Radio Galaxy Zoo)**
  Pros: "Easy" **Cons:** bias / errors, limited to human capability
- **Combine all of the above!**
  Pros: Very complete / diverse **Cons:** difficult to balance
- **Self / Active - Learning or Unsupervised**
  Train with one sample, then use one of the above to refine « new candidates », or try various flavor of unsupervised methods
  Pros: limits defined by the method and the data themselves, less human bias.