Exoplanet Imaging Data Challenge:

benchmarking image processing methods for exoplanet direct detection

Faustine Cantalloube[1], Carlos Gomez-Gonzales[2,3], Olivier Absil[4], Carles Cantero[4,5]

Bacher R., Bonse J. M., Bottom M., Dahlqvist C.-H., Desgrange C., Flasseur O., Fuhrmann T., Henning Th., Jensen-Clem R., Kenworthy M., Mawet D., Mesa D., Meshkat T., Mouillet D., Müller A., Nasedkin E., Pairet B., Pierard S., Ruffio J.-B., Samland M., Stone J. and , Van Droogenbroeck M.

European Research Council
Established by the European Commission

CONTEXT

The **Exoplanet Imaging Data Challenge** (EIDC) is a community-wide effort meant to offer a platform to enable a fair and common comparison of the various image processing methods dedicated to exoplanet direct detection.
https://exoplanet-imaging-challenge.github.io/

Open-source: data hosted on *ZenoDo*, competition on *CodaLab*
Benchmarking: (1) to support observers / users
(2) to guide publications of new algorithms
Sparking collaborations in the post-processing community

Phase 1: *from 09/2019 to 10/2020*

Focused on **detection** capabilities of the algorithms.
Pre-phase to receive feedbacks ended on 01/2020.
A **workshop** took place in 01/2020 to discuss the outcome.

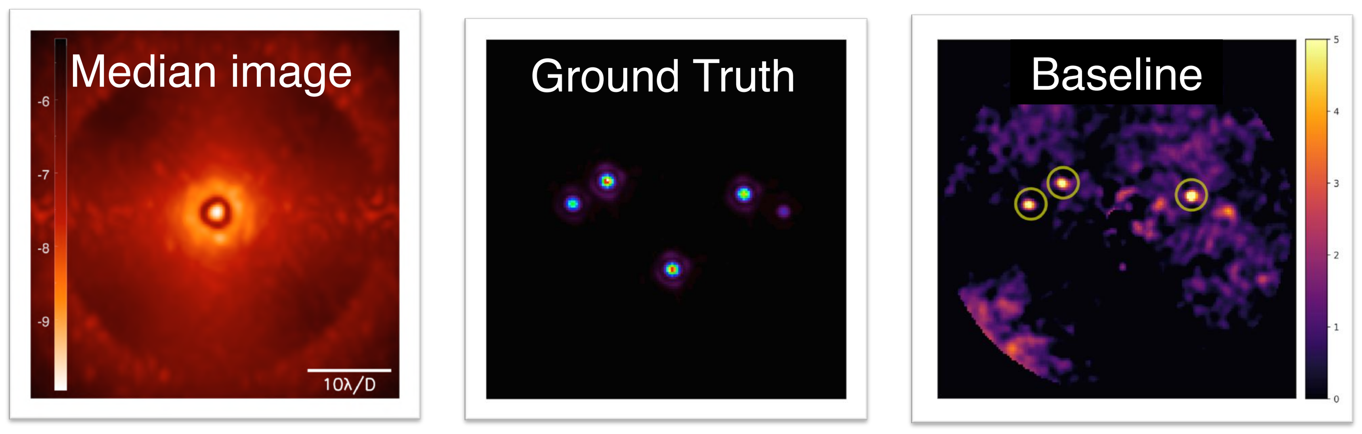
DATA SETS

Because the performance of a given image processing method may be dependent upon the instrument and the observing conditions, we used several datasets from different high-contrast instruments.

High-contrast instruments:

(1) ADI subchallenge: temporal cube in pupil tracking (PT)
9 data sets from 3 instruments
VLT/SPHERE-IRDIS using an Apodized Lyot Coronagraph (H-band)
Keck/NIRC2 using an Annular Groove Phase Mask (L-band)
LBT/LMIRCam without coronagraph (L-band)

(2) ADI+mSDI subchallenge: multispectral cube in PT
10 datasets from 2 instruments
VLT/SPHERE-IFS using an Apodized Lyot Coronagraph
Gemini-S/GPI using an Apodized Lyot Coronagraph



Synthetic planetary signals Injections:

In each dataset, we injected from **0 to 5** synthetic planetary signals, using the **inverse** parallactic angles to smear out potential signals.
The injection separation and contrast are **randomly** picked in a range close to the detection limit (contrast curve) from the chosen baseline.
The **Baseline** algorithm is an annular Principal Component Analysis.
For the ADI+mSDI injections, **spectral features** are injected.
The injections are made using the VIP package: the synthetic planetary signals are injected without smearing, without photometric variation in time, and assuming a given center fixed for all images.

PHASE 1: EVALUATION

Required input from participants:

From running a given algorithm on **all** the pre-reduced datasets, the participants had to provide:
- A detection map for each dataset,
- A single threshold value for all datasets.

Counting the detections:

By definition, any signal above threshold triggers a detection.
We considered only one detection per unit of resolution element (~1λ/D), computed from the instrumental point-spread function.
For various thresholds, we counted the true positives (TP), the true negatives (TN), the false positives (FP), the false negatives (FN).
From these we derived:

- True positive rate: TPR = TP/(TP+FN)
- False discovery rate: FDR = FP/(FP+TP)
- False positive rate: FPR = FP/(FP+TN)

At the submitted threshold, we compute:

- F1-score = 2 TP / (2 TP+FP+FN)

All these scores must be between 0 and 1.
In the absence of injections (no positive), the TPR and the F1-score are undefined.
To study the sensitivity of the algorithms, we display the TPR, FPR and FDR scores with respect to a varying threshold value (**FIG. 1**).

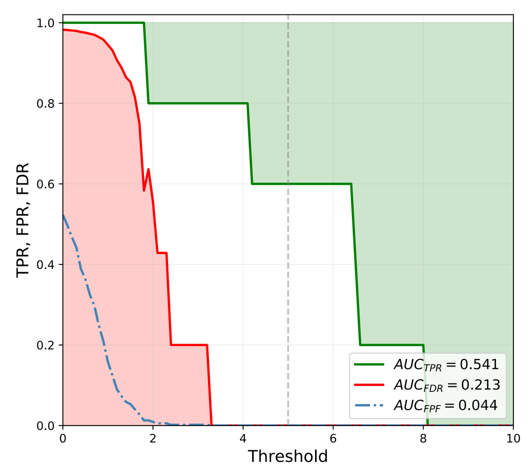


FIG.1: The green area (TPR) and the red area (FDR) must be minimal. The blue line (FPR) gives information about the residuals.

The comparison metrics will be refined in a future dedicated paper.

PHASE 1: SUBMISSIONS

For this first phase, **65 people** registered on the *CodaLab* platform.

(1) ADI subchallenge: 22 valid submissions from **12** participants:
We separated the submissions in 3 families:
(i) Speckle subtraction techniques: the most widely used techniques, providing either a residual map or a detection map.
→ **12 submissions:**
cADI, PCA, LOCI, STIM map, RSM map
(ii) Inverse problem approaches: these techniques make a model of the expected planetary signal and track it in the data.
→ **5 submissions:**
ANDROMEDA, FMMF, PACO, TRAP
(iii) Supervised machine learning: after applying PCA, the algorithm is trained to classify detection vs. non-detections.
→ **5 submissions:**
SODIRF, SODINN

(2) ADI+mSDI subchallenge: 4 valid submissions from **3** participants. 1 submission is a speckle subtraction technique and the 3 others are based on an inverse problem approach.
PCA-ASDI, PACO-ASDI, FMMF, ANDROMEDA

SOME RESULTS...

In the corresponding proceeding, you will find more details about the submitted algorithms and the comparison.

Subchallenge 1: The detection maps below are for the VLT/SPHERE-IRDIS dataset containing 5 injections.

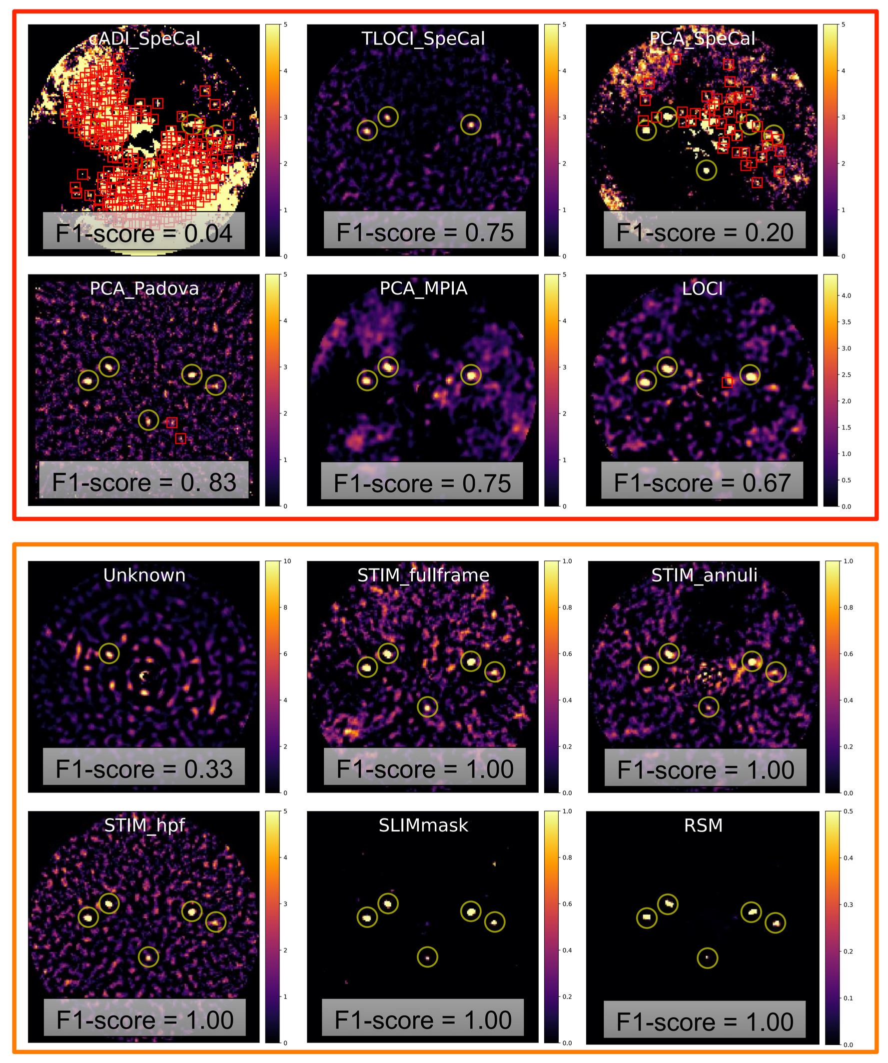
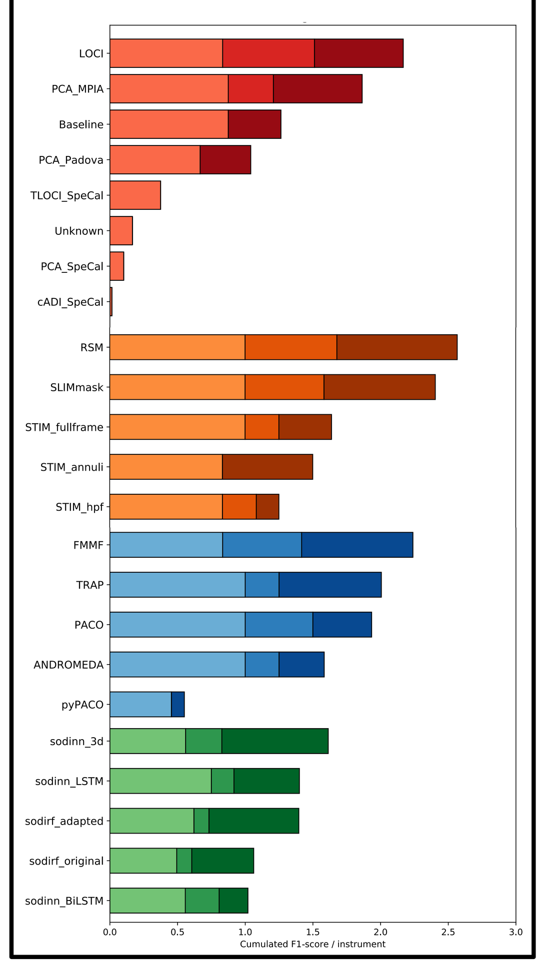


FIG. 3: Ranking based on the F1-score for the three instruments.



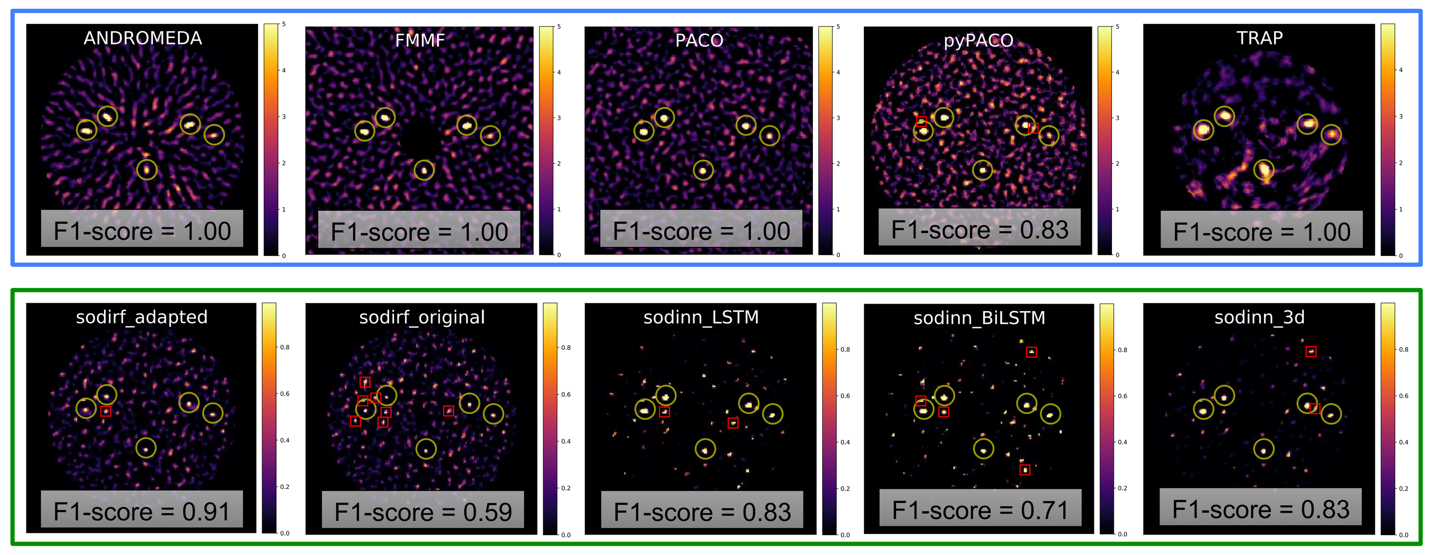
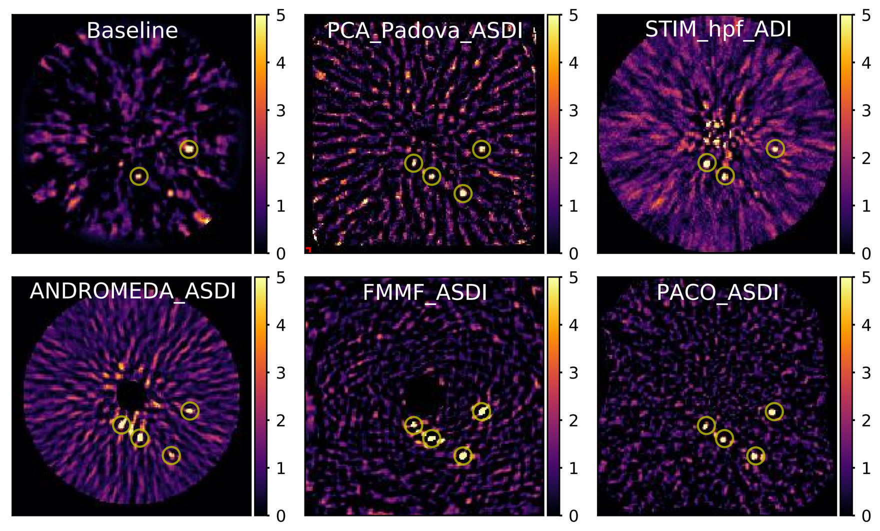
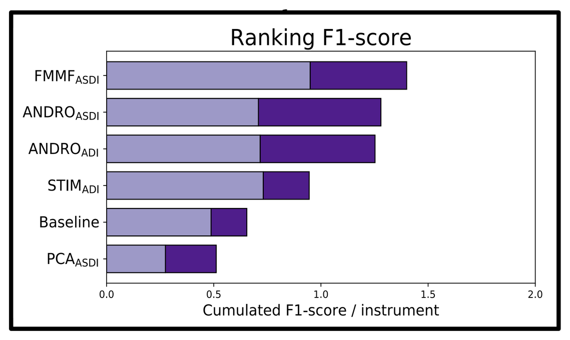


FIG.1: Detection maps (from 0 to threshold). FP are indicated with red squares, TP with yellow circles.

Subchallenge 2: The detection maps below are for the Gemini-S/GPI dataset containing 4 injections.





THE FUTURE OF THE EIDC

The data will be permanently hosted by *Zenodo.*
For the next phases we intend to:
(1) Include the characterization of companions
(2) Add the detection of extended sources
(3) Add hyperspectral data
(4) ⋯ more to come !

And have a report of the different phases of the EIDC published for the SPIE Astronomical Telescopes+Instrumentation conference

Affiliations: [1]Max-Planck-Institute for Astronomy, Heidelberg, Germany, [2]Barcelona Supercomputing Center, Spain, [3]Université Grenoble Alpes, CNRS, IPAG, Grenoble, France,
[4]Space Sciences, Technologies & Astrophysics Research (STAR) Institute, Université de Liège, Allée du six Août 19c, B-4000 Liège, Belgium, [5]Montefiore Institute,Université de Liège, 4000 Liège, Belgium.

Acknowledgements: We thank the collaborators who provided us with pre-reduced data from the various instruments constituting the data challenge (J. Stone, T. Meshkat, D. Mouillet, J.-B. Ruffio).
We thank all the participants who submitted their results, as well as the numerous collaborators who helped designing the data challenge, from the data set to the metric used for comparing the various submissions.