24th July 2023

ASTRONOMICAL RADIO SOURCE DETECTORS

Patrick ARLT
Computer Science Department

Supervisors:
Prof. Svyatoslav Voloshynovskyy
Dr. Olga Taran

UNIVERSITÉ DE GENÈVE
FACULTÉ DES SCIENCES

CONTENT

INTRODUCTION

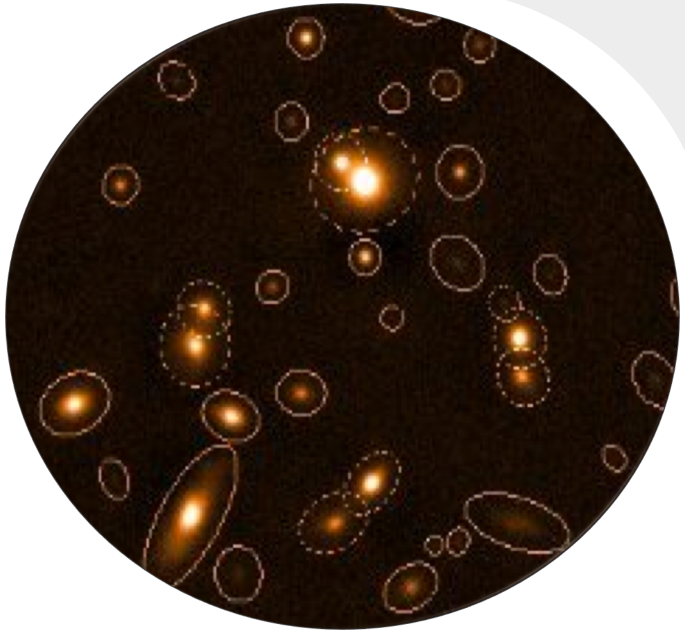
THESIS OBJECTIVES

1. Investigate the state-of-the-art PyBDSF source detector
2. Exploration of source detector based on morphological processing
3. Comparison of PyBDSF and morphological approaches and deep neural network algorithm

CONTEXT & SOURCE DETECTION APPROACHES

TRADITIONAL APPROACHES

- Morphological detector
- PyBDSF
  - Karabo
  - Original PyBDSF



MACHINE LEARNING BASED APPROACH

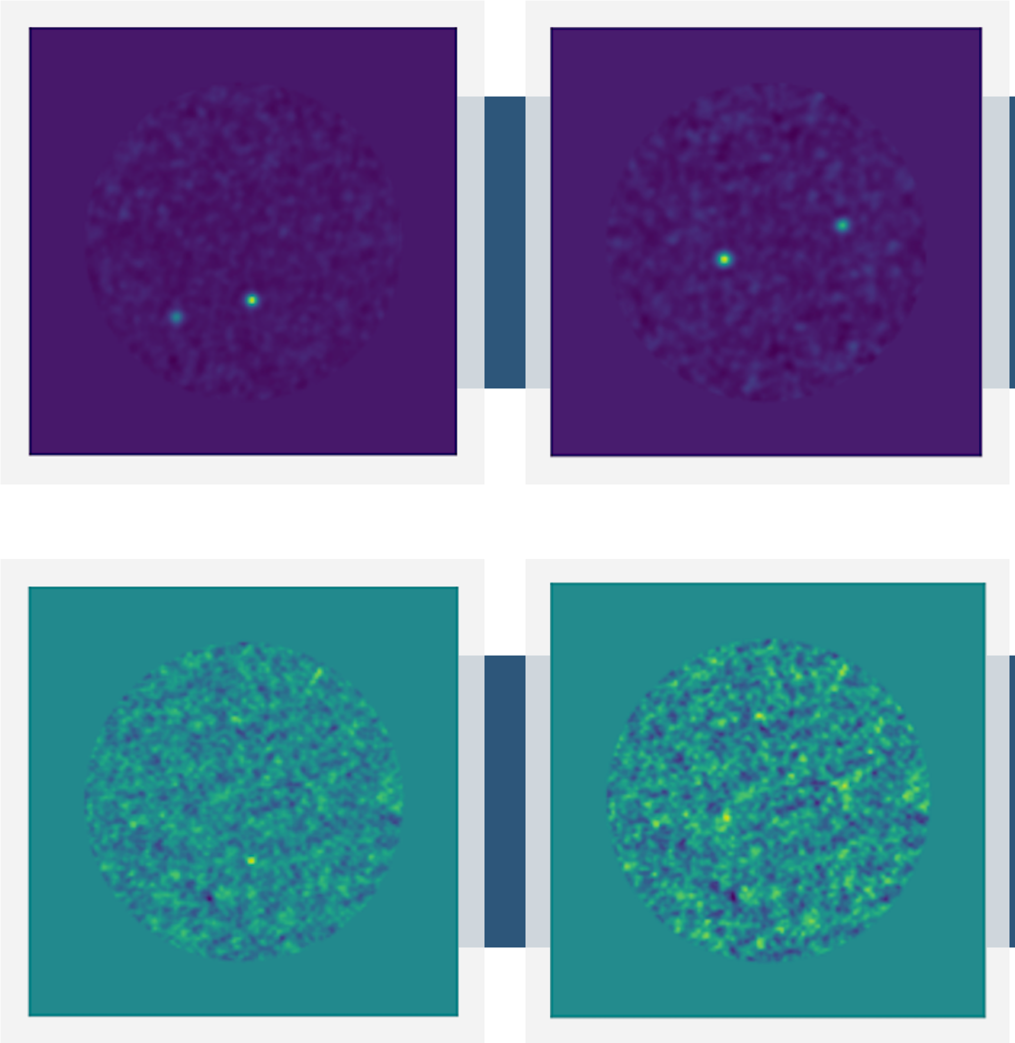
- Convolutional auto encoder

DATASET

- Number of sources: 0 – 5
- Source position: random
- Source Type: Gaussian
- Sky models size: 512x512
- Pwv: 1’796

- 2 subsets:
  - Noiseless
  - Noisy

- Training set: 7’331
- Validation Set: 917
- Testing Set: 916



O. Taran, O. Bait, M. Dessauges-Zavadsky, T. Holotyak, D. Schaerer, and S. Voloshynovskiy, "Challenging interferometric imaging: Machine learning-based source localization from uv-plane observations," *Astronomy & Astrophysics (A&A)*, vol. 674, 2023.

METRICS

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

$$Completeness = \frac{TP}{TP + FN}$$

TP: True Positives

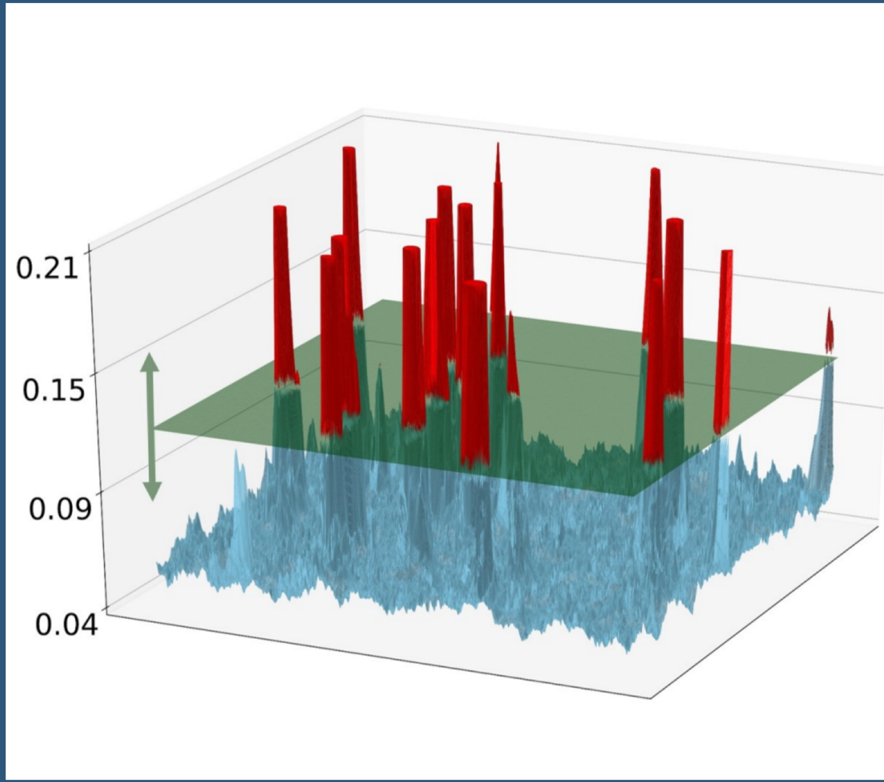
FP: False Positives

FN: False Negatives

EXPERIMENTS

MORPHOLOGICAL DETECTOR

Thresholded Blob Detection (TBD) technique



A. V. Sadr, E. E. Vos, B. A. Bassett, Z. Hosenie, N. Oozeer, and M. Lochner, “Deepsource: Point source detection using deep learning,” MNRAS, vol. 484, no. 2, pp. 2793–2806, April 2019

RESULTS

	Purity	Completeness
Noiseless	0.929	0.879
Noisy	1	0.013

Table 1: Result of the morphological detection approach with the grid search done on the noiseless subset: $area\_lim = 105$, $jump\_lim = 2$, $ignore\_border = 10$, $loc\_det = peak$, and $threshold = 1$

	Purity	Completeness
Noiseless	0.932	0.35
Noisy	0.31	0.926

Table 2: Result of the morphological detection approach with the grid search done on the noisy subset: $area\_lim = 50$, $jump\_lim = 3$, $ignore\_border = 150$, $loc\_det = mean$, and $threshold = 1$

Python Blob Detection and Source Finder

- Extract sources from radio astronomical data
  - Make available their properties

PyBDSF

- Plenty of parameters
- Demand expert domain knowledge
- More complicated to use

N. Mohan and D. Rafferty, “Pybdsf: Python blob detection and source finder,” Astrophysics Source Code Library, pp. ascl–1502, 2015

Karabo

- Only few parameters
- Intuitive
- Easy-to-use

https://github.com/i4Ds/Karabo-Pipeline/tree/main/karabo

RESULTS KARABO

	Purity	Completeness
Noiseless	0.3433	0.9806
Noisy	0.9772	0.1661

Table 3: Karabo source detection efficiency.

RESULTS PYBDSF

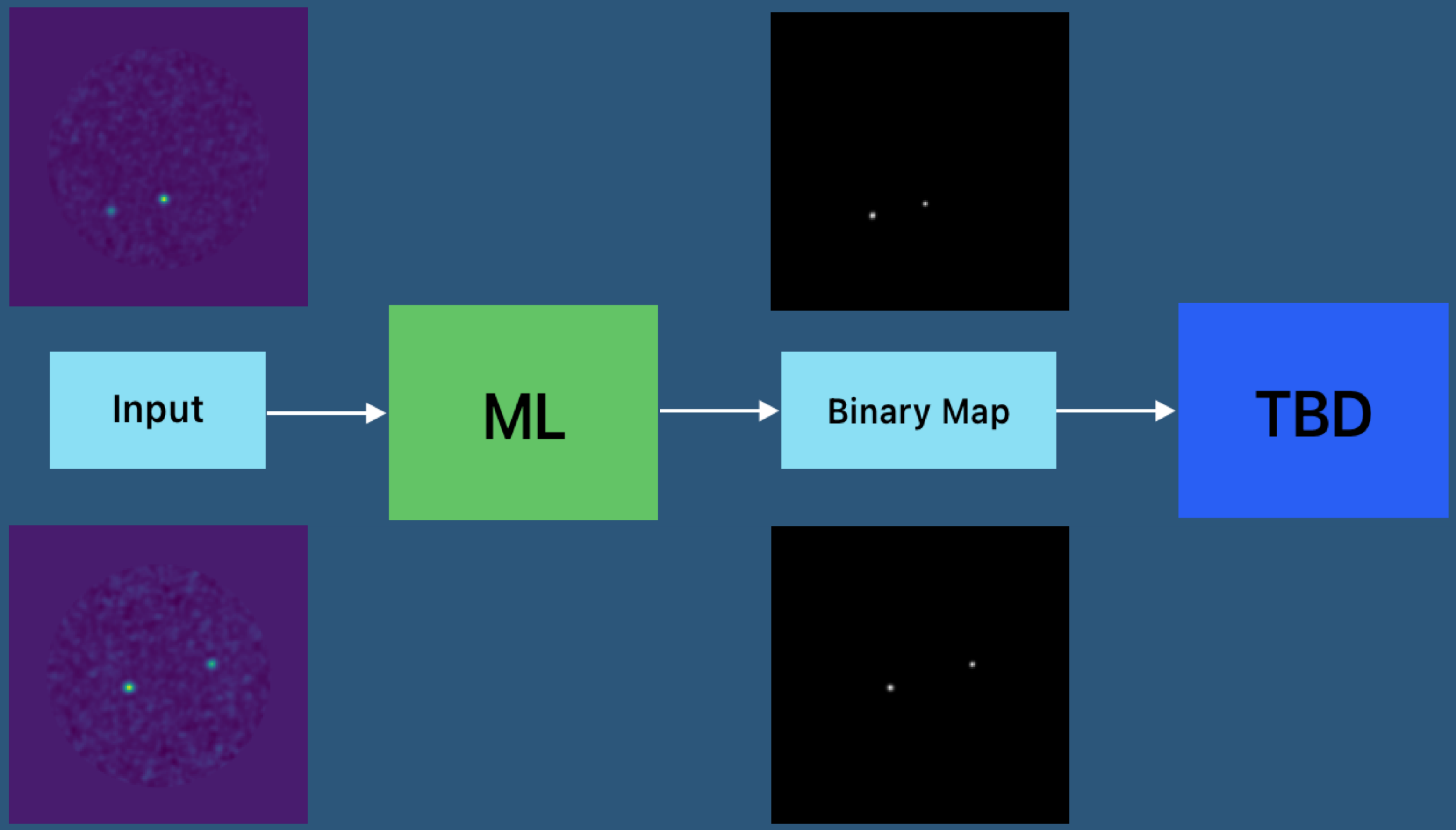
Dataset	Purity	Completeness
Noiseless	0.921	0.93
Noisy	1	0.023

Table 4: PyBDSF efficiency with *rms_map* parameter set to *True*

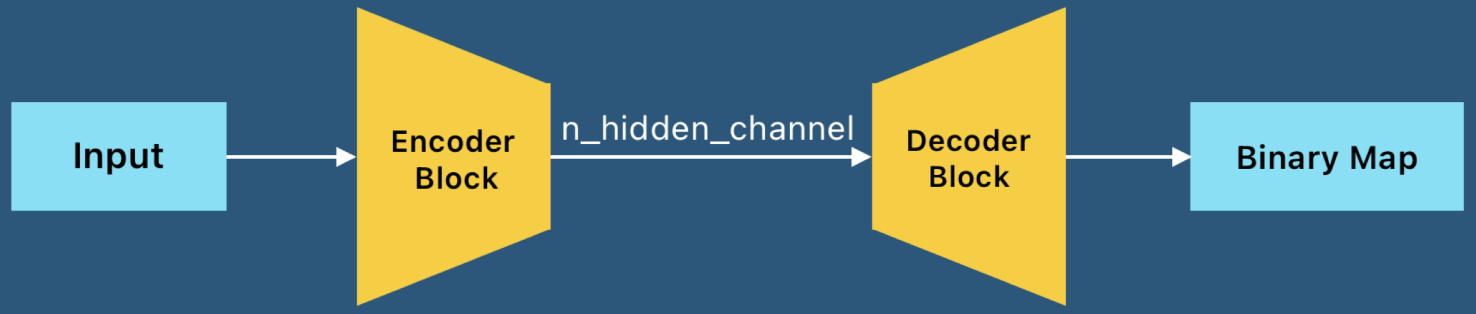
dataset	Purity	Completeness
Noiseless	0.428	0.962
Noisy	0.984	0.178

Table 5: PyBDSF efficiency with parameters *rms_map* set to *False* and *rms_value* set to *None*

MACHINE LEARNING APPROACH



AUTO ENCODER



RESULTS

	Purity	Completeness
noiseless	0.894	0.846
noisy	0	0

Table 6: Efficiency of the source detection process when using the CNN model trained on noiseless data

	dataset	Purity	Completeness
jump_lim = 1	noiseless	0.562	0.392
	noisy	0.764	0.436
jump_lim = 2	noiseless	0.32	0.64
	noisy	0.537	0.658

Table 7: Efficiency of the source detection process when using the CNN model trained on noisy data

CONCLUSION

SUMMARY

Data type	Approaches	Purity	Completeness
Noiseless	TBD	0.929	0.879
	PyBDSF	0.921	0.93
	CNN autoencoder	0.894	0.846
Noisy	TBD	0.31	0.926
	PyBDSF	0.984	0.178
	CNN autoencoder	0.764	0.436

Table 8: Comparision of all approaches explored

FUTURE WORK

- Train on both noiseless and noisy images
- Extend the dataset (more complicated sources)
- Explore more advanced architecture i.e. U-net, Transformers, DDPM, etc

Thank you for your attention

Questions ?