Finanziato dall'Unione europea NextGenerationEU
Ministero dell'Università e della Ricerca
Italiadomani PIANO NAZIONALE DI RIPRESA E RESILIENZA
ICSC Centro Nazionale di Ricerca in HPC, Big Data and Quantum Computing

Updates on Radio U-Net: application to the full LoTSS survey

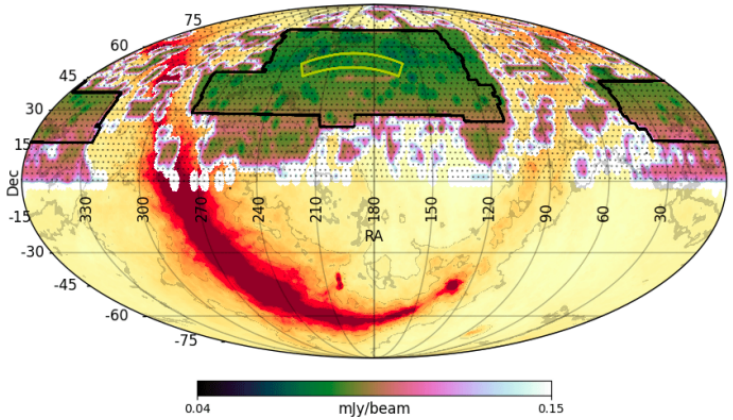
Chiara Stuardi
INAF-IRA
with
Claudio Gheller (INAF-IRA),N. Sanvitale (INAF-IRA),G. Di Gennaro (INAF-IRA),
F. Braga (UniBo) F. Vazza (UniBo), A. Botteon (INAF-IRA)

Spoke 3 II Technical Workshop, Perugia May 26-29, 2025

Scientific Rationale

Current radio surveys are challenging our detection and cataloging strategies

- large data size (PB/year)

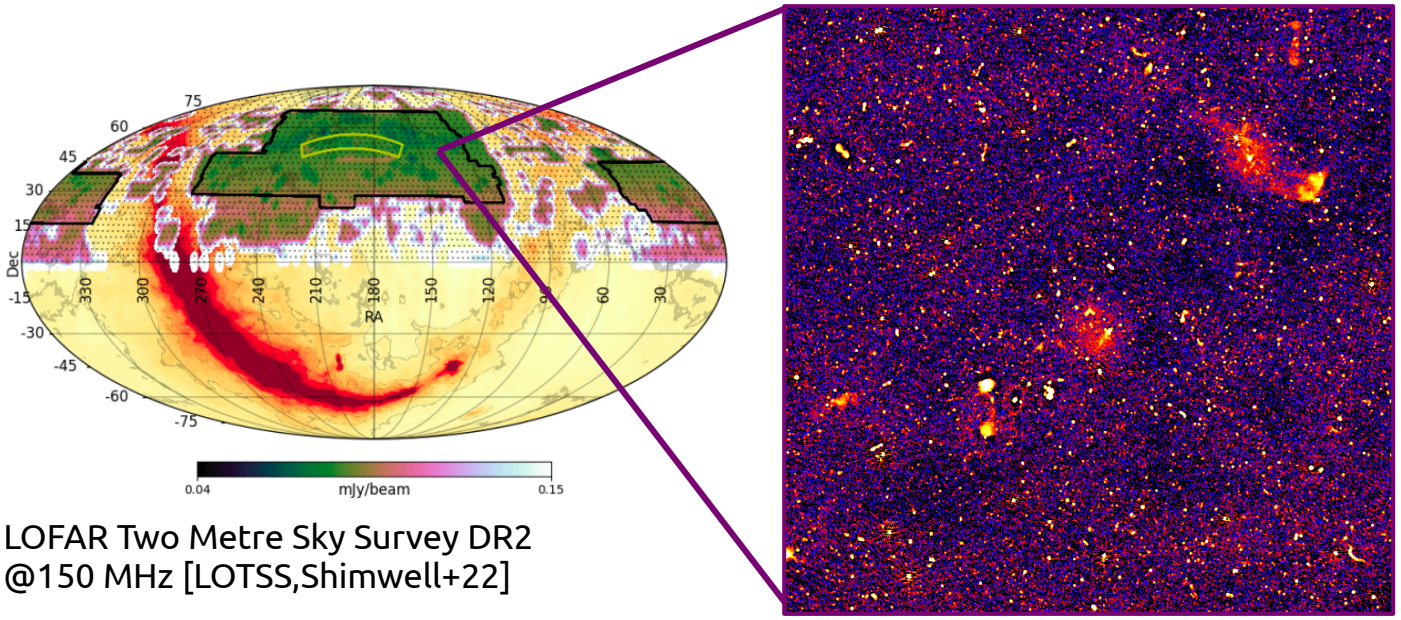


LOFAR Two Metre Sky Survey DR2
@150 MHz [LOTSS,Shimwell+22]

Scientific Rationale

Current radio surveys are challenging our detection and cataloging strategies

- large data size (PB/year)
- time-consuming and computationally expensive data reduction and imaging procedures
- non-Gaussian noise and imaging artifacts

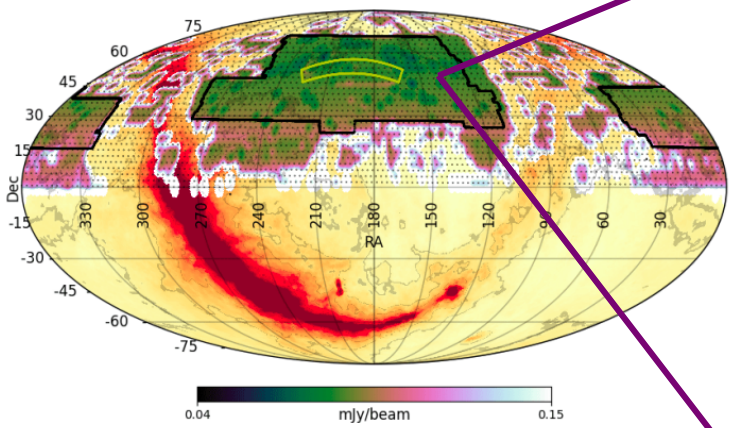


LOFAR Two Metre Sky Survey DR2
@150 MHz [LOTSS,Shimwell+22]

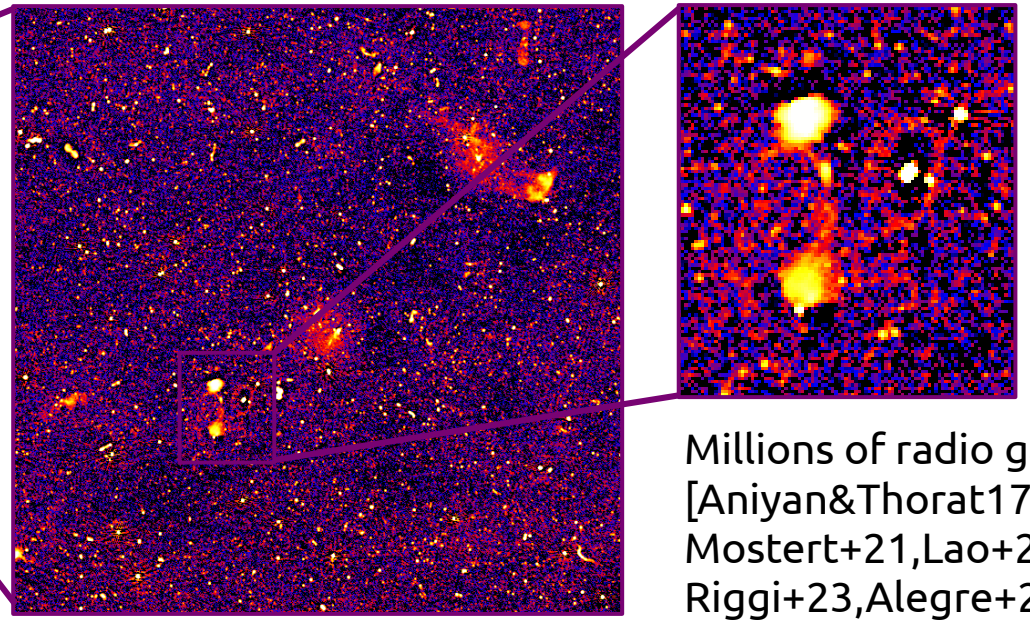
Scientific Rationale

Current radio surveys are challenging our detection and cataloging strategies

- large data size (PB/year)
- time-consuming and computationally expensive data reduction and imaging procedures
- non-Gaussian noise and imaging artifacts



LOFAR Two Metre Sky Survey DR2 @150 MHz [LOTSS,Shimwell+22]

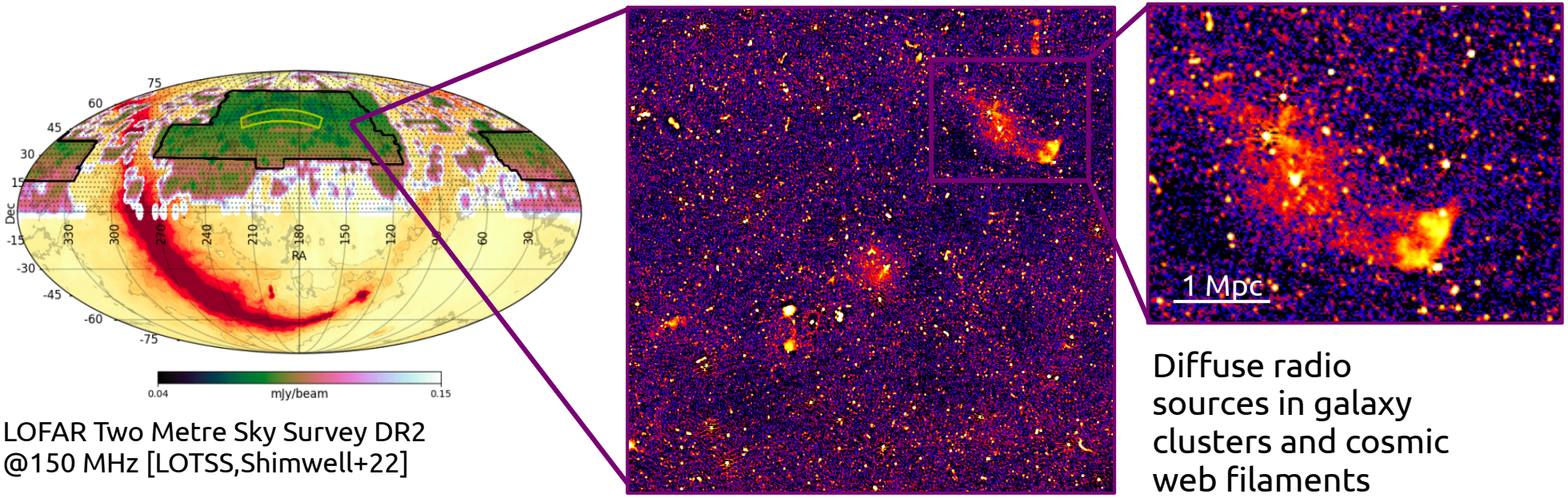


Millions of radio galaxies [Aniyan&Thorat17,Lukic+18, Mostert+21,Lao+23, Riggi+23,Alegre+24,Gupta+24 Riggi+24,Slijepcevic+24]

Scientific Rationale

Current radio surveys are challenging our detection and cataloging strategies

- large data size (PB/year)
- time-consuming and computationally expensive data reduction and imaging procedures
- non-Gaussian noise and imaging artifacts
- rare sources with complex and irregular morphology



LOFAR Two Metre Sky Survey DR2 @150 MHz [LOTSS,Shimwell+22]

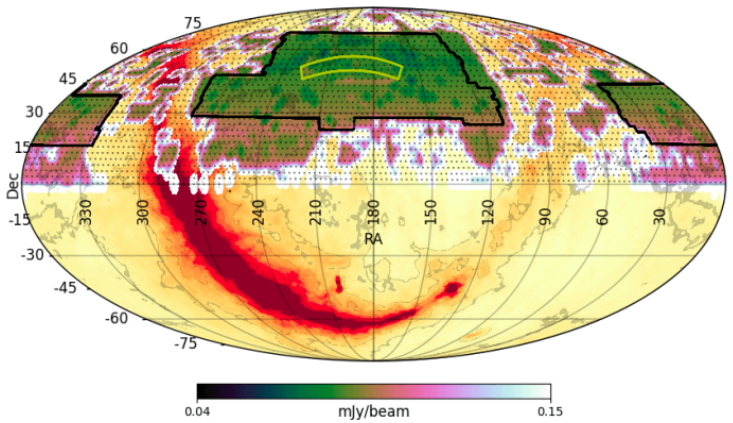
Diffuse radio sources in galaxy clusters and cosmic web filaments

Scientific Rationale

Current radio surveys are challenging our detection and cataloging strategies

- large data size (PB/year)
- time-consuming and computationally expensive data reduction and imaging procedures
- non-Gaussian noise and imaging artifacts
- rare sources with complex and irregular morphology

  → new strategies to minimize human intervention in data processing



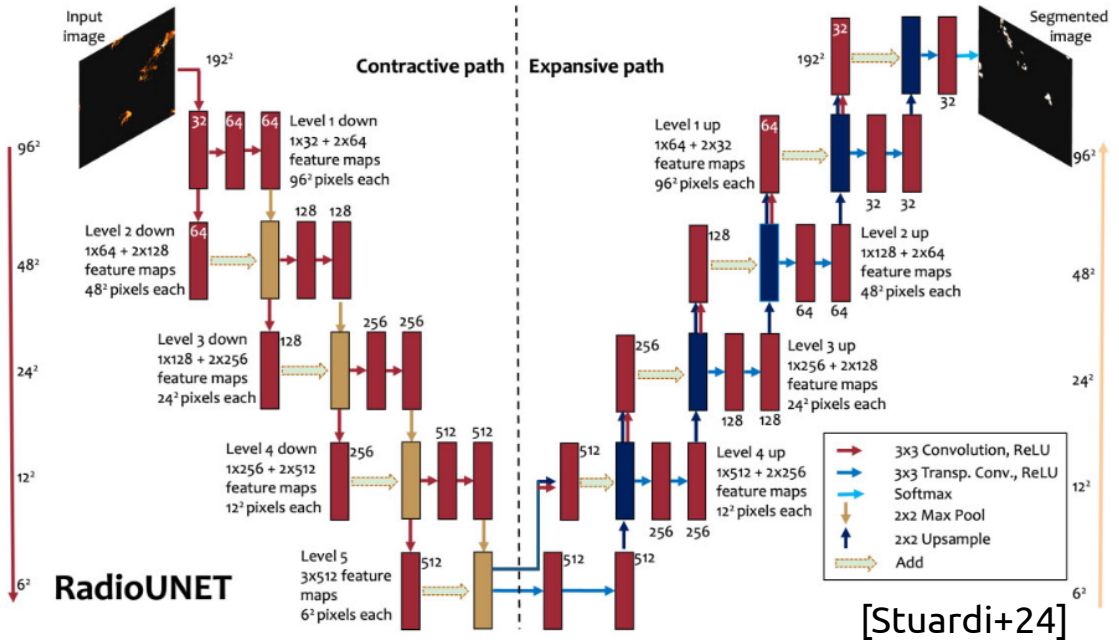
LOFAR Two Metre Sky Survey DR2
@150 MHz [LOTSS,Shimwell+22]

Square Kilometre Array, operational in 2030

Technical Objectives, Methodologies and Solutions

Objective: Developement and distribution of a machine leraning tool for detection of diffuse emission in radio surveys

- **U-Net architecture adopted perform fast segmentation of large datasets**

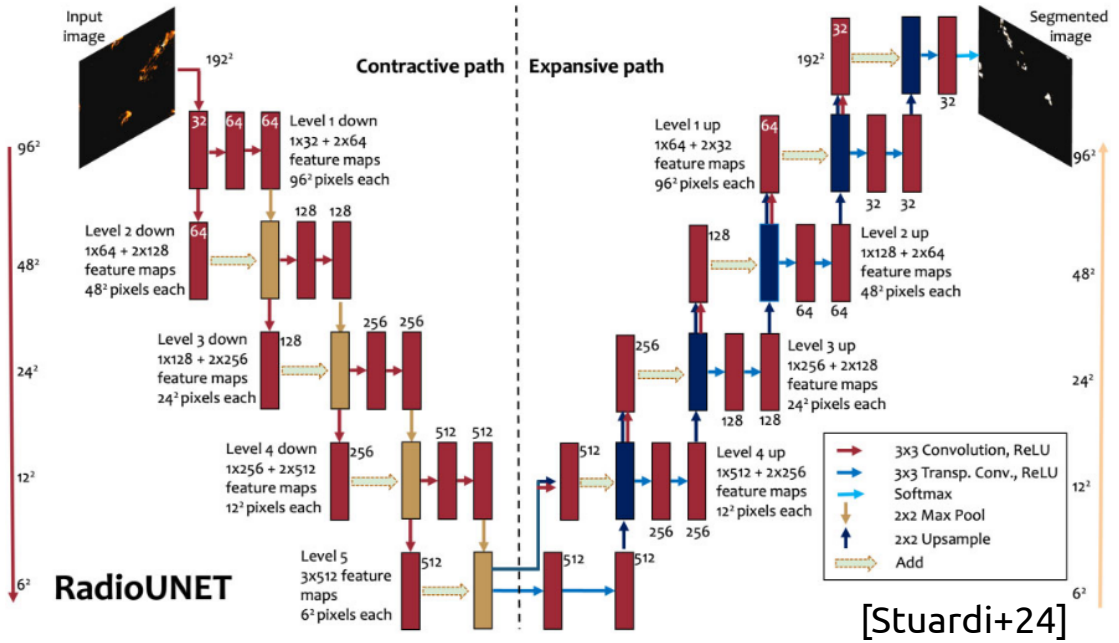


[Stuardi+24]

Technical Objectives, Methodologies and Solutions

Objective: Developement and distribution of a machine leraning tool for detection of diffuse emission in radio surveys

- **U-Net architecture adopted perform fast segmentation of large datasets**
- **Scalbility from CPU to GPU, run on Leonardo @CINECA**

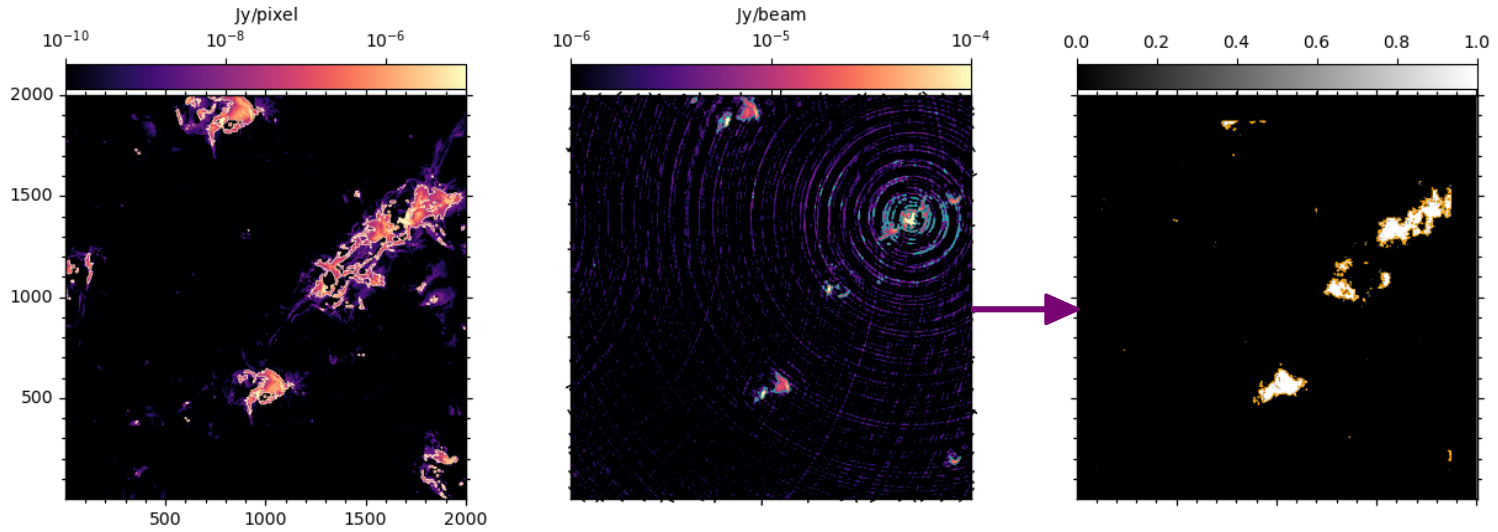


[Stuardi+24]

Technical Objectives, Methodologies and Solutions

Objective: Developement and distribution of a machine leraning tool for detection of diffuse emission in radio surveys

- **U-Net architecture adopted perform fast segmentation of large datasets**
- **Scalbility from CPU to GPU, run on Leonardo @CINECA**
- **Supervised learning with trainig on comological simulations** [Gheller&Vazza22]

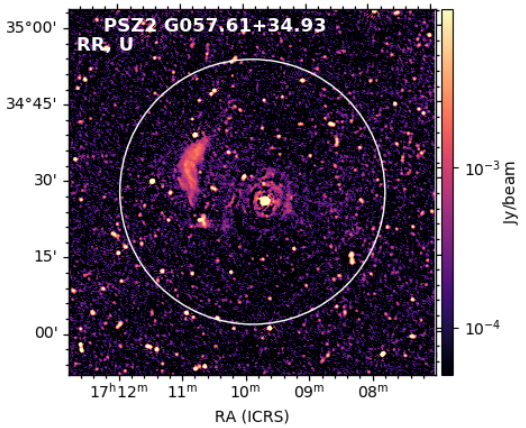


Sky image – Reference mask | Input image | Output segmented image

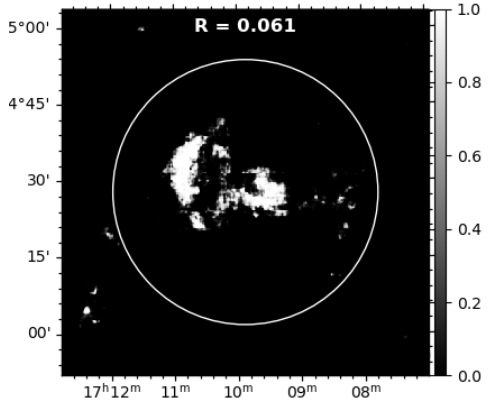
Technical Objectives, Methodologies and Solutions

<u>Objective</u>: Developement and distribution of a machine leraning tool for detection of diffuse emission in radio surveys

- **U-Net architecture adopted perform fast segmentation of large datasets**
- **Scalbility from CPU to GPU, run on Leonardo @CINECA**
- **Supervised learning with trainig on comological simulations**
- **Application and performance verification on LOFAR Two Metre Sky Survey (LoTSS)**



LoTSS image



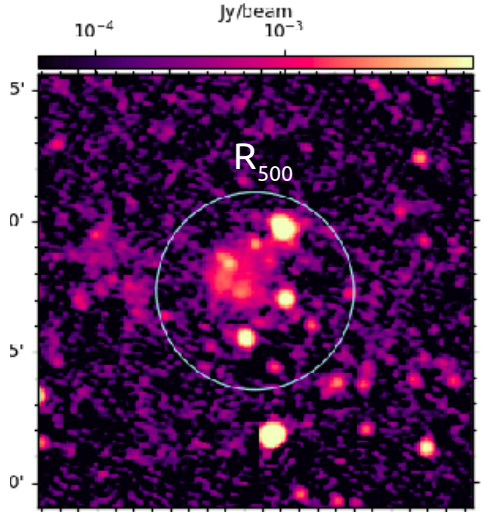
Radio U-Net segmented image

Main Results

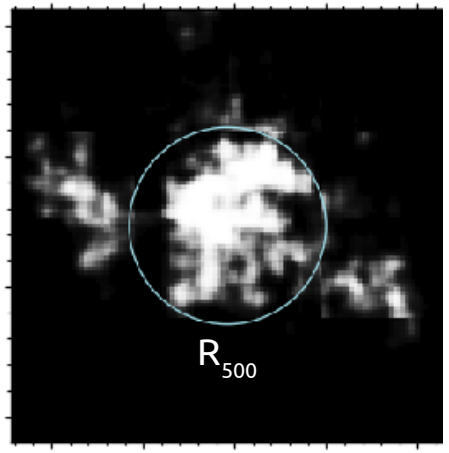
Paper published on MNRAS (KPI) ✓
+ follow-up A&A letter published
Private code on Spoke 3 GitHub 100%

Main Results

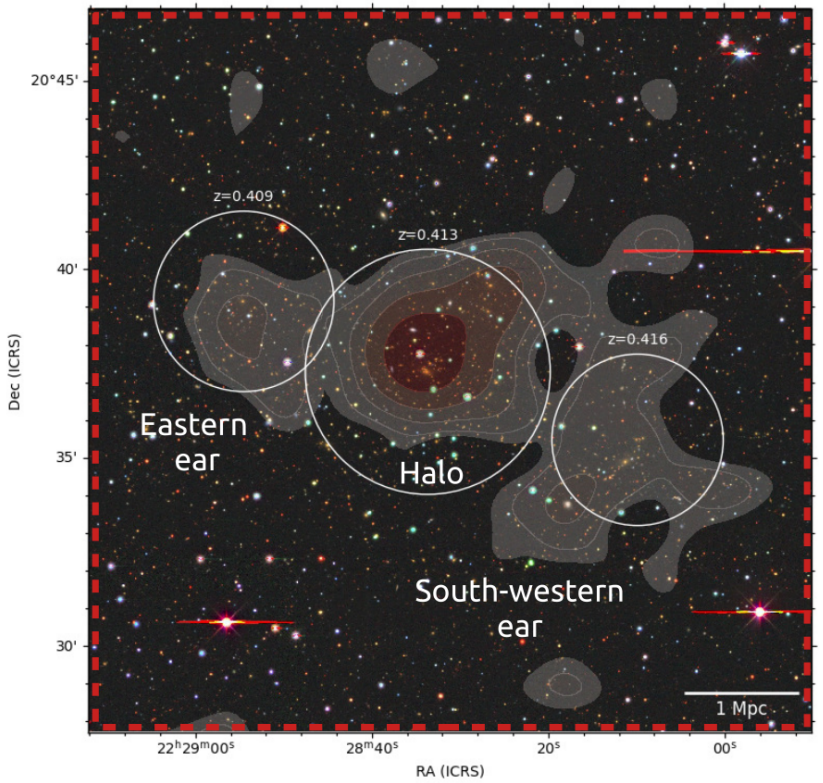
LoTSS data



Radio U-Net segmented

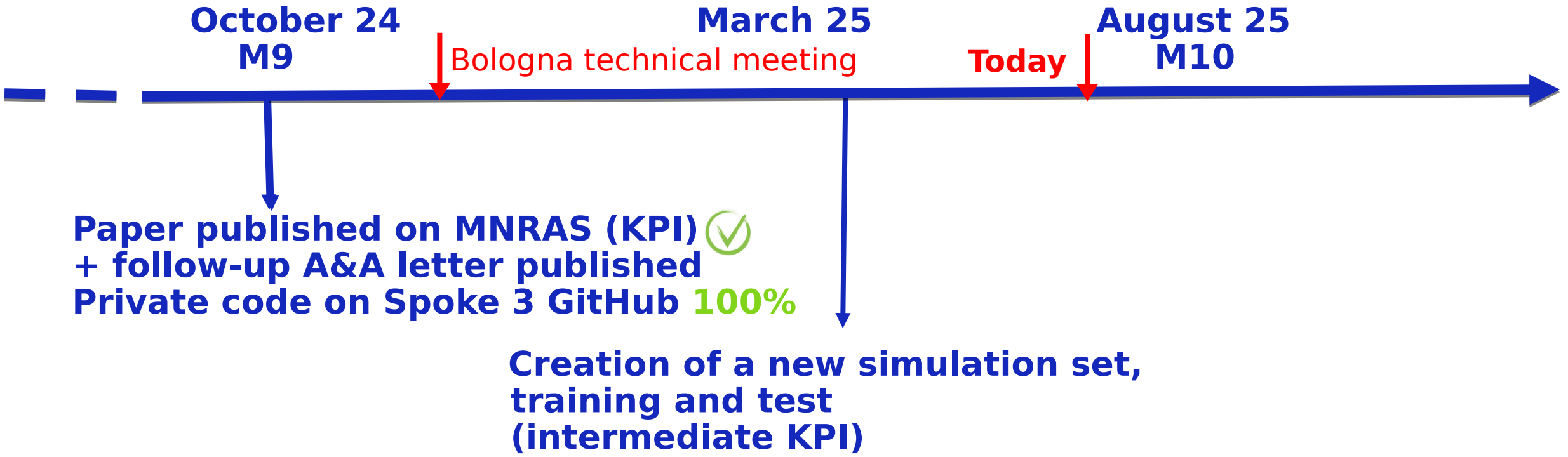


Detection of diffuse radio emission beyond galaxy clusters and below classical detection limits

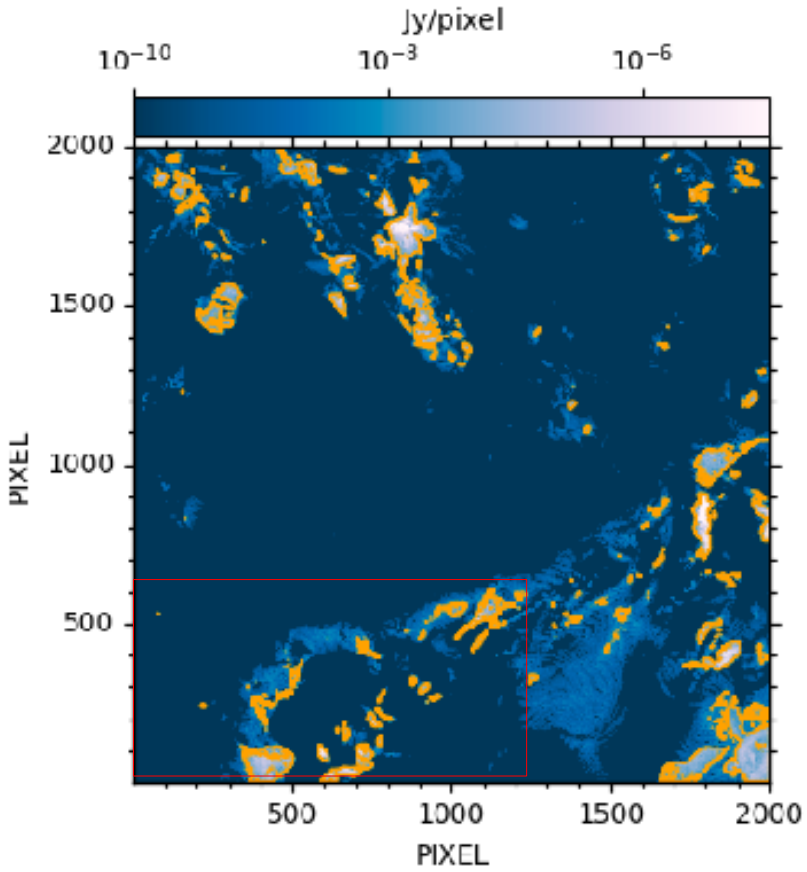


[Stuardi+25]

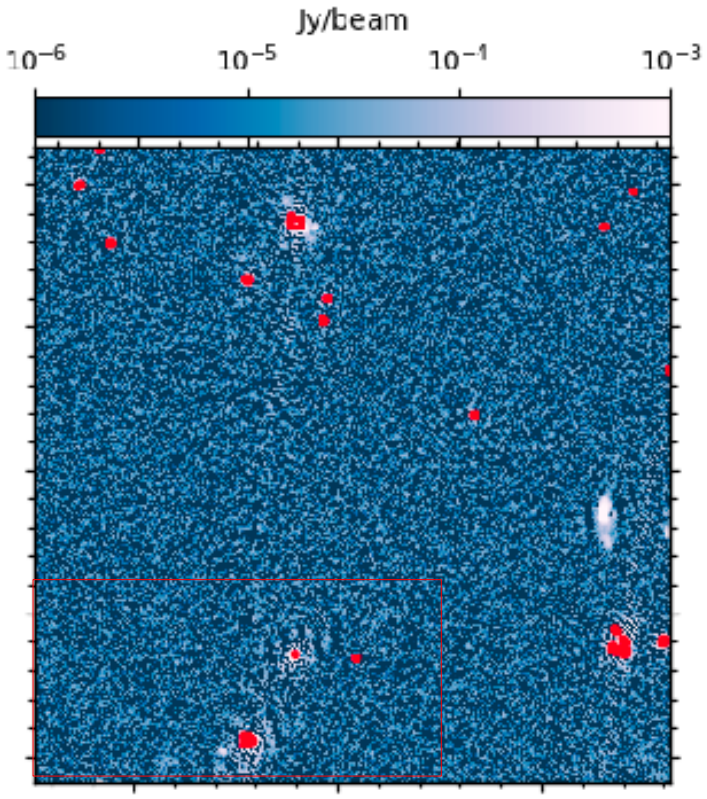
Main Results



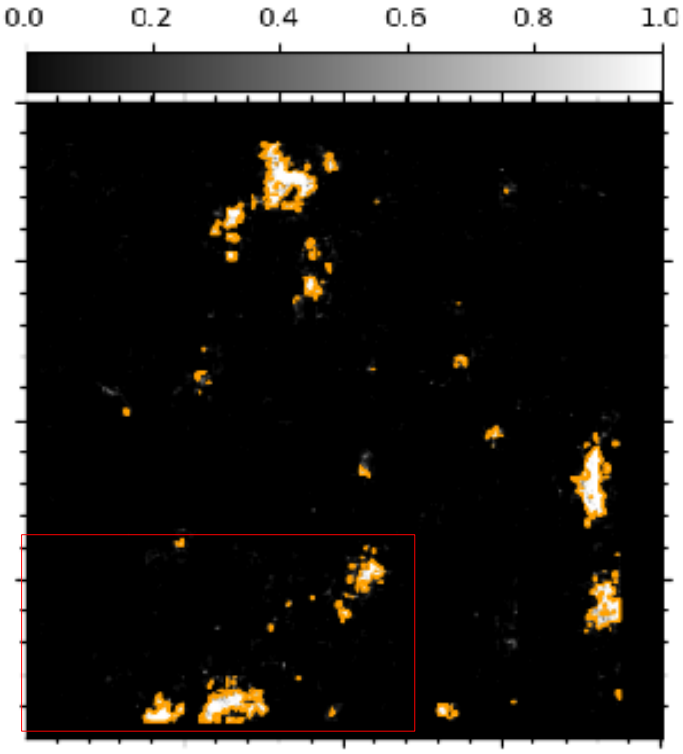
Main Results



Reference mask (no AGN)

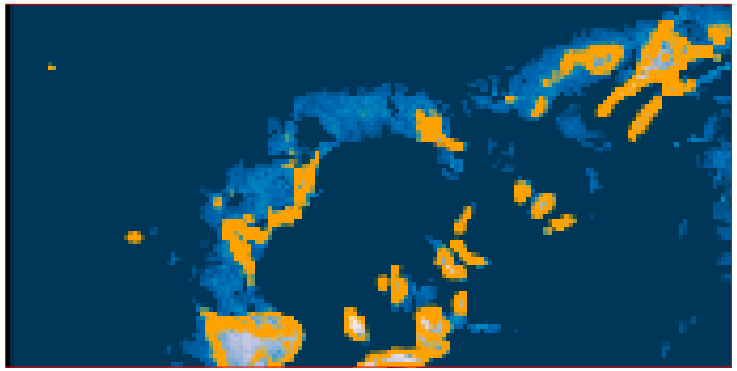


Input image (with AGN)

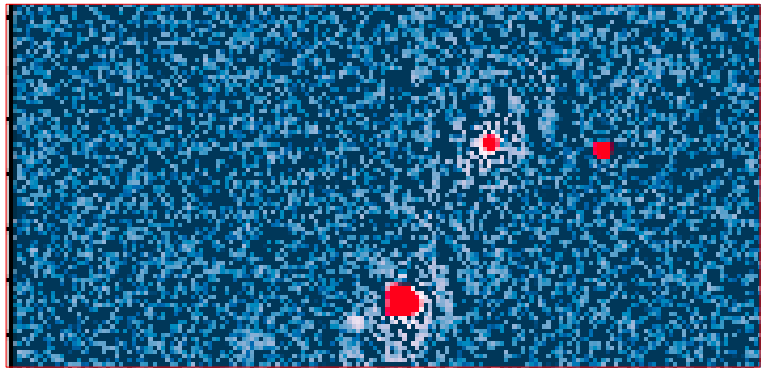


Output segmented image

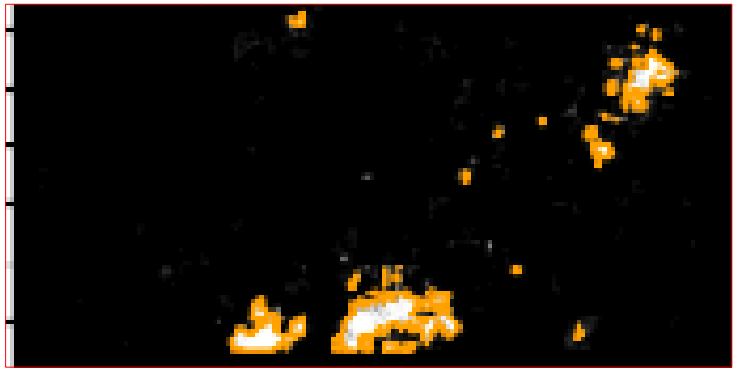
Main Results



Reference mask (no AGN)

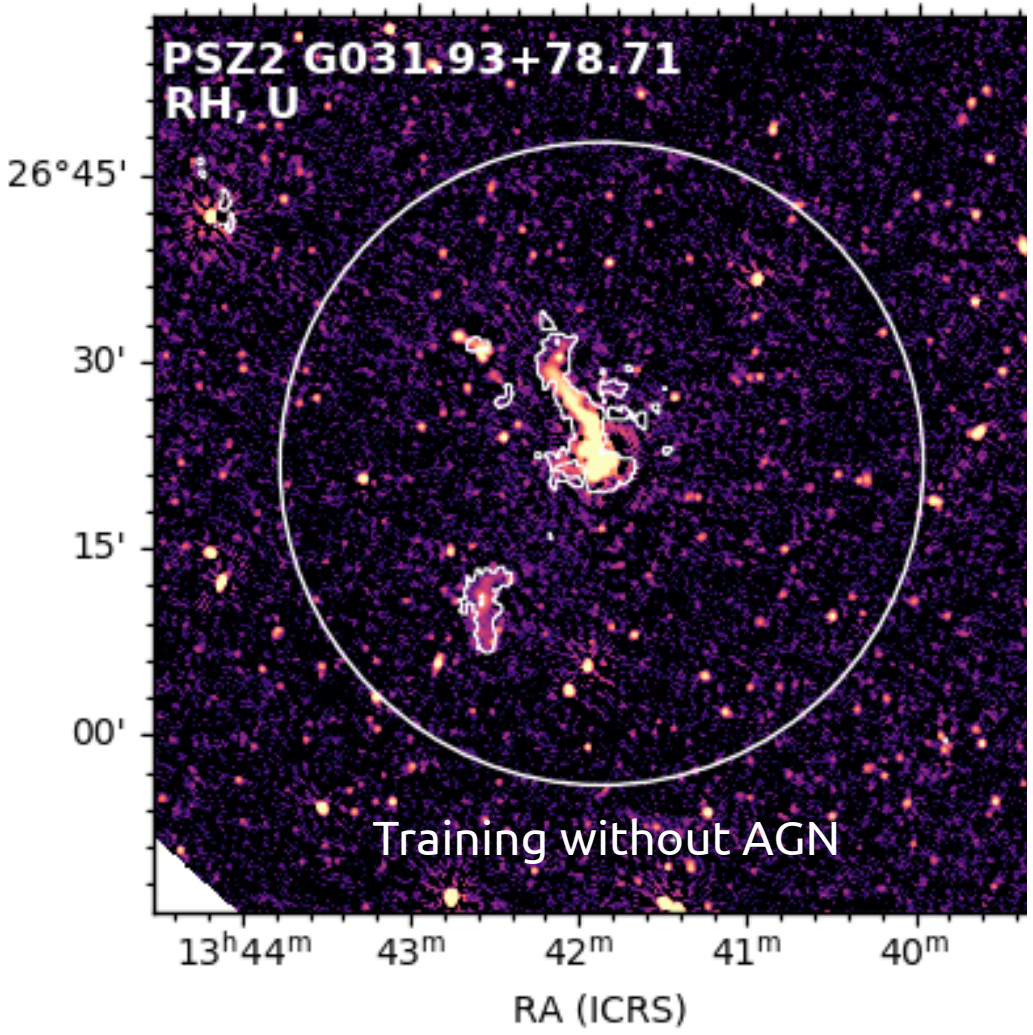


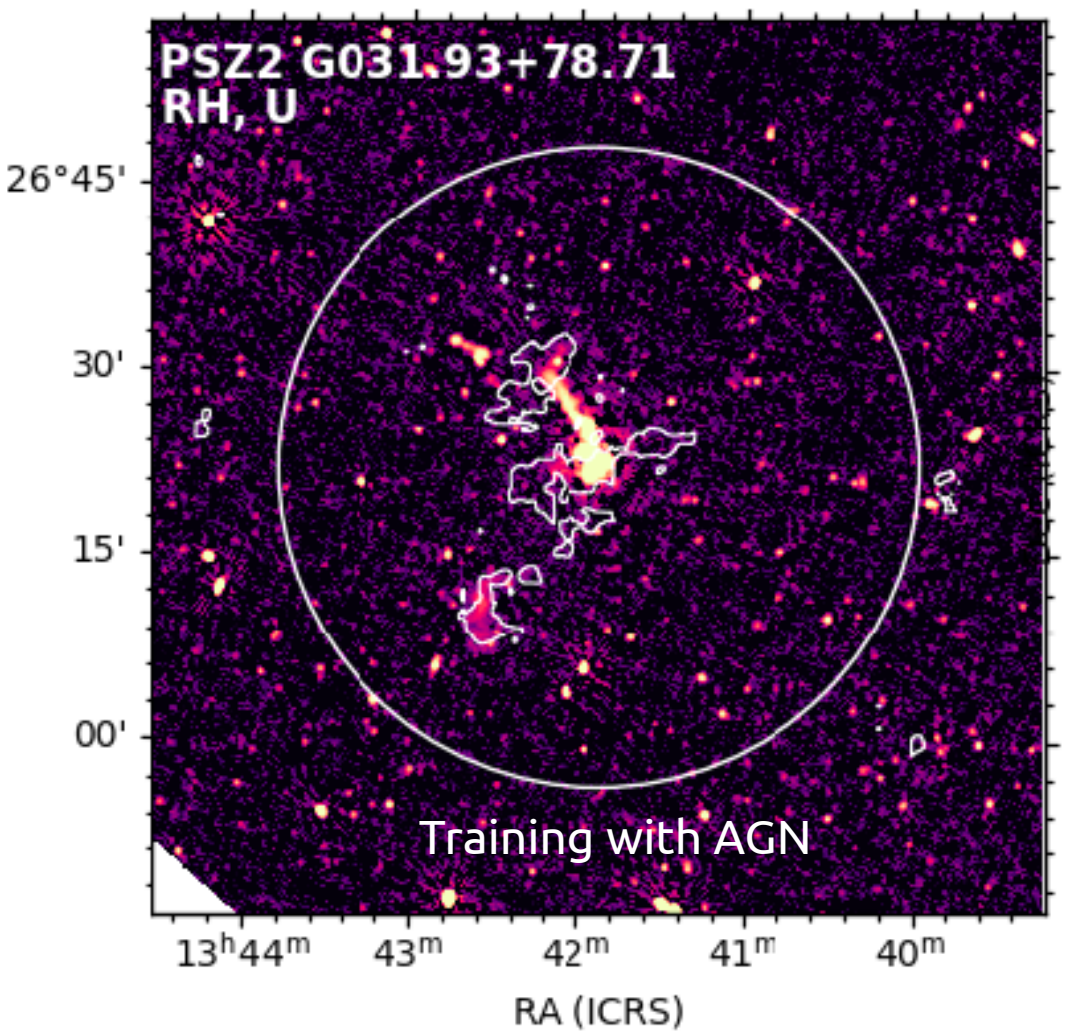
Input image (with AGN)



Output segmented image

Main Results





Final Steps

October 24
M9

March 25

Bologna technical meeting

Today

August 25
M10

Paper published on MNRAS (KPI) ✓
+ follow-up A&A letter published
Private code on Spoke 3 GitHub 100%

Creation of a new simulation set
training and test 100%
(intermediate KPI) ✓

Application of the network
to the LoTSS DR3 and submit paper
(KPI)

Final Steps

October 24
M9

March 25

Bologna technical meeting

Today

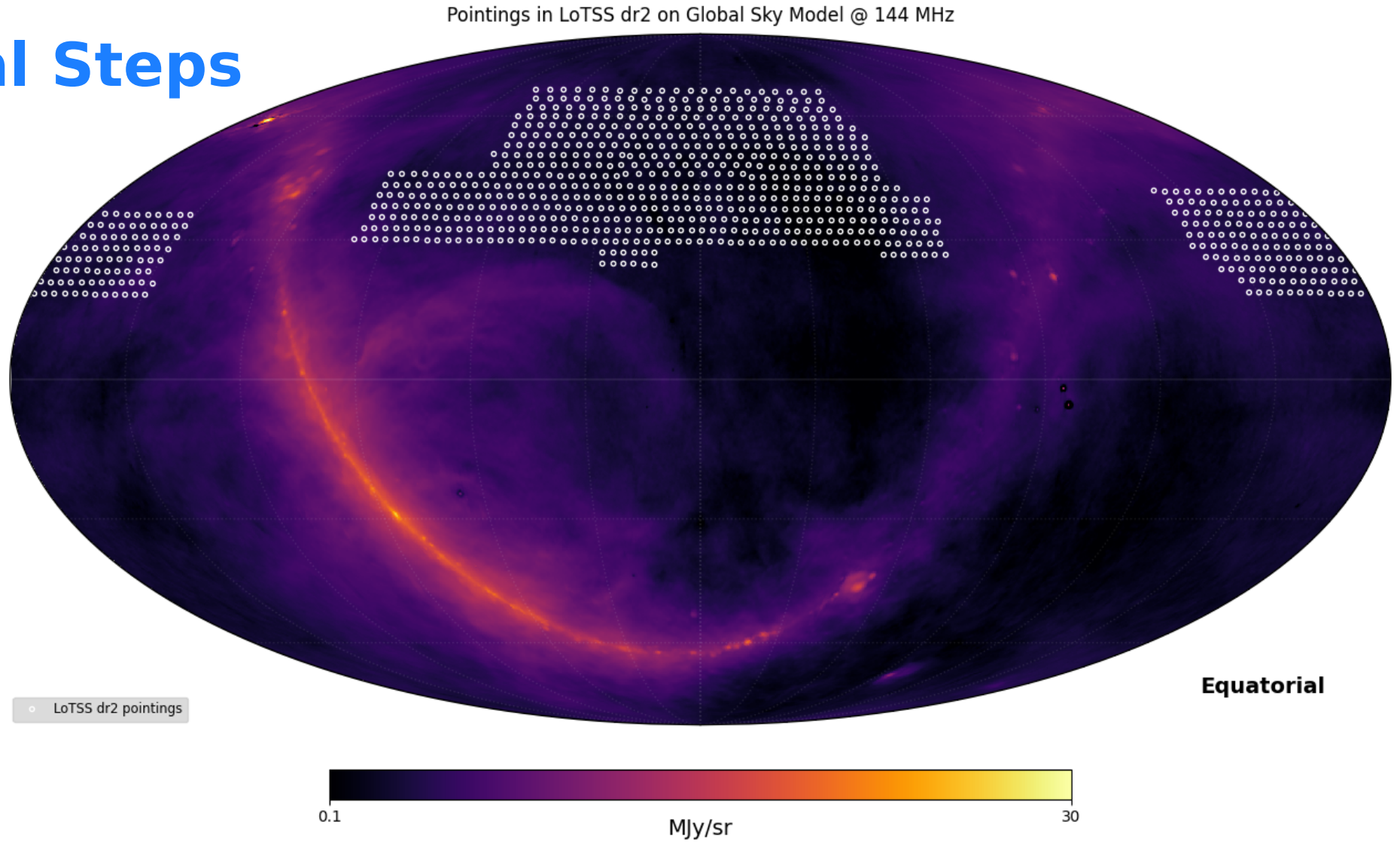
August 25
M10

Paper published on MNRAS (KPI) ✓
+ follow-up A&A letter published
Private code on Spoke 3 GitHub 100%

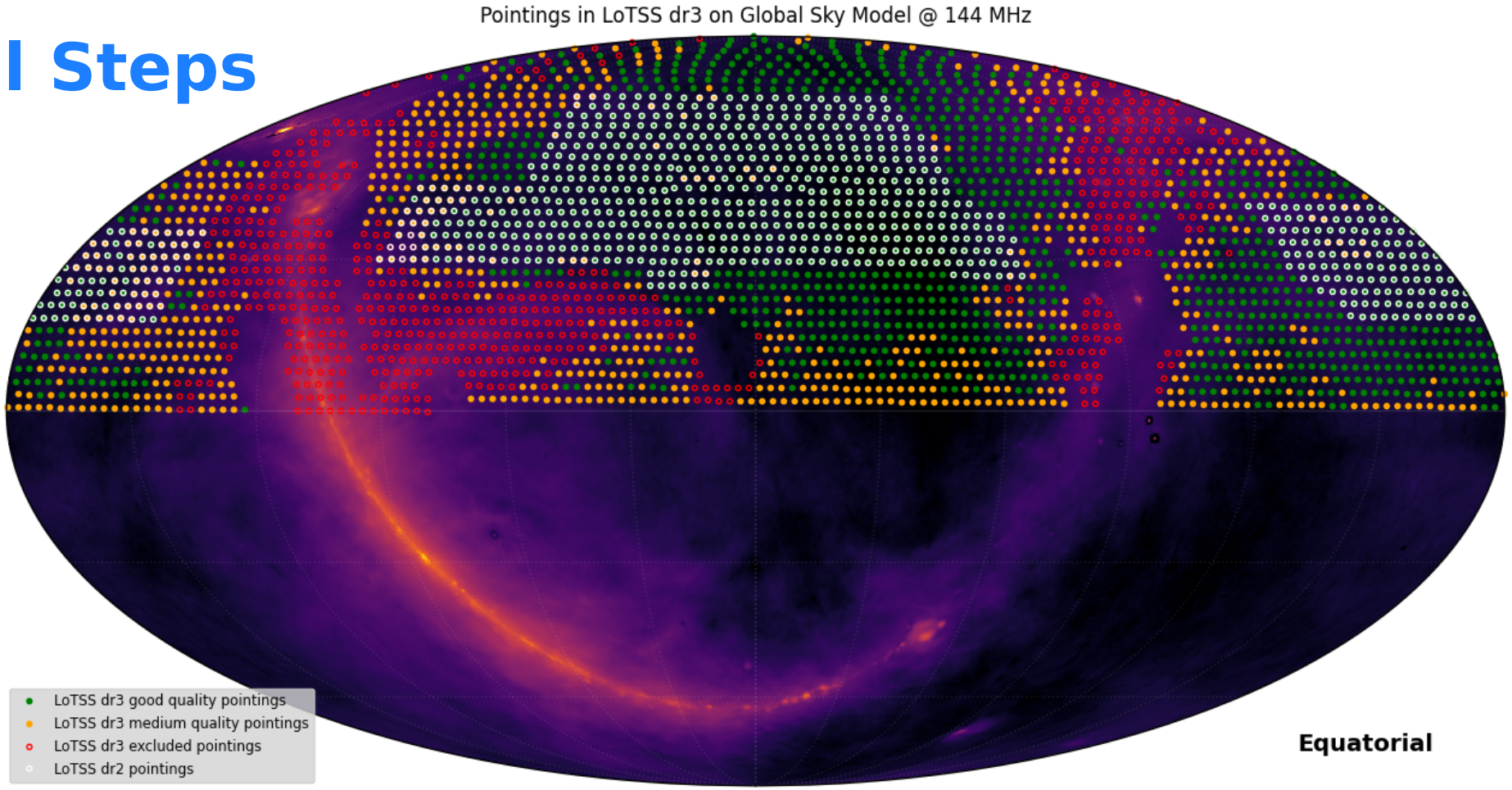
Creation of a new simulation set
training and test 100%
(intermediate KPI) ✓

Application of the network
to the LoTSS DR3 and submit paper
(KPI) 60%

Final Steps

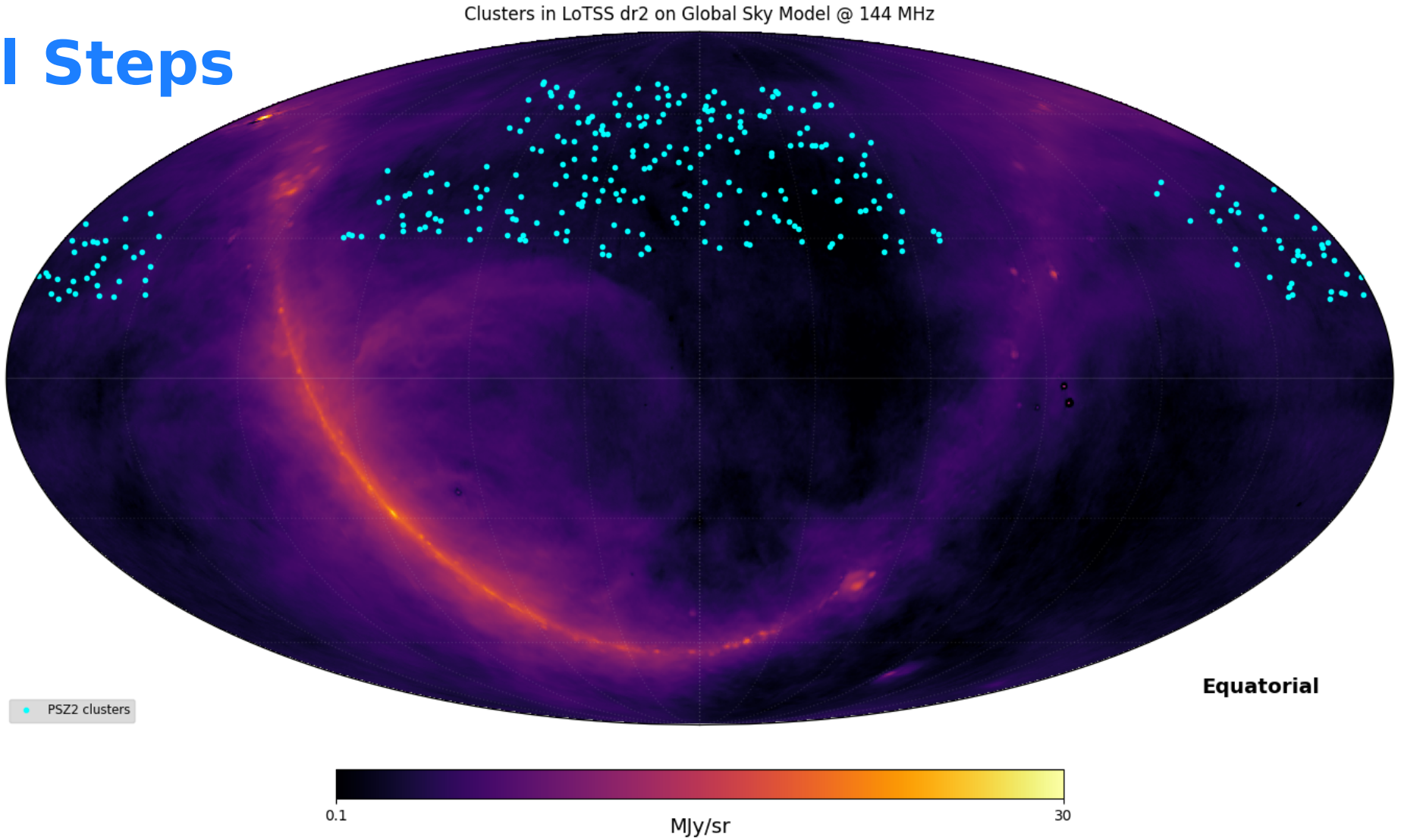

Pointings in LoTSS dr2 on Global Sky Model @ 144 MHz
LoTSS dr2 pointings
Equatorial
0.1
30
MJy/sr

Final Steps

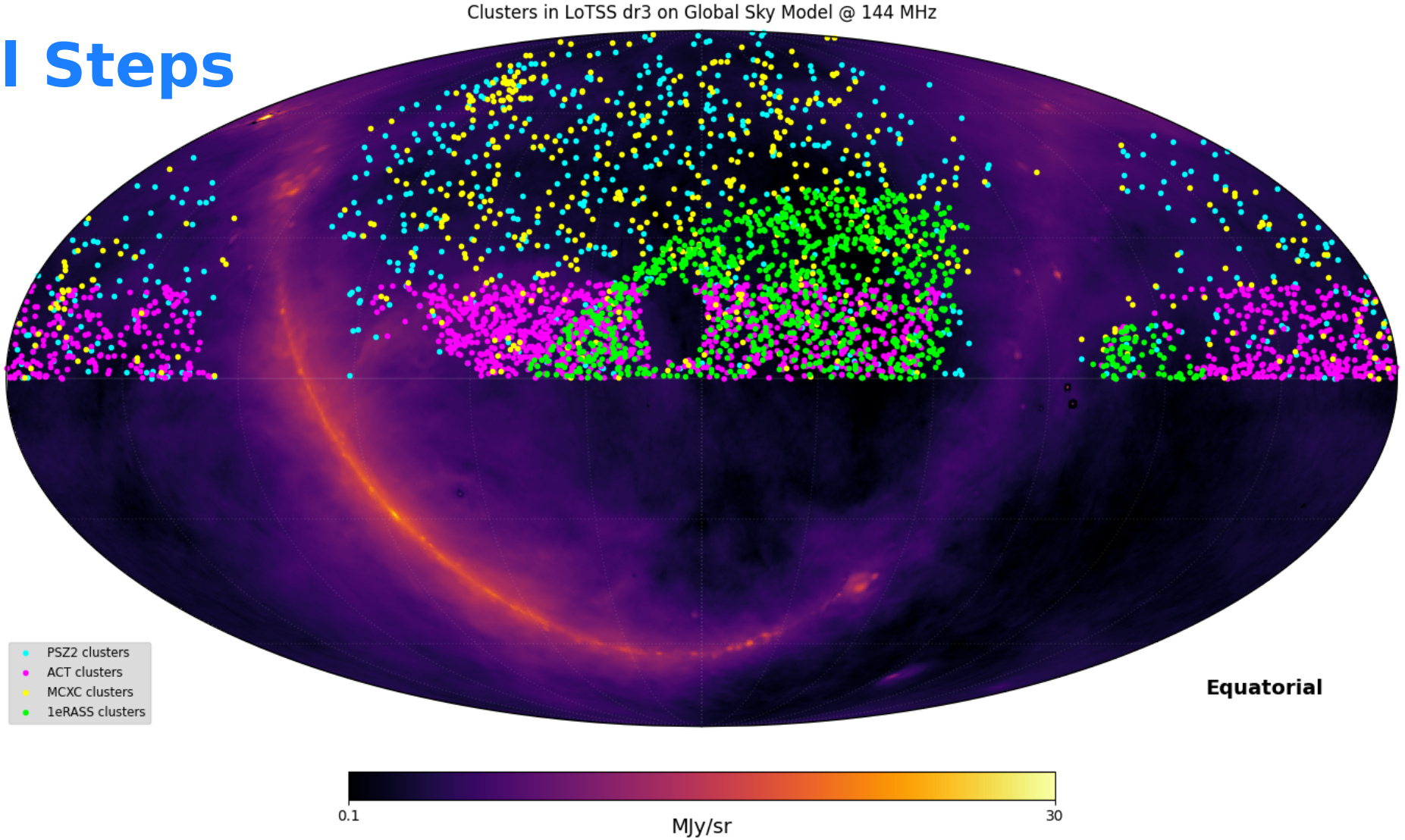


- Adapt the pipeline to process entire pointings
- New problem: Galactic diffuse emission

Final Steps


Clusters in LoTSS dr2 on Global Sky Model @ 144 MHz
PSZ2 clusters
Equatorial
0.1
30
MJy/sr

Final Steps



309 → 3821 clusters

[Stuardi & LoTSS collaboration in prep.]

Final Steps

December 25
M11(?)

Publish the LoTSS dr3 paper, segmentation masks and the code (KPI) 30%

TOTAL: 80%