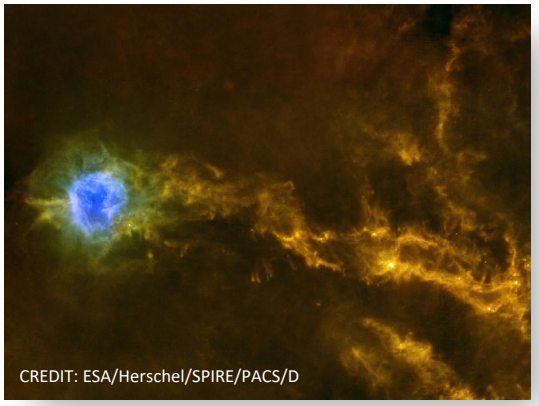
Segmentation of the Galactic ISM Filaments using Deep Learning and Hi-GAL Catalogue

Siouar BENSAID (LAM & LIS, France)

Annie ZAVAGNO (LAM, France)

François-Xavier DUPÉ (LIS & LAM, France)



ESA/ESO SCIOPS WORKSHOP 16-20 May 2022 - Garching

BigSF: Multidisciplinary Project

LAM/LIS (Marseille)

Annie Zavagno (Astro)

Siouar Bensaid (ML)

François-Xavier Dupé (ML)

Morgan Gray (ML)

Jean-Charles Lambert (HPC)

Delphine Russeil (Astro)

Doris Arzoumanian (NAOJ, Astro)

INAF-IAPS (Rome)

Eugenio Schisano (Astro)

Gianluca Li Causi (Astro)

Sergio Molinari (Astro)

Davide Elia (Astro)

INAF-CAO (Naples)

Massimo Brescia (Astro / ML)

Guiseppe Riccio (Astro / ML)

Stefano Cavuoti (Astro / ML)

Sponsorship

Fondation A*Midex

CNRS – MITI

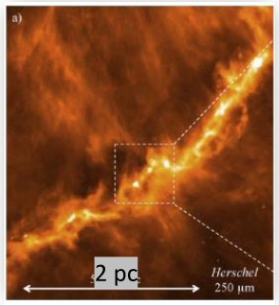
Institut Universitaire de France

BigSF: Bridging Astro & AI for SF Study

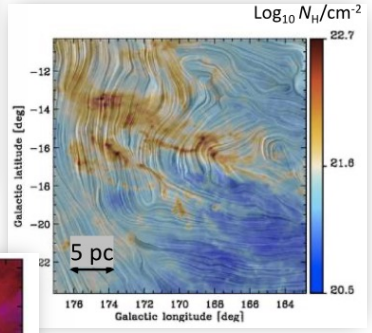
Key Question: How does the environment affect star formation (SF)?

Filaments

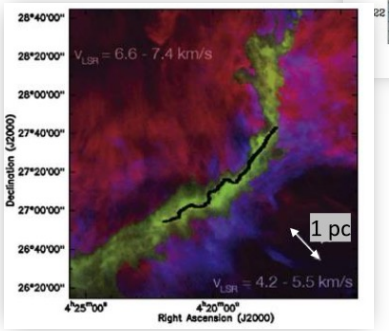
$N(H_2)$



$N(H_2)$ & Magnetic field



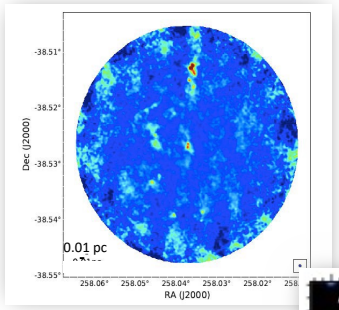
Velocity field



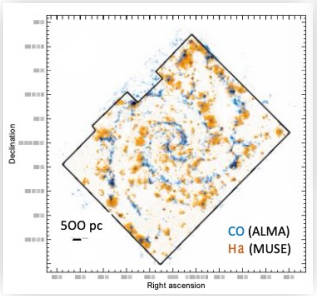
back to the initial form

Galaxies

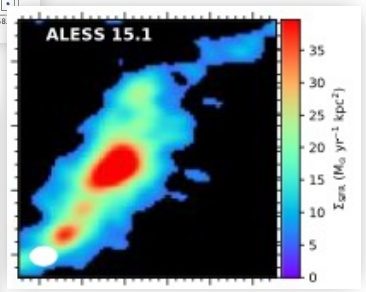
< 0.01 pc



~100 pc



~500 pc

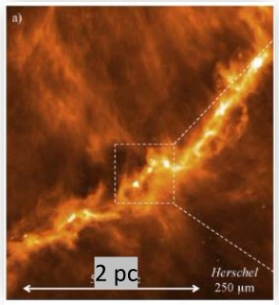


BigSF: Bridging Astro & AI for SF Study

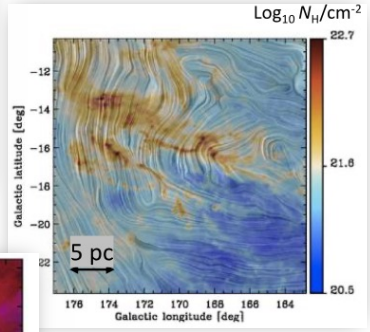
Key Question: How does the environment affect star formation (SF)?

Filaments

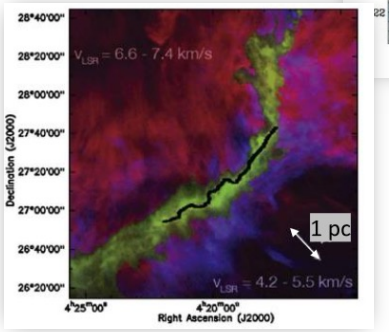
$N(H_2)$



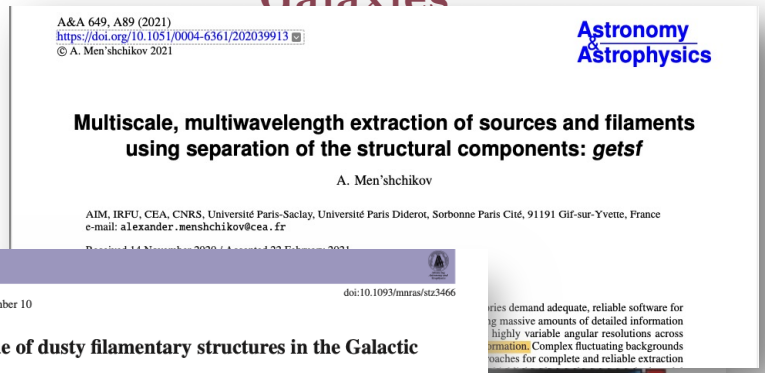
$N(H_2)$ & Magnetic field



Velocity field



Galaxies



A&A 649, A89 (2021)
https://doi.org/10.1051/0004-6361/202039913
© A. Men'shchikov 2021

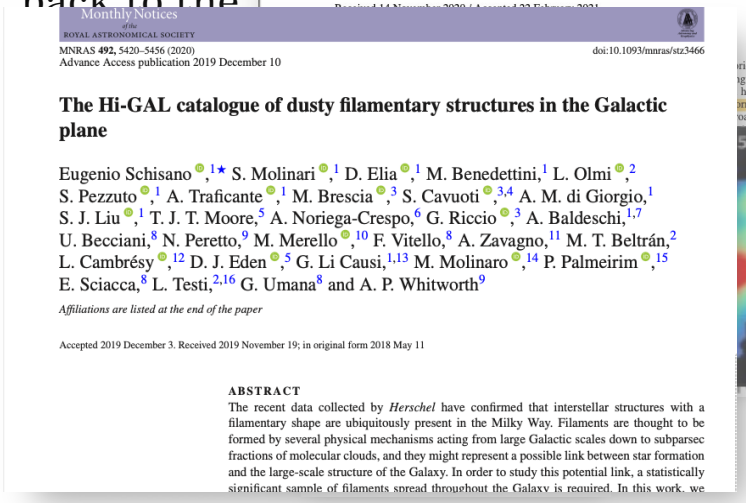
Astronomy & Astrophysics

Multiscale, multiwavelength extraction of sources and filaments using separation of the structural components: *getsf*

A. Men'shchikov

AIM, IRFU, CEA, CNRS, Université Paris-Saclay, Université Paris Diderot, Sorbonne Paris Cité, 91191 Gif-sur-Yvette, France
e-mail: alexander.menshchikov@cea.fr

back to the



Monthly Notices of the ROYAL ASTRONOMICAL SOCIETY

MNRAS **492,** 5420–5456 (2020) doi:10.1093/mnras/stz3466
Advance Access publication 2019 December 10

The Hi-GAL catalogue of dusty filamentary structures in the Galactic plane

Eugenio Schisano,[1]★ S. Molinari,[1] D. Elia,[1] M. Benedettini,[1] L. Olmi,[2] S. Pezzuto,[1] A. Traficante,[1] M. Brescia,[3] S. Cavuoti,[3,4] A. M. di Giorgio,[1] S. J. Liu,[1] T. J. T. Moore,[5] A. Noriega-Crespo,[6] G. Riccio,[3] A. Baldeschi,[1,7] U. Becciani,[8] N. Peretto,[9] M. Merello,[10] F. Vitello,[8] A. Zavagno,[11] M. T. Beltrán,[2] L. Cambrésy,[12] D. J. Eden,[5] G. Li Causi,[1,13] M. Molinaro,[14] P. Palmeirim,[15] E. Sciacca,[8] L. Testi,[2,16] G. Umana[8] and A. P. Whitworth[9]

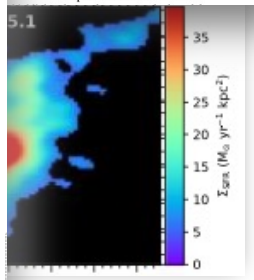
Affiliations are listed at the end of the paper

Accepted 2019 December 3. Received 2019 November 19; in original form 2018 May 11

ABSTRACT

The recent data collected by *Herschel* have confirmed that interstellar structures with a filamentary shape are ubiquitously present in the Milky Way. Filaments are thought to be formed by several physical mechanisms acting from large Galactic scales down to subparsec fractions of molecular clouds, and they might represent a possible link between star formation and the large-scale structure of the Galaxy. In order to study this potential link, a statistically significant sample of filaments spread throughout the Galaxy is required. In this work, we

ries demand adequate, reliable software for
g massive amounts of detailed information
highly variable angular resolutions across
ormation. Complex fluctuating backgrounds
oaches for complete and reliable extraction



Classical signal & image extraction methods

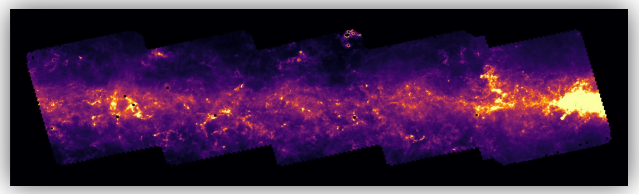
Big Data → use **AI** to detect filaments

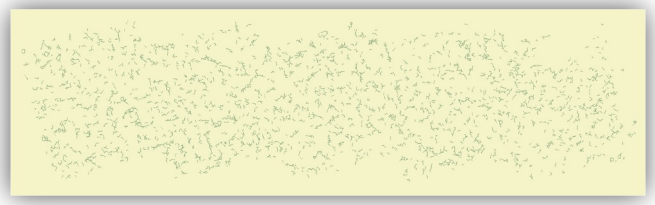
Dataset: Hi-GAL Catalogue

Column Density $N(H_2)$ Maps

Filaments Masks

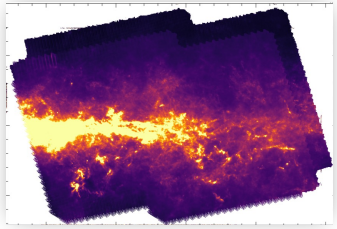
002-011°

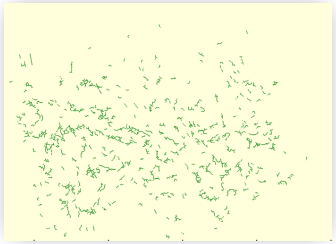




37 Mosaics

358-000°





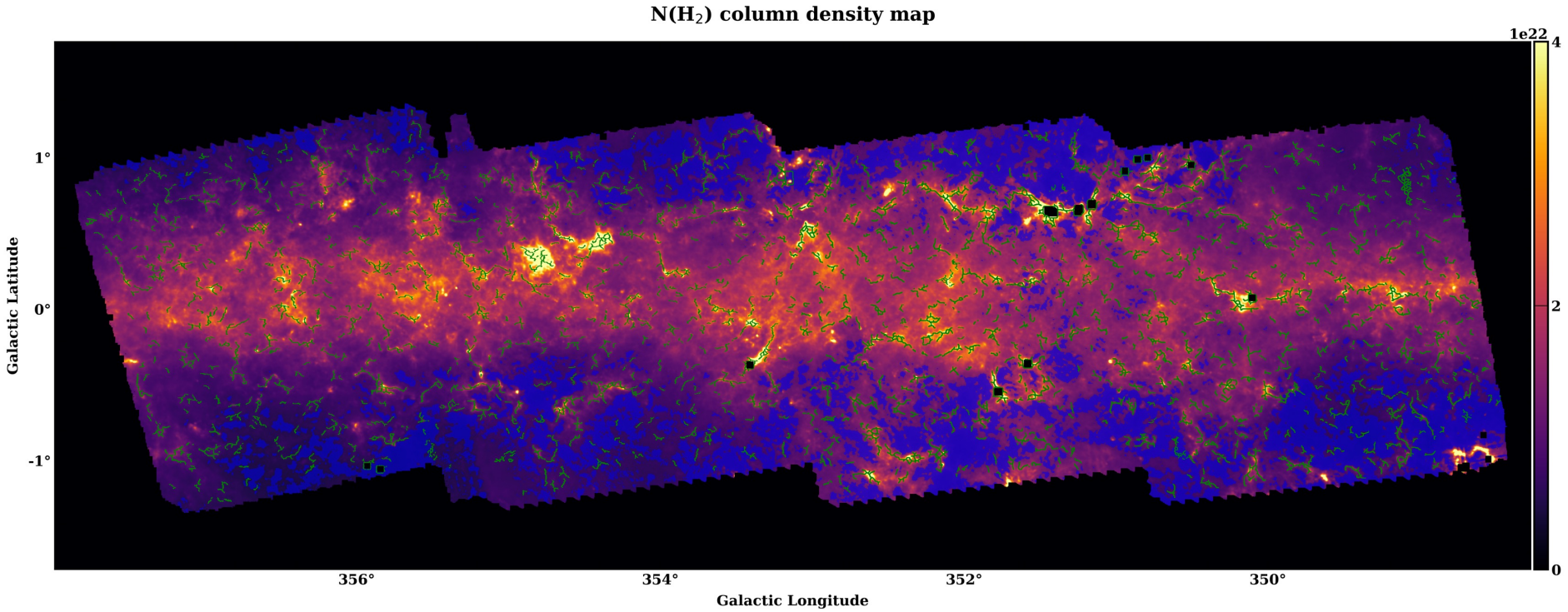
Herschel Infrared Galactic plane Survey (Molinari et al. 2010)

Wavelengths: 70 to 500 μm

Mosaic: spans ~10° in Galactic longitude

~32,000 filaments

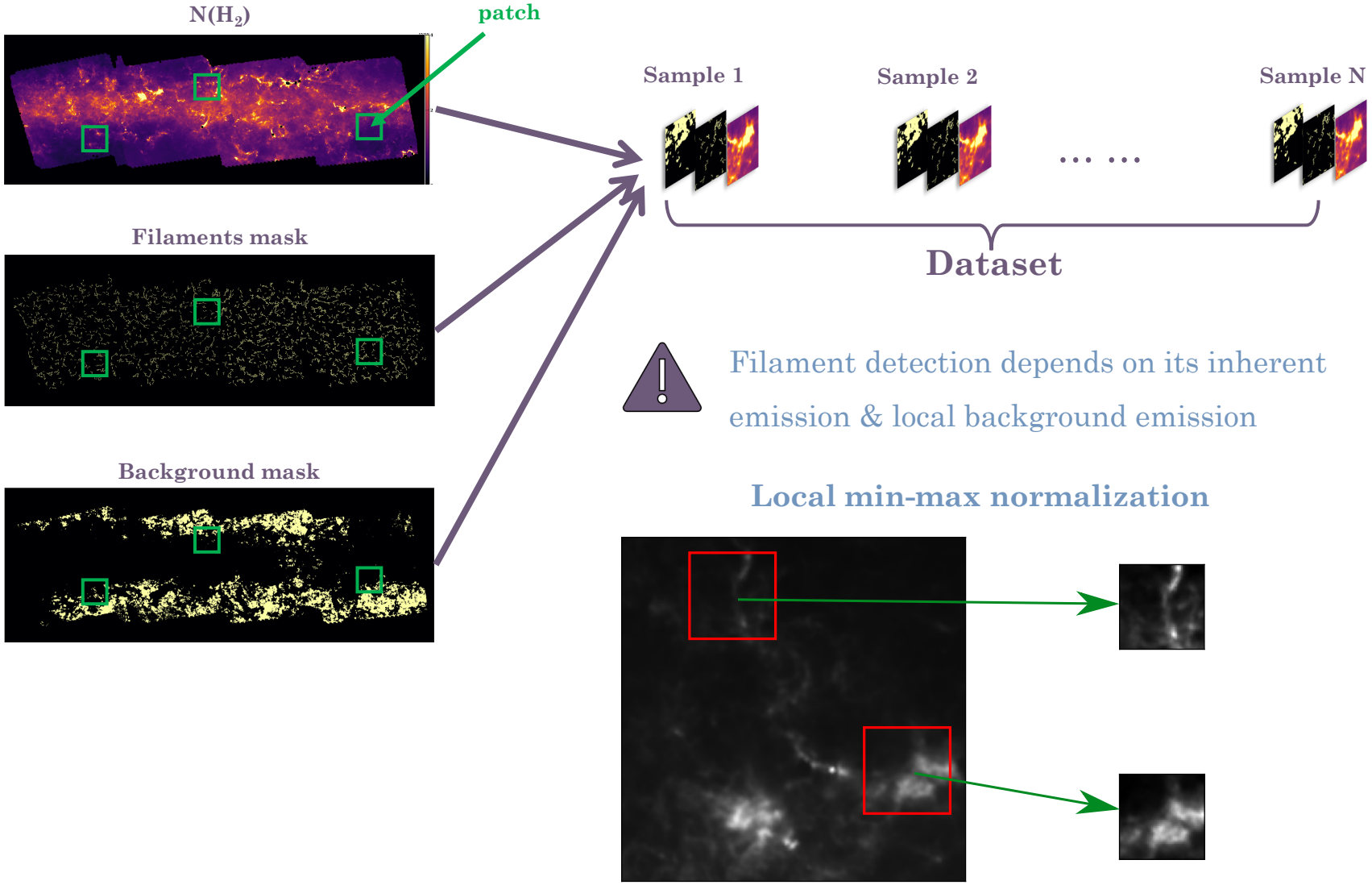
Dataset: Hi-GAL Catalogue



- High dynamic range: $10^{20} \rightarrow 10^{23}$ cm^{-2}
- Incomplete ground-truth
- **Local background** definition → zones of low intensity where no filaments are observed

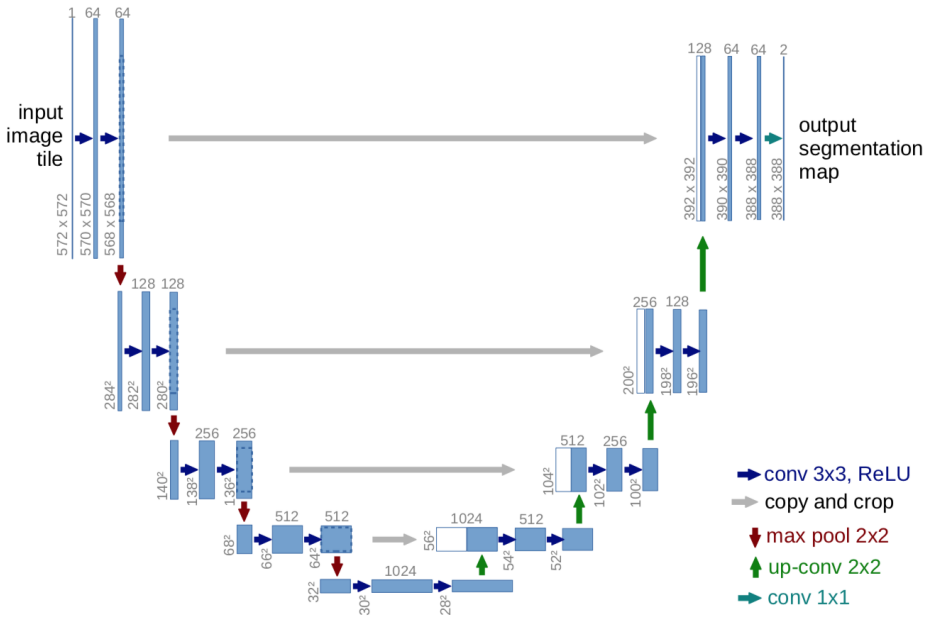
→ Labeled data: filaments (green) & background (blue)

Dataset: Hi-GAL Catalogue (cont'd)



U-Net for Filaments Segmentation

State-of-the-art in automatic segmentation



Patch-size of 32 × 32 (p = 32)

N ~ 51,000

Batch-size = 64 patches

epochs = 100

Training/validation/test sets: 80%,10%, 10%

Data augmentation (rotations & flips)

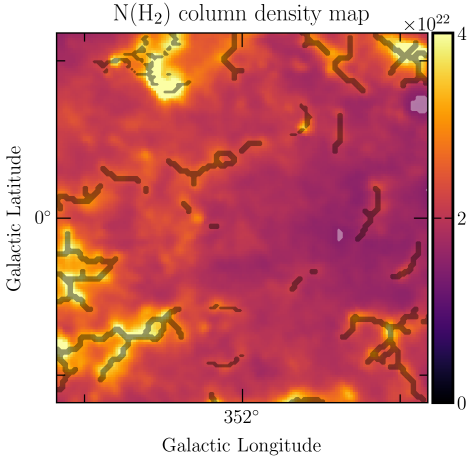
learning rate: $lr_{\text{init}} \in [10^{-5}, 10^{-4}, 10^{-3}, 10^{-2}]$

Zhang, J. https://towardsdatascience.com/unet-line-by-line-explanation-9b191c76baf5

Ronneberger+2015 arXiv:1505.04597

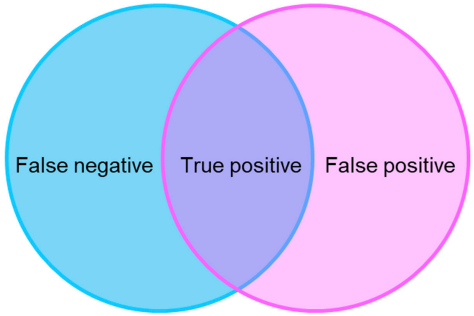
Hi-GAL Structures Recovered

Filaments in Hi-GAL column density maps



scores on test set

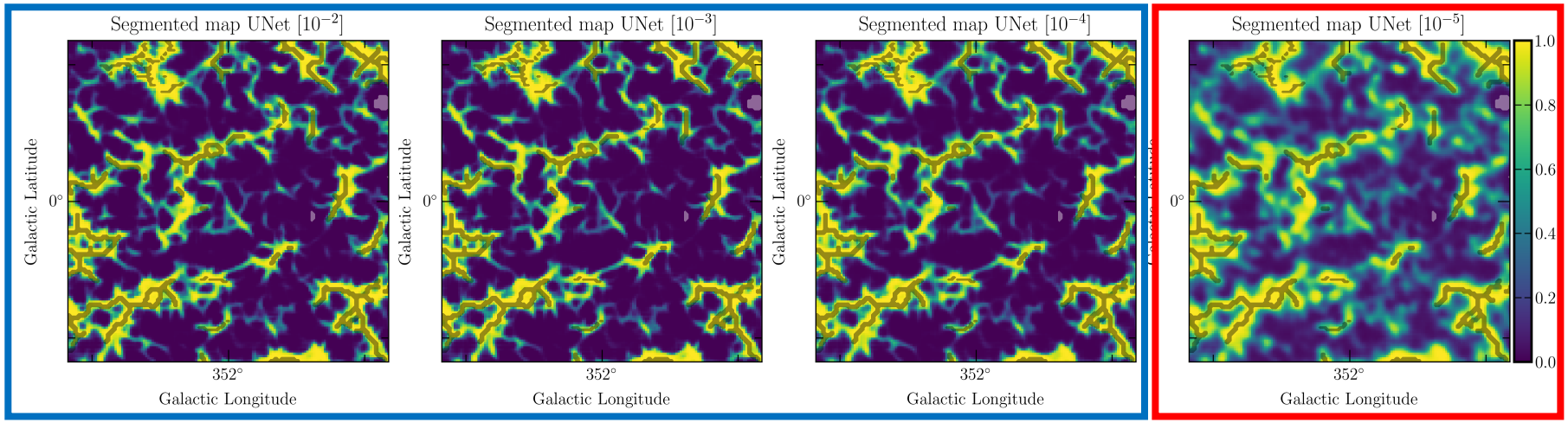
Score (%) / Model	Dice [0.2]	Dice [0.4]	Dice [0.6]	Dice [0.8]
UNet [10^{-2}]	93.19	94.34	94.07	92.06
UNet [10^{-3}]	93.13	94.37	94.16	92.43
UNet [10^{-4}]	92.75	94.09	93.79	91.9
UNet [10^{-5}]	82.4	89.6	91.3	87.66



$$Dice = \frac{2 \times TP}{(TP + FN) + (TP + FP)} = \frac{2 \times}{+}$$

Thicker filaments

2 times more pixels segmented as filaments at $thr = 0.8$

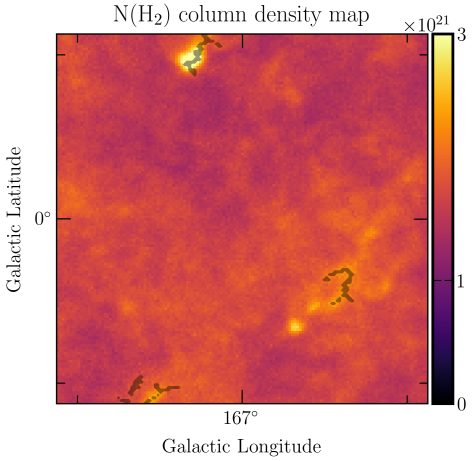


Salient structures

Fainter structures

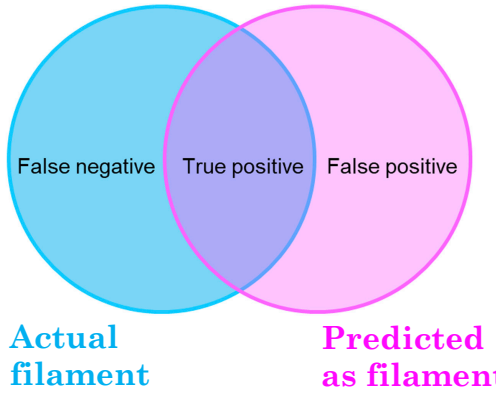
Hi-GAL Structures Recovered

Filaments in Hi-GAL column density maps



scores on test set

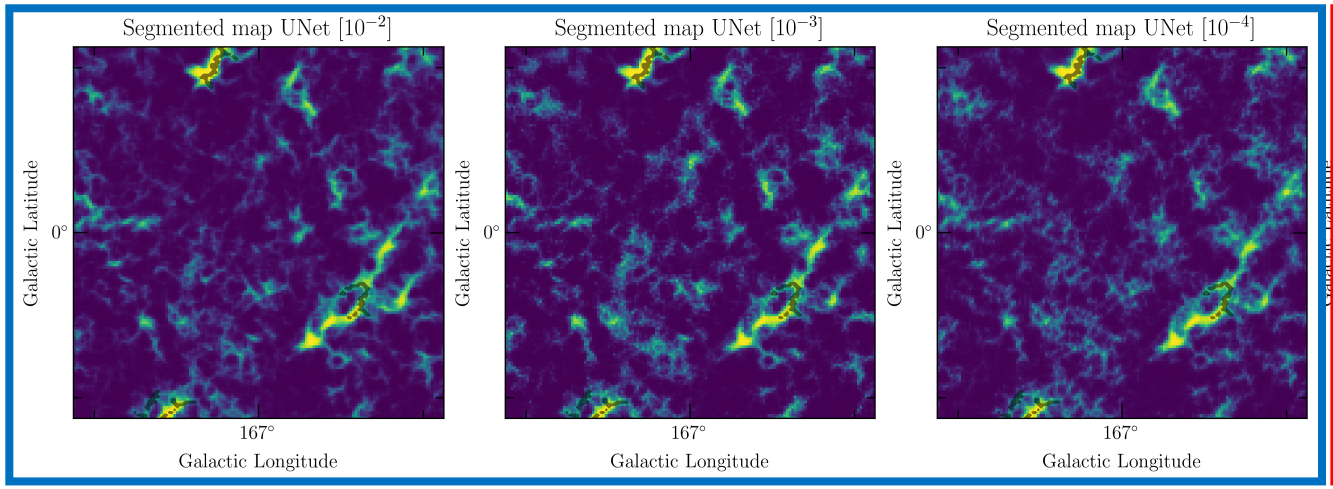
Score (%) / Model	Dice [0.2]	Dice [0.4]	Dice [0.6]	Dice [0.8]
UNet [10^{-2}]	93.19	94.34	94.07	92.06
UNet [10^{-3}]	93.13	94.37	94.16	92.43
UNet [10^{-4}]	92.75	94.09	93.79	91.9
UNet [10^{-5}]	82.4	89.6	91.3	87.66

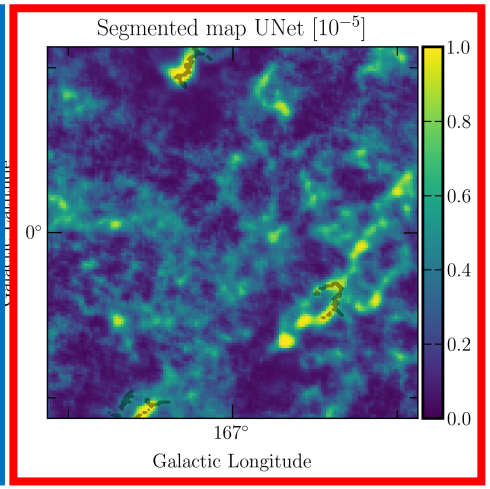


$$Dice = \frac{2 \times TP}{(TP + FN) + (TP + FP)} = \frac{2 \times}{+}$$

Thicker filaments

2 times more pixels segmented as filaments at $thr = 0.8$



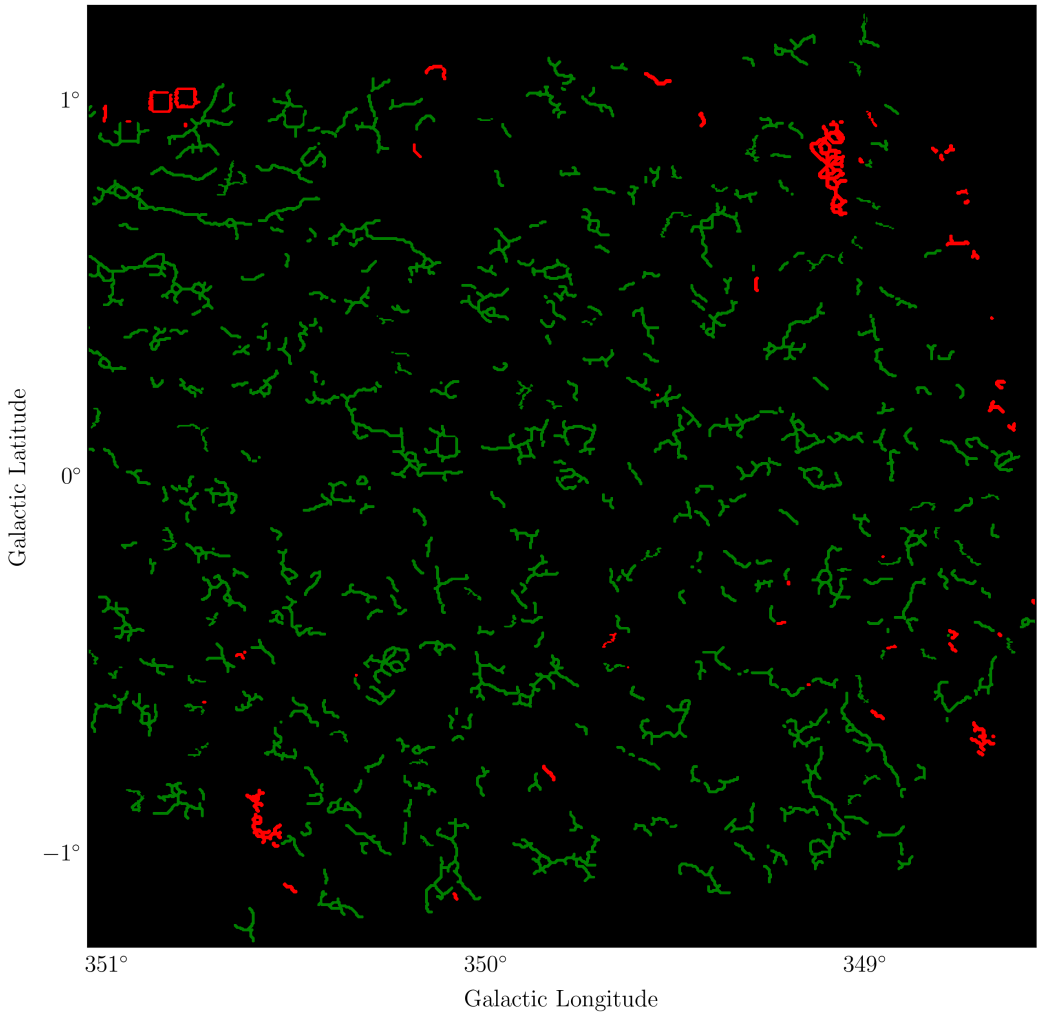


Salient structures

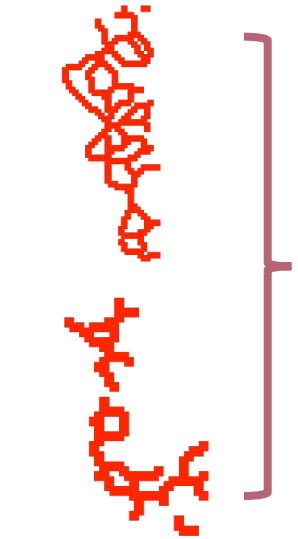
Fainter structures

Hi-GAL Missed Structures

Groundtruth & Missed structures



~15% of MS rate at *thr* = 0.8



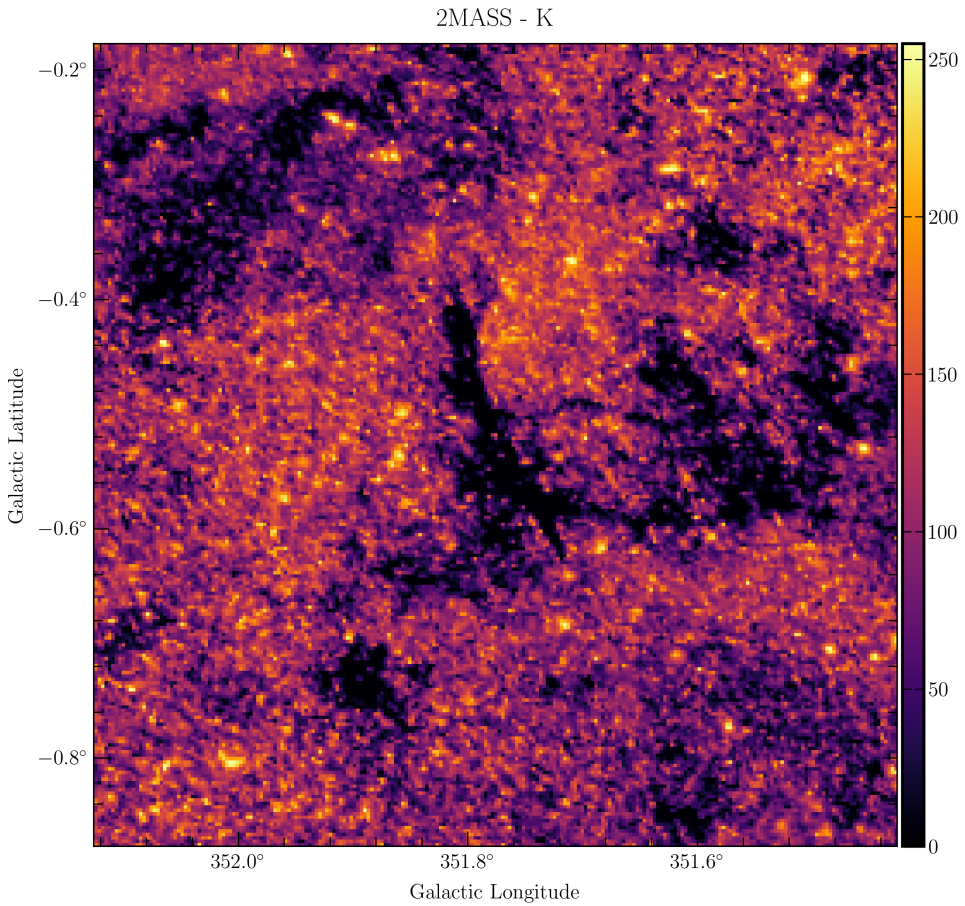
Mislabeled filaments

saturated pixels

New Structures Revealed

Empirical check of the new revealed structure using other wavelengths

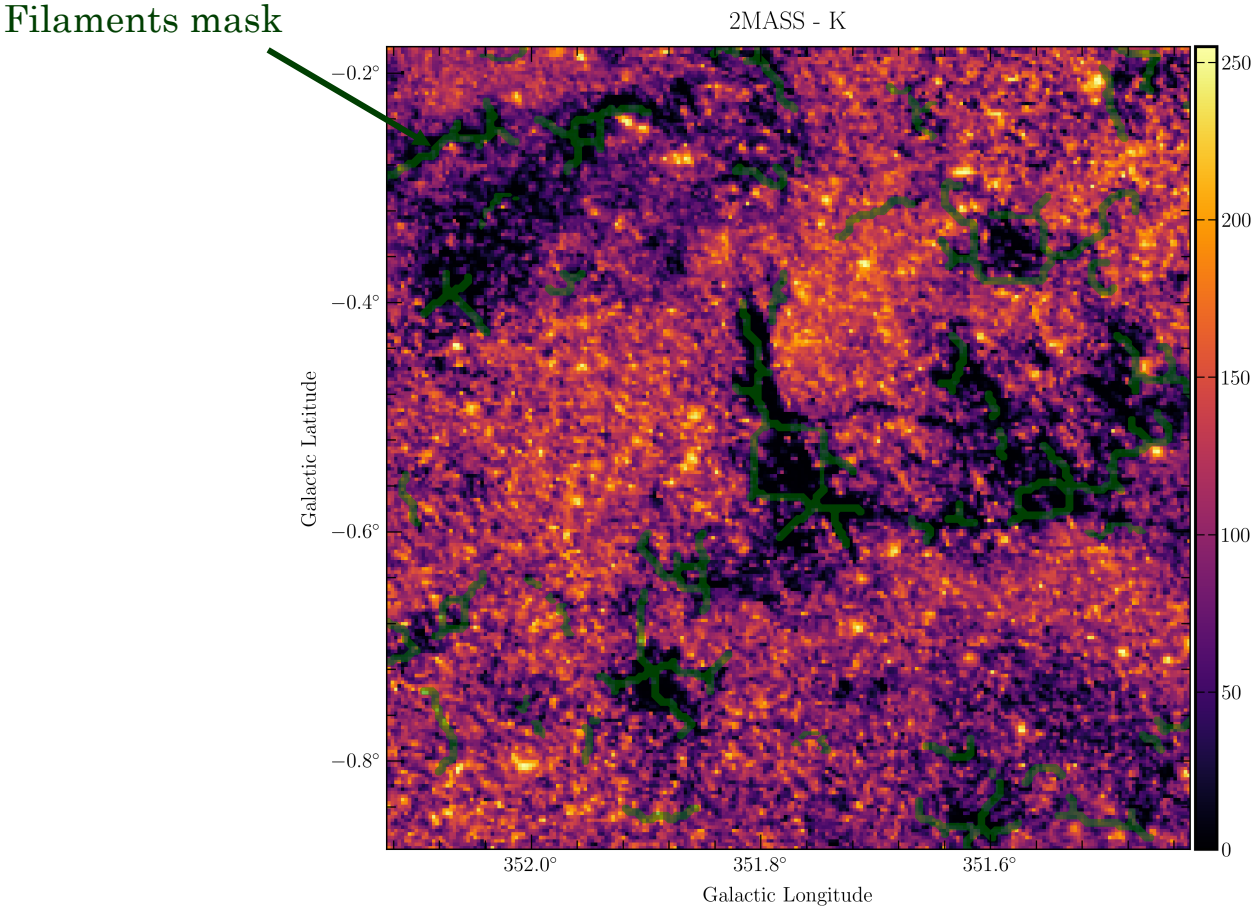
Filament G351.776-0.527 observed in 2MASS K band (near-infrared) image



New Structures Revealed

Empirical check of the new revealed structure using other wavelengths

Filament G351.776-0.527 observed in 2MASS K band (near-infrared) image

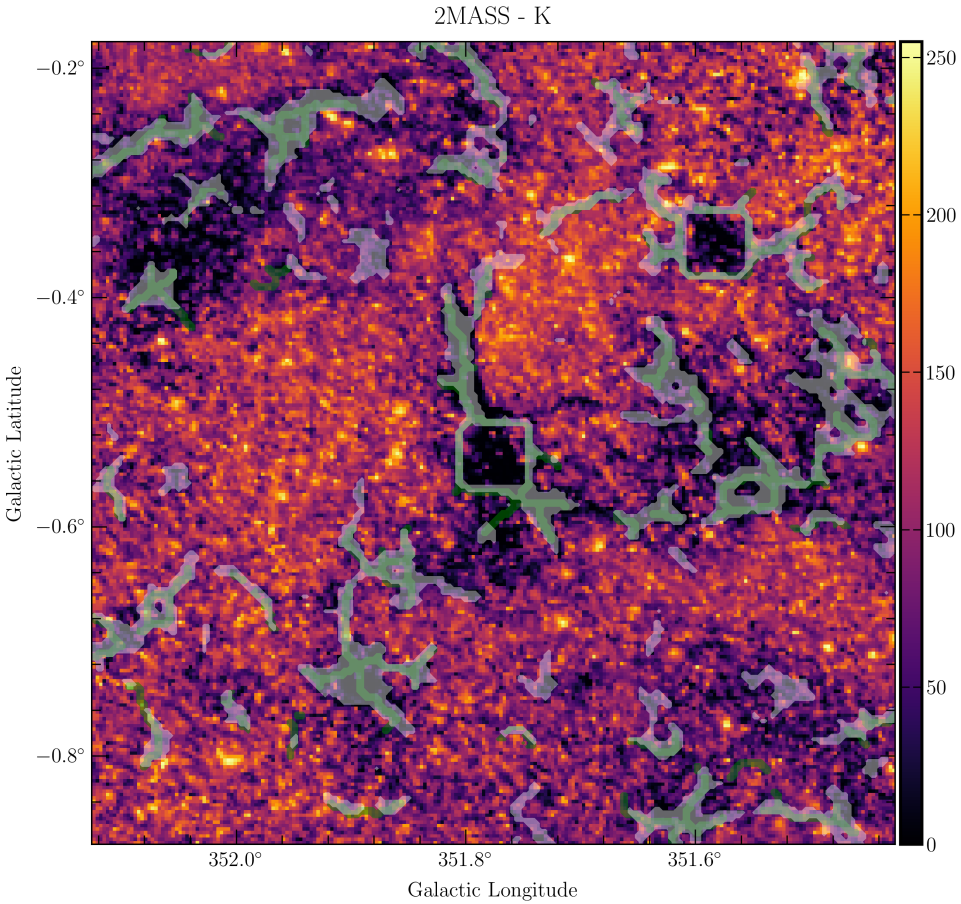


New Structures Revealed

Empirical check of the new revealed structure using other wavelengths

Filament G351.776-0.527 observed in 2MASS K band (near-infrared) image

UNet[10^{-4}] segmentation at classification threshold = 0.8

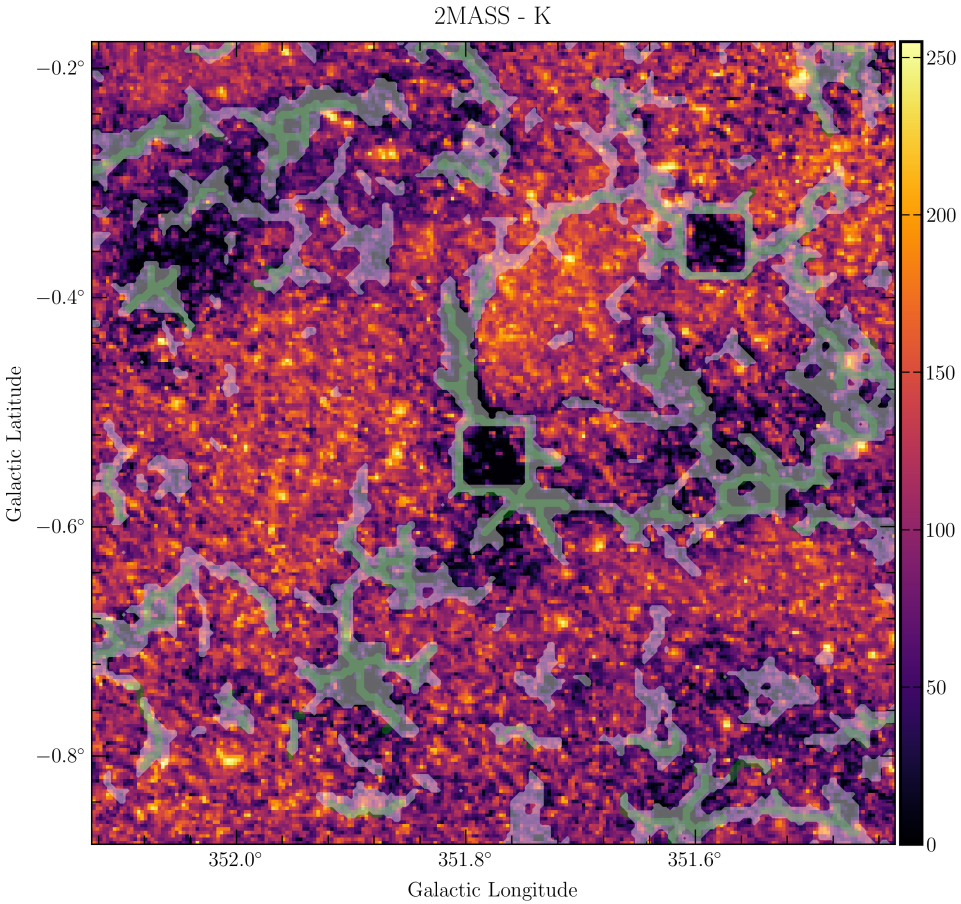


New Structures Revealed

Empirical check of the new revealed structure using other wavelengths

Filament G351.776-0.527 observed in 2MASS K band (near-infrared) image

UNet[10^{-4}] segmentation at classification threshold = 0.5

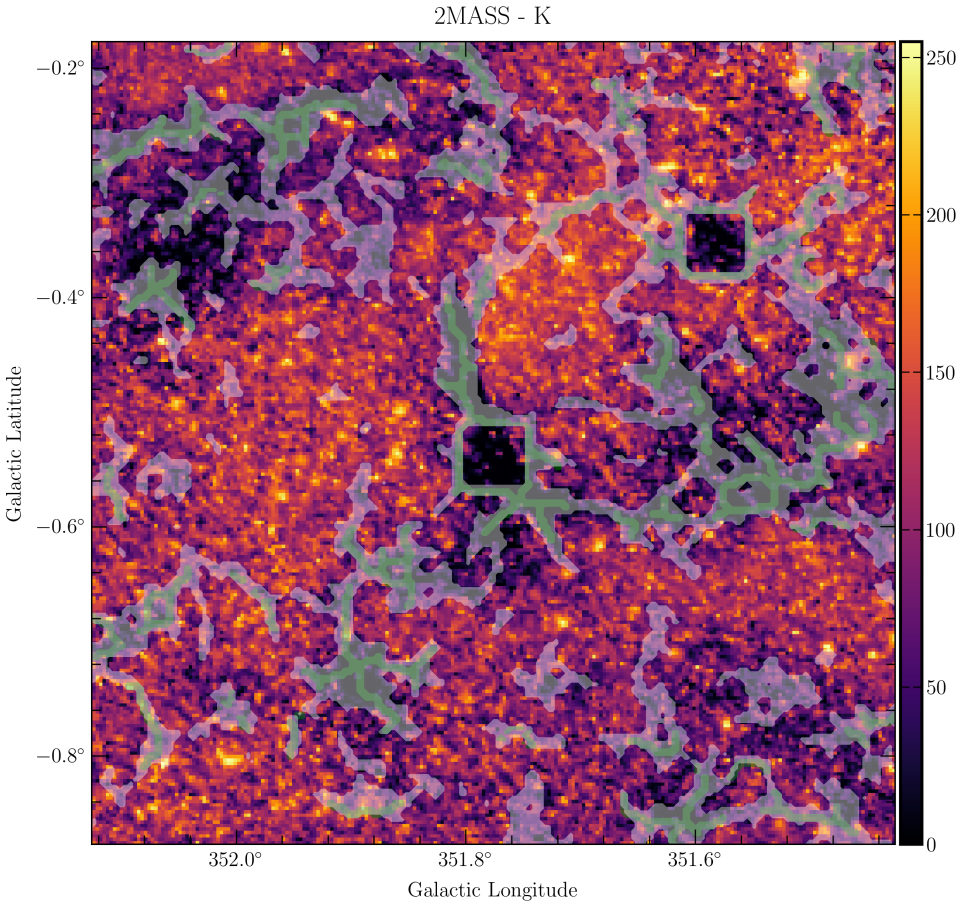


New Structures Revealed

Empirical check of the new revealed structure using other wavelengths

Filament G351.776-0.527 observed in 2MASS K band (near-infrared) image

UNet[10^{-4}] segmentation at classification threshold = 0.38 (optimal)



Conclusion & Perspectives

- ✓ **Proof of concept:** Use Deep Learning method & Hi-GAL catalogue to segment filaments
- ✓ Existing Hi-GAL structures successfully recovered
- ✓ **New revealed** structures corresponding to filaments when checking at other wavelengths

- **Incomplete** ground-truth → explore other methods (e.g. semi-supervised)
- Biased by the eye of expert → bridge **simulation and real data** for more robust validation of the results

- From revealing to **extracting** filaments: Converge to a **unanimous definition of filament** in the scientific community

Thank you for your attention!

Annex

What is a filament?

"Two dimensional, cylindric-like structure that is elongated and shows a higher brightness contrast with respect to its surroundings"

Schisano et al. (2020)