Carving out the low surface brightness astronomical signal

(using NoiseChisel, part of GNU Astronomy Utilities or Gnuastro)

Mohammad Akhlaghi
Staff Researcher at
Centro de Estudios de Física del Cosmos de Aragón (CEFCA), Teruel, Spain

Shiraz University Gnuastro Workshop
24th of October 2023

Contact: mohammad@akhlaghi.org
Slides available at: http://akhlaghi.org/pdf/noisechisel.pdf

Ayuda ICTS-MRR-2021-03-CEFCA, financiada por:
MINISTERIO DE CIENCIA E INNOVACIÓN
Plan de Recuperación, Transformación y Resiliencia
ESPAÑA PUEDE.
Financiado por la Unión Europea NextGenerationEU

CEFCA
Centro de Estudios de Física del Cosmos de Aragón

Preamble: J-PAS survey

(imaging $+8000$ deg^2 in 54 narrow-band filters)

LSST has 6 filters in optical (as a comparison)

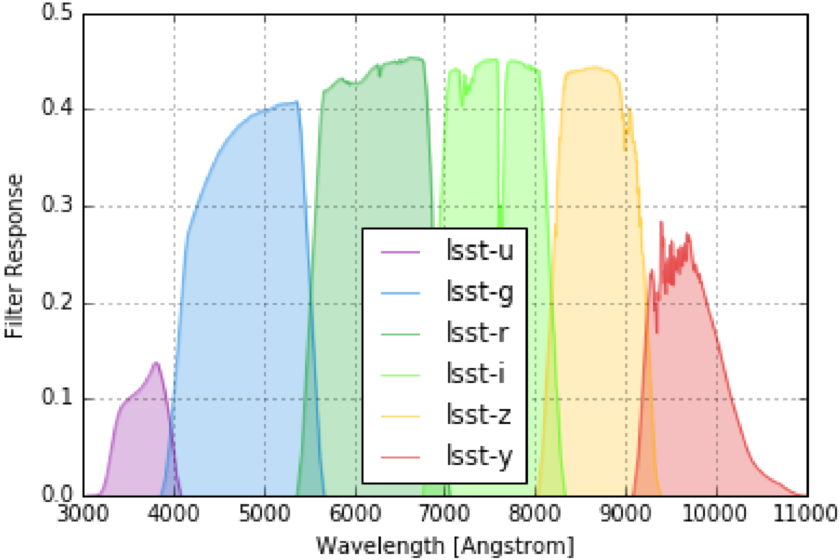


Image from community.lsst.org

J-PAS has 54 narrow, 2 medium, and 1 broad filter in the optical

(the broad filter, iSDSS, is not shown here; the two medium bands are on the two extremes)

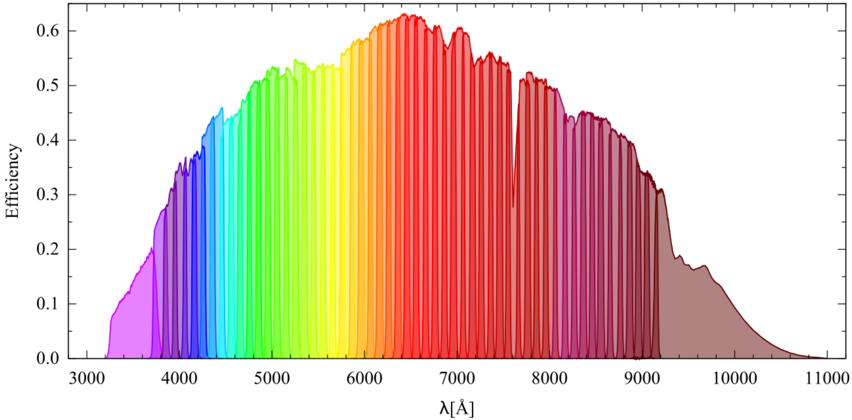
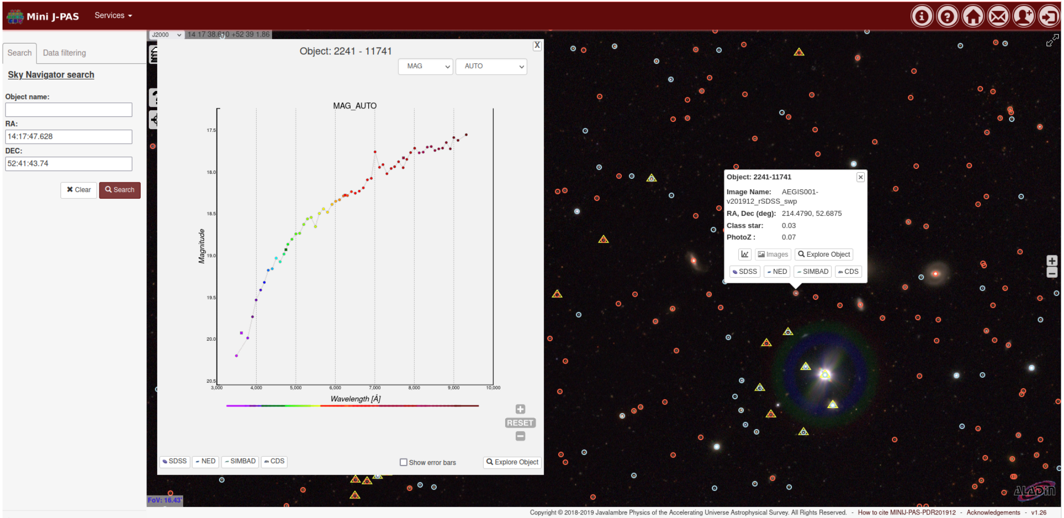


Image from Bonoli et al. 2021 (`2021A&A...653A..31B`)

J-PAS also has a pretty nice web navigator



Screenshot from http://archive.cefca.es/catalogues/minijpas-pdr201912/navigator.html)

Conducted at OAJ (Observatorio Astrofísico de Javalambre)

Image from Wikipedia

OAJ is in Teruel, Spain

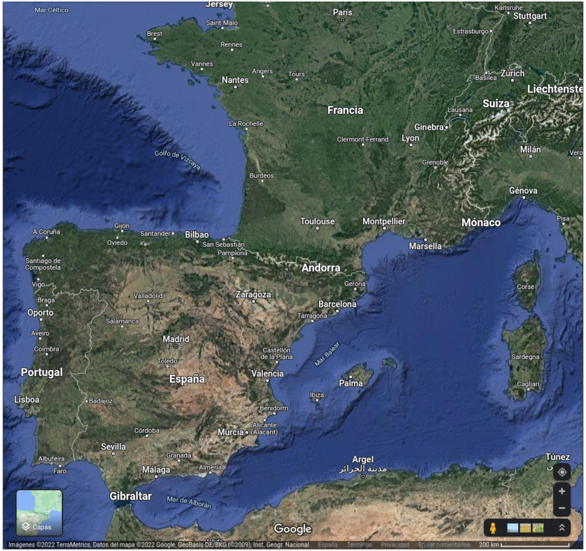


Image from Google Maps

OAJ is in Teruel, Spain

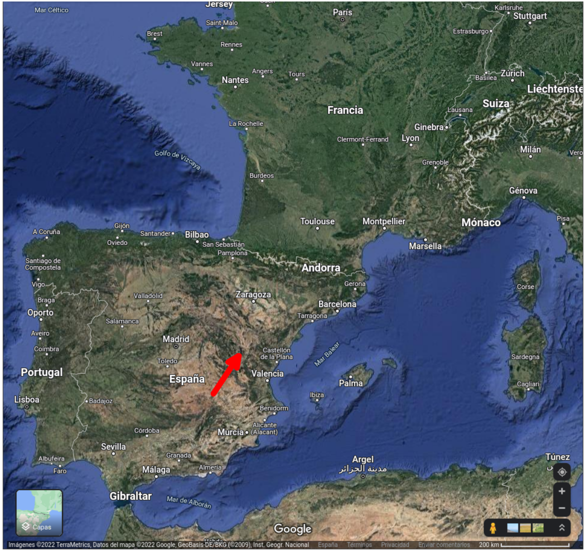
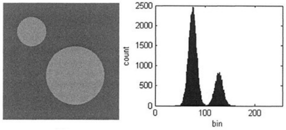


Image from Google Maps

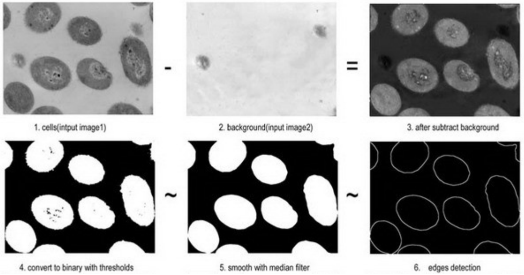
The problem of Detection

Traditional detection (sharp edges, high S/N)

When there are sharp and high S/N edges, a sufficiently high threshold can avoid the noise, so we can call them Signal-based detection.

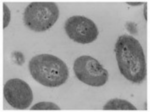
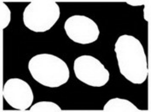
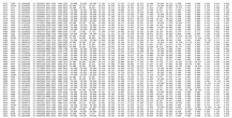


© what-when-how.com



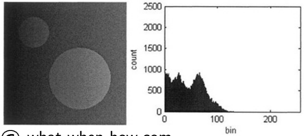
© OriginLab.com

Outline is:

 → Threshold →  → Measure → 

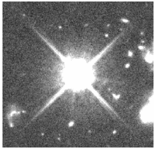
But the traditional method isn't sufficient for astronomical objects, because ...

... they don't have such sharp edges.
... they can have a huge diversity of shapes and sizes.

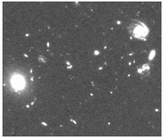


© what-when-how.com

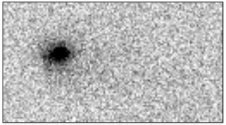
A star:



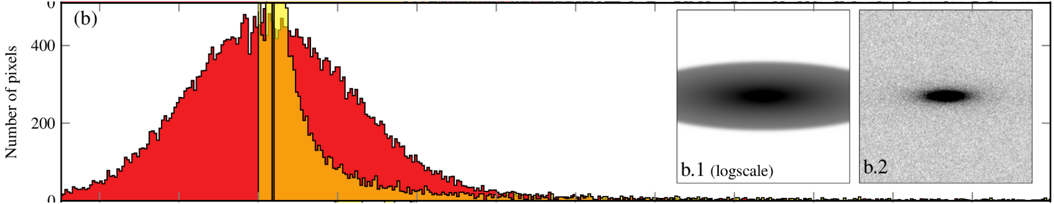
Some galaxies:



A main-belt comet:



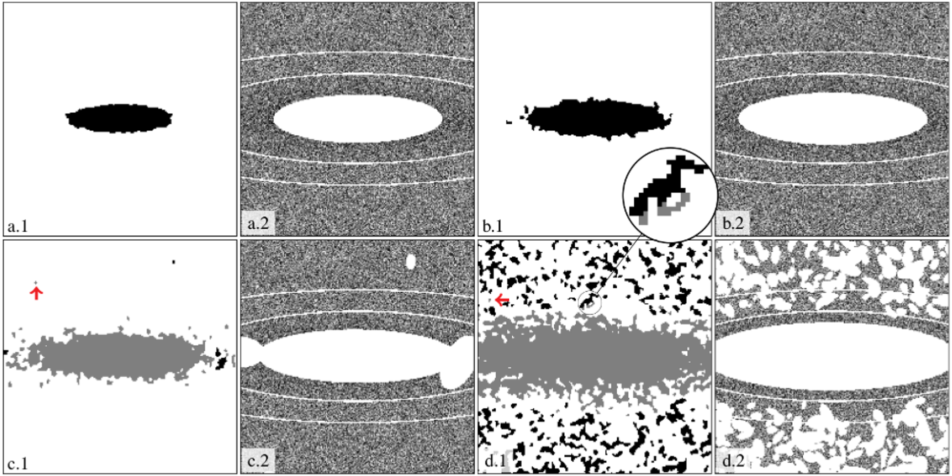
A mock galaxy: (Yellow/left-image: only signal. Red/right-image: signal and noise.)



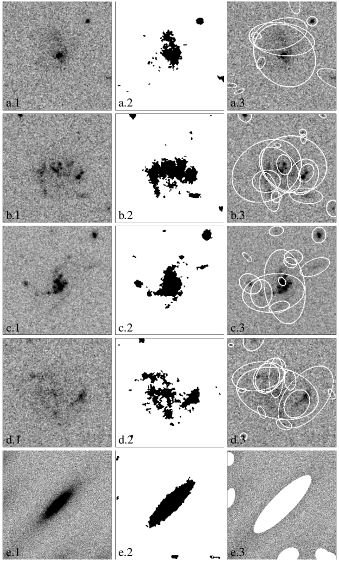
So for astronomical objects ...

... a threshold designed to avoid the noise (signal-based detection) will miss a lot of the signal. Decreasing the threshold will result in many false detections.

Our only hope is to model the brighter parts (implicitly assuming the outskirts fit the same model).

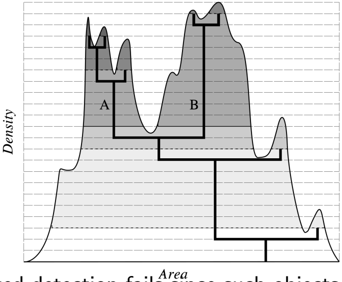


Examples on real galaxy images



Real observed galaxies:

- Are not a clean ellipse.
  - Can be clumpy.
  - Can be diffuse.
  - Can have spiral arms.
  - Can be on the edge of the image.
- SExtractor's deblending uses layers and the parent is used to identify true peaks (systematic biases):



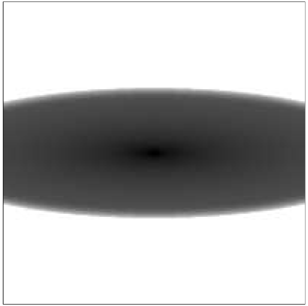
Signal-based detection fails since such objects do not satisfy its *a priori* assumptions.

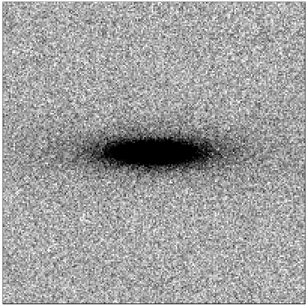
NoiseChisel – Detection – Basics

Aims:

- Threshold must be independent of the Sky.
- Impose negligible assumptions on signal.
- Accurately remove false detections.
- Use the actual data, not *a priori* models.

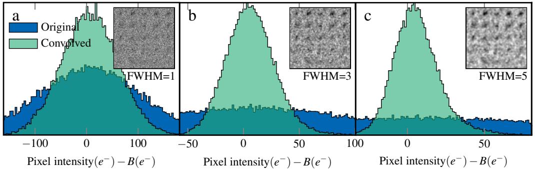
Model profile for demonstration:

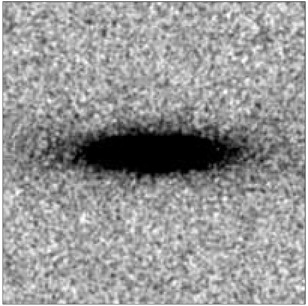




NoiseChisel – Detection – Convolution

- Convolution decreases dynamic range.
- **So:** Gaussian kernel, FWHM= 2pixels.



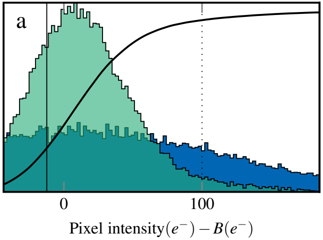


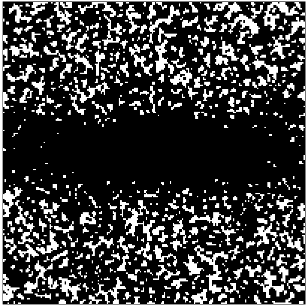
Since we are concerned with the sampling (noise) a-priori knowledge of the PSF (which relates to the signal) is no longer necessary and the same parameters work accurately on space-based and ground-based images.

✔ An assumption removed. Works on any image.

NoiseChisel – Detection – Threshold

- Use the cumulative pixel distribution.
- The threshold is set to the 0.3 quantile of the image.





> Since the threshold is now independent of Sky, we can accurately estimate the Sky once detection is complete.

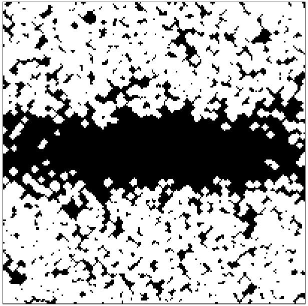
Threshold no longer defined by Sky.

NoiseChisel – Detection – Erode

Erosion: Foreground becomes background if touching.

- **Or:** we expand the holes.
- **Or:** we carve off the signal.

NoiseChisel name: a tool to carve off noise



No assumption on the shape of the object.

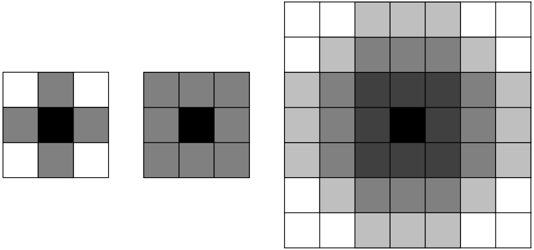
NoiseChisel – Detection – Open

Definitions:

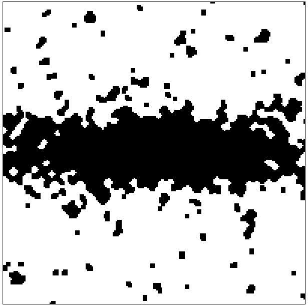
- **Dilation:** Inverse of erosion.
- **Opening:** Erosion followed by dilation.

In practice:

- Separates all the steps below.
- We use eight connectivity here (and four connectivity in the previous step.)

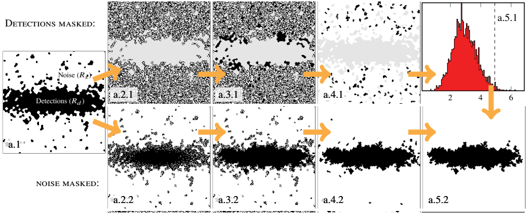


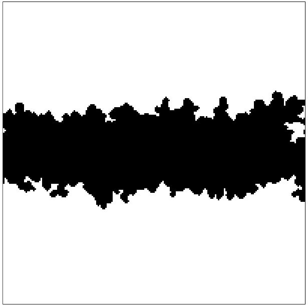
a b c



NoiseChisel – Detection – Remove false detections

- Use the ambient noise as a reference.
- The S/N of definite false detections is used:

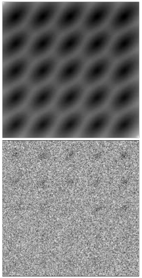


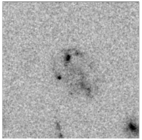


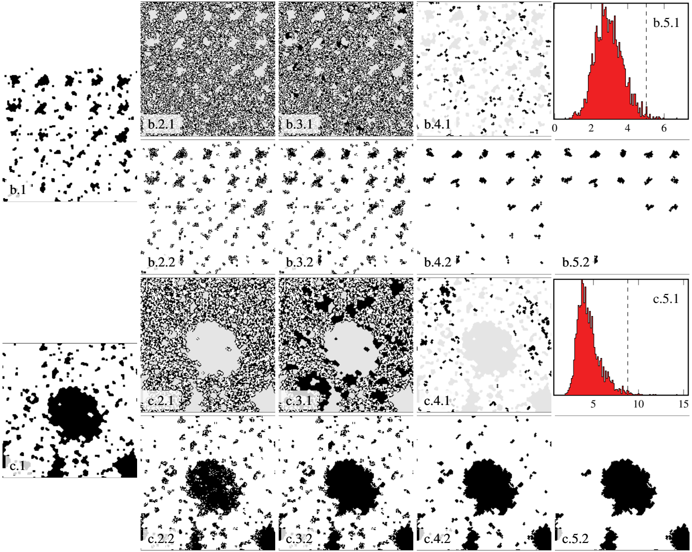
False detections are successfully removed with high accuracy.

✓ False detections are now identified for any image without hand-input values.

NoiseChisel – Detection – Remove false detections – Two more examples

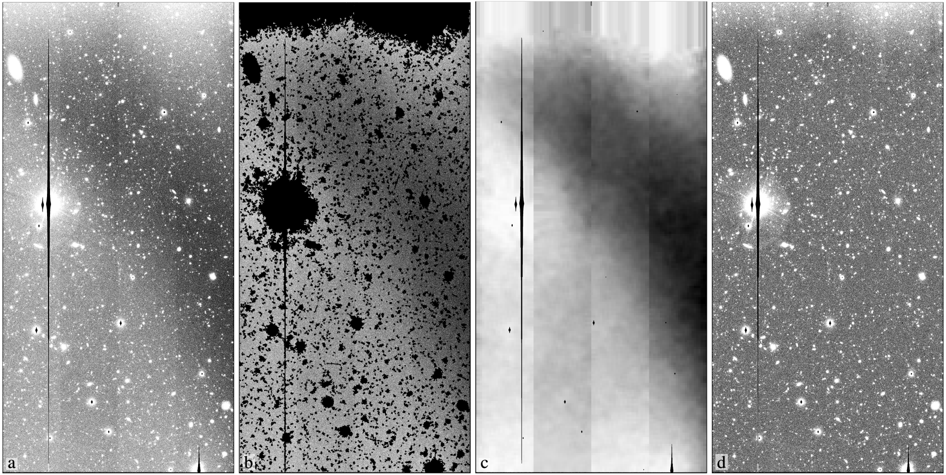




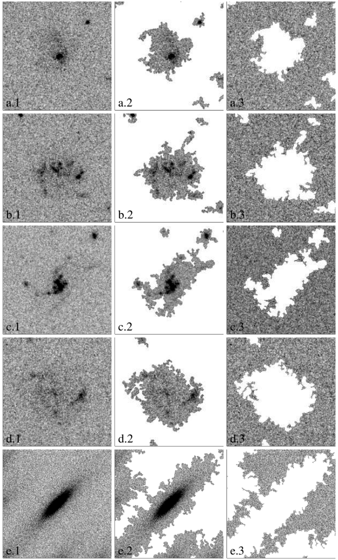


NoiseChisel - Accurate Sky estimation (after detection)

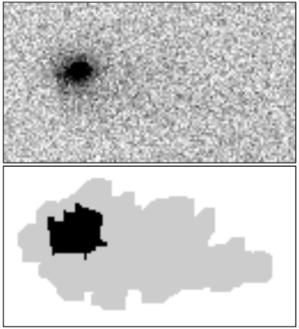
- Subaru SuprimeCam image after basic reduction.
- Sources are first detected, then the Sky value is measured (non-parametrically).
- CCD amplifier signature (bias subtraction residuals) removed.



NoiseChisel – Detection – Other real targets

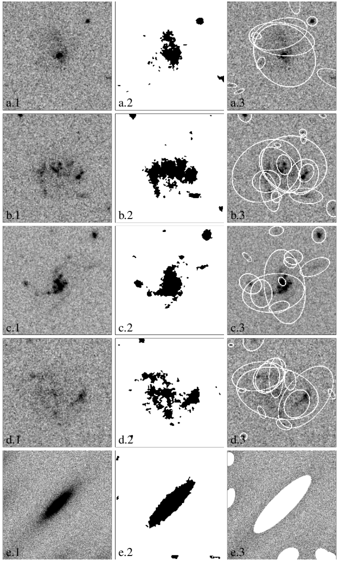


Detection of the diffuse and low surface brightness tail of a main-belt comet (image by H. Hsieh):



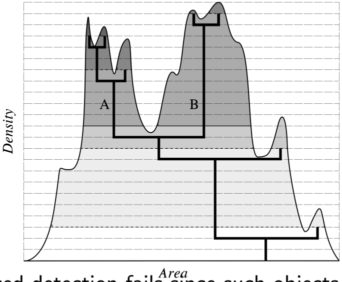
Noise-based detection: Works on any image with any target shape.

Just as a reminder... (of the old method, shown before)



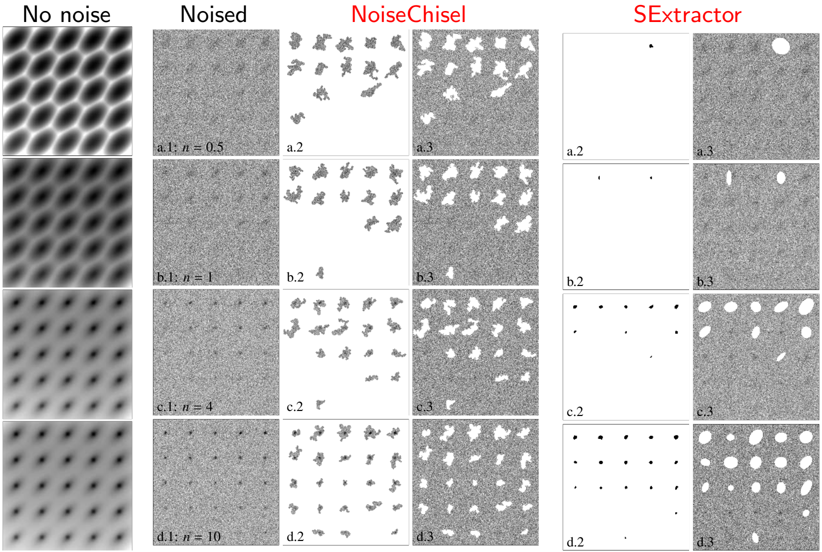
Real observed galaxies:

- Are not a clean ellipse.
    - Can be clumpy.
    - Can be diffuse.
    - Can have spiral arms.
    - Can be on the edge of the image.
- SExtractor's deblending uses layers and the parent is used to identify true peaks (systematic biases):

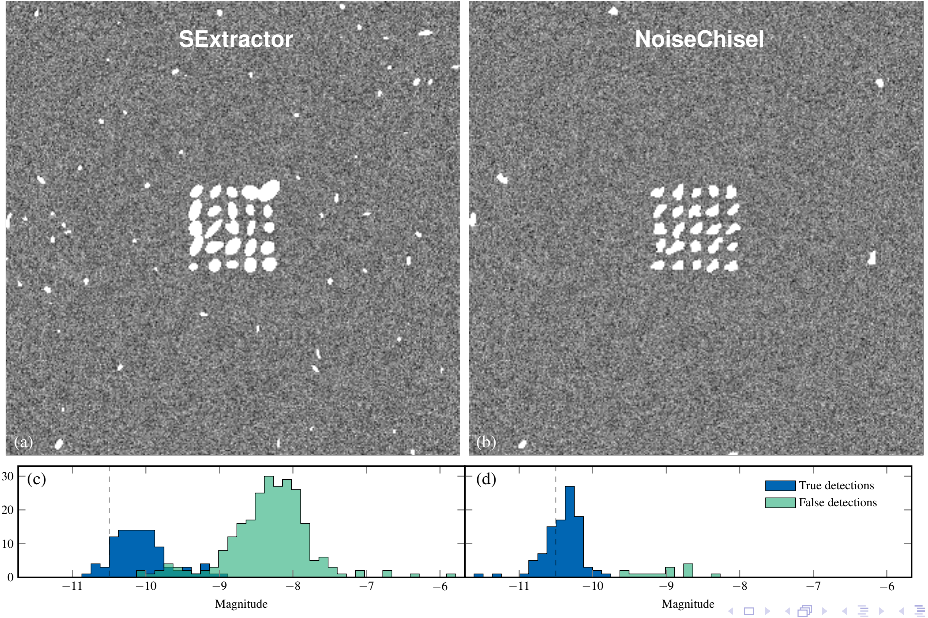


Signal-based detection fails since such objects do not satisfy its *a priori* assumptions.

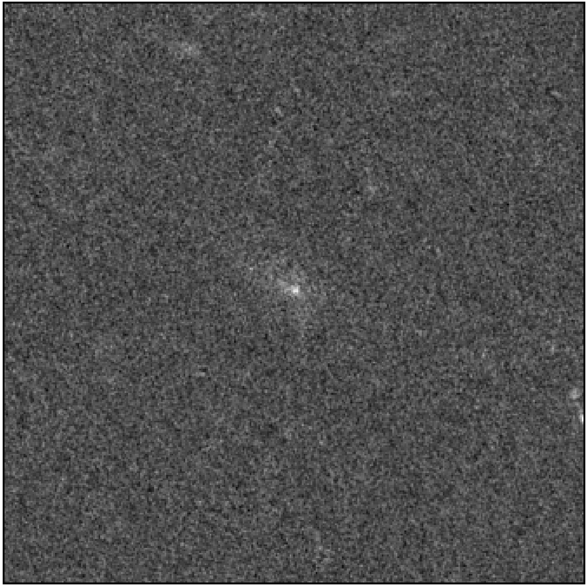
Rough completeness demo on mocks: NoiseChisel/SExtractor



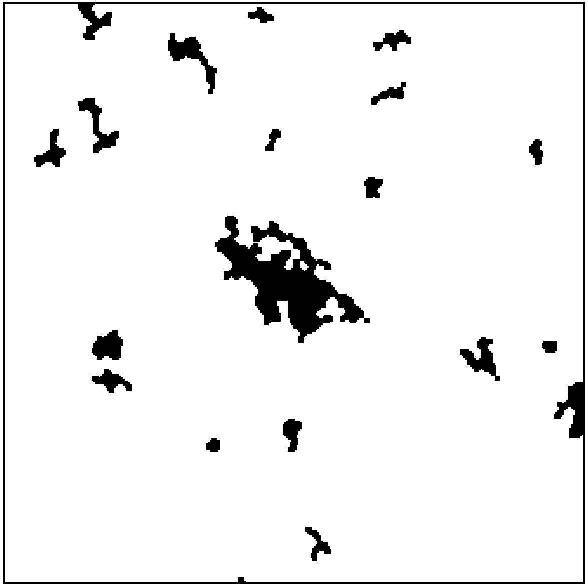
Purity and Magnitude dispersion test



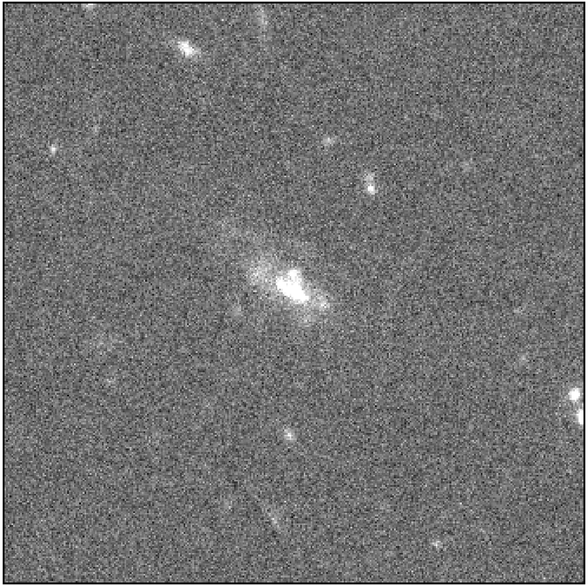
New growth demo: input (F435W, $\sim$ 700sec HST exposure)



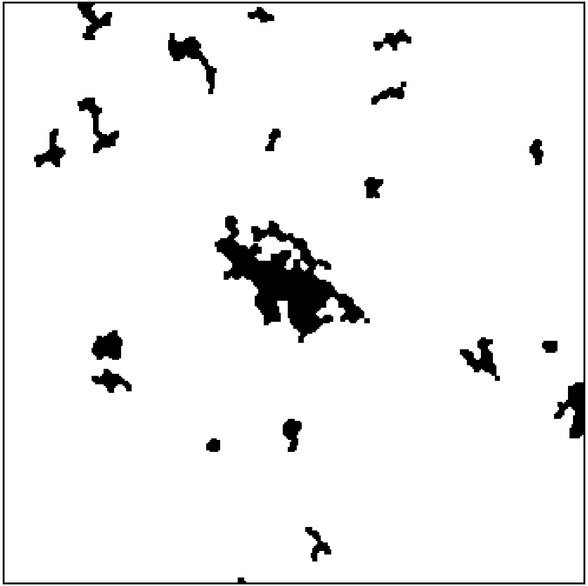
New growth demo: No growth



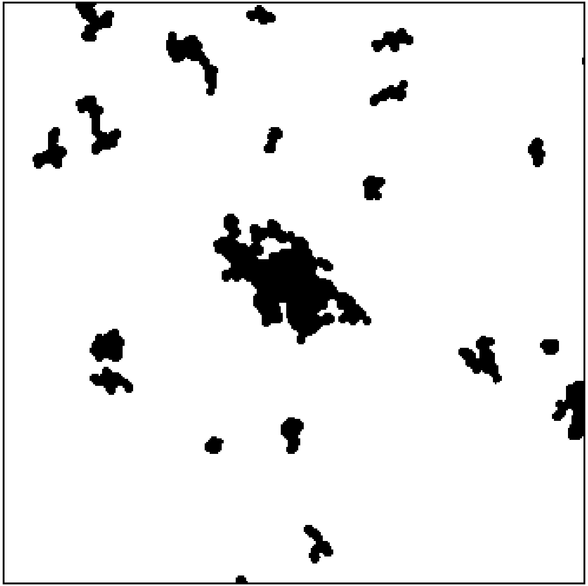
HST F435W ($\sim$ 1600sec exposure)



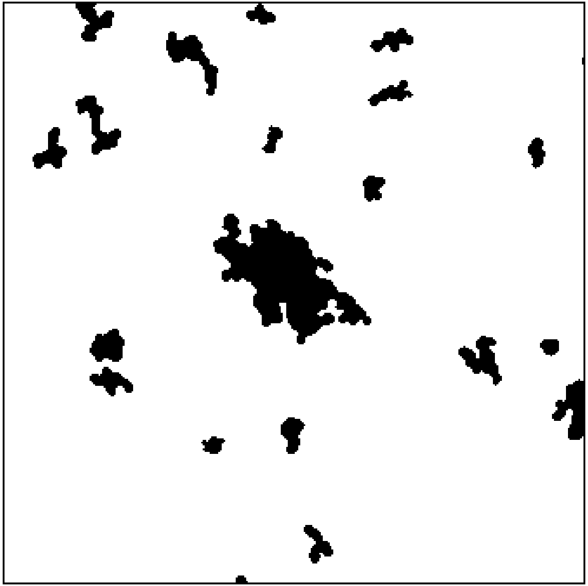
New growth demo: No growth



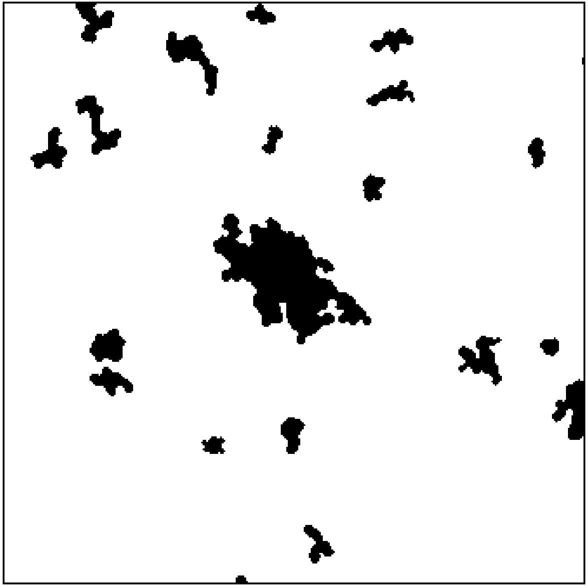
New growth demo: growth quantile: 0.99



New growth demo: growth quantile: 0.97

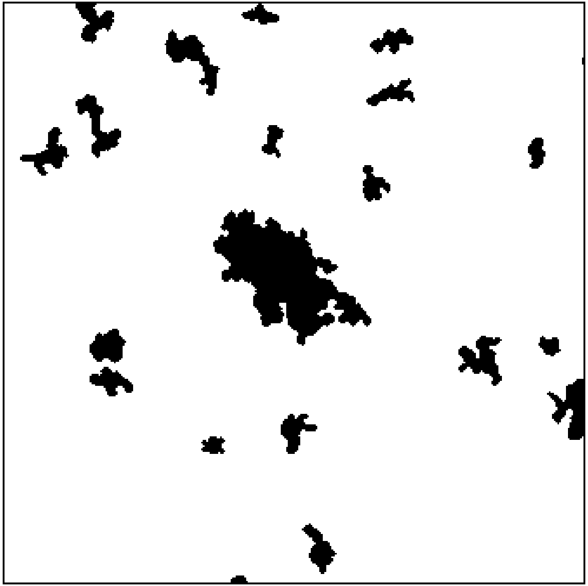


New growth demo: growth quantile: 0.95

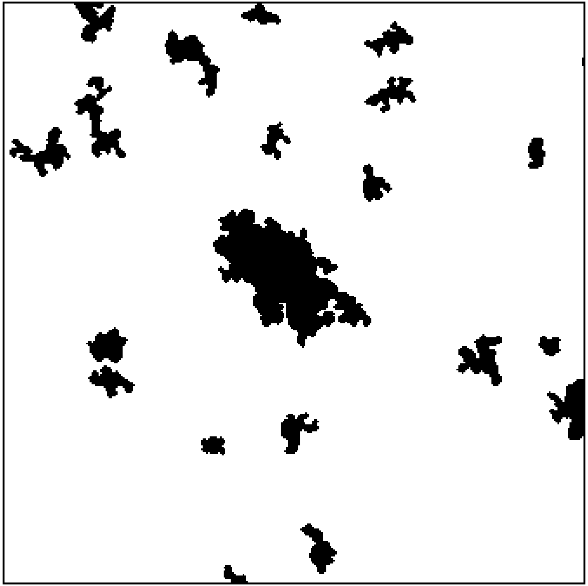


New growth demo: growth quantile: 0.92

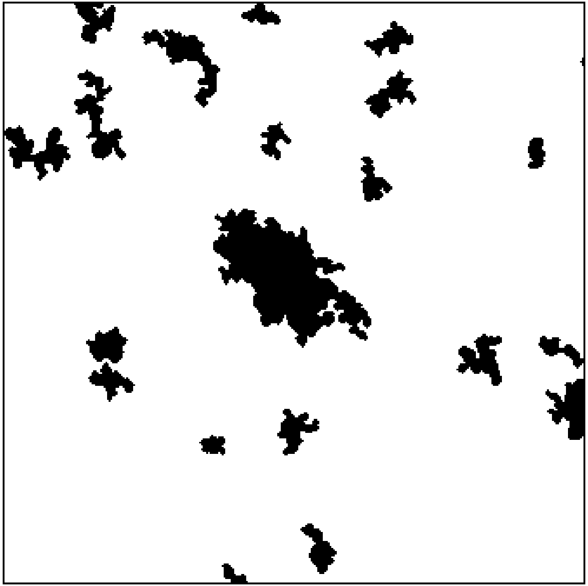
New growth demo: growth quantile: 0.90



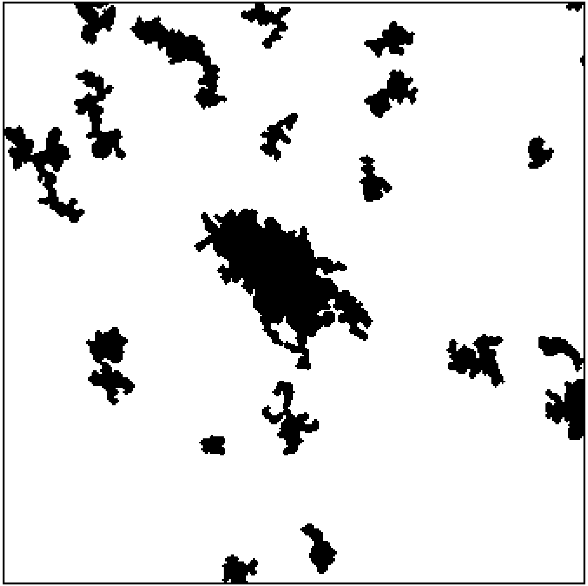
New growth demo: growth quantile: 0.85



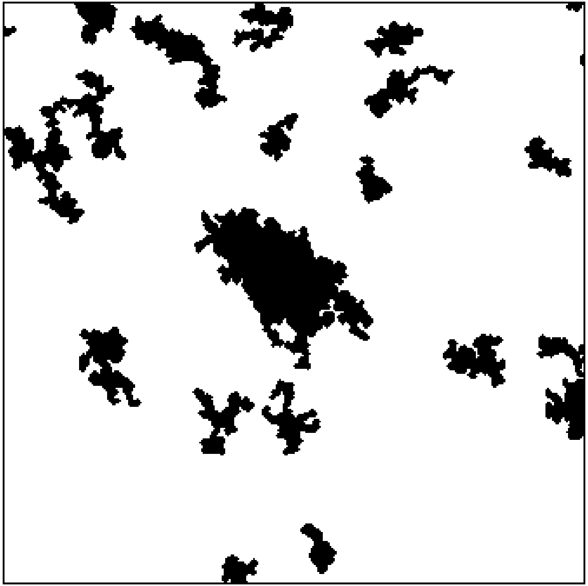
New growth demo: growth quantile: 0.80



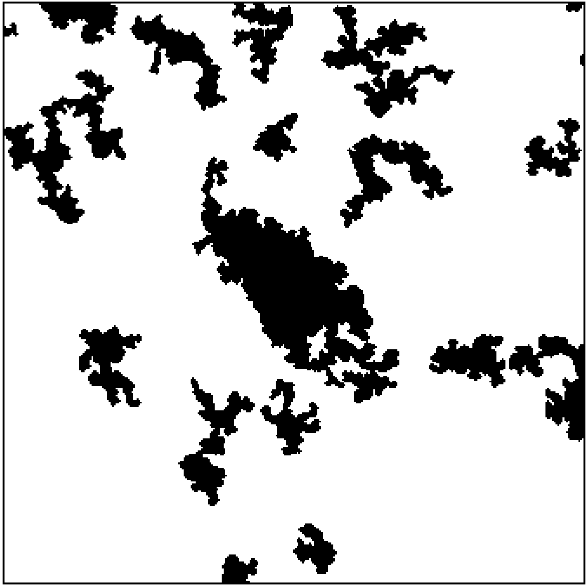
New growth demo: growth quantile: 0.75



New growth demo: growth quantile: 0.70



New growth demo: growth quantile: 0.65



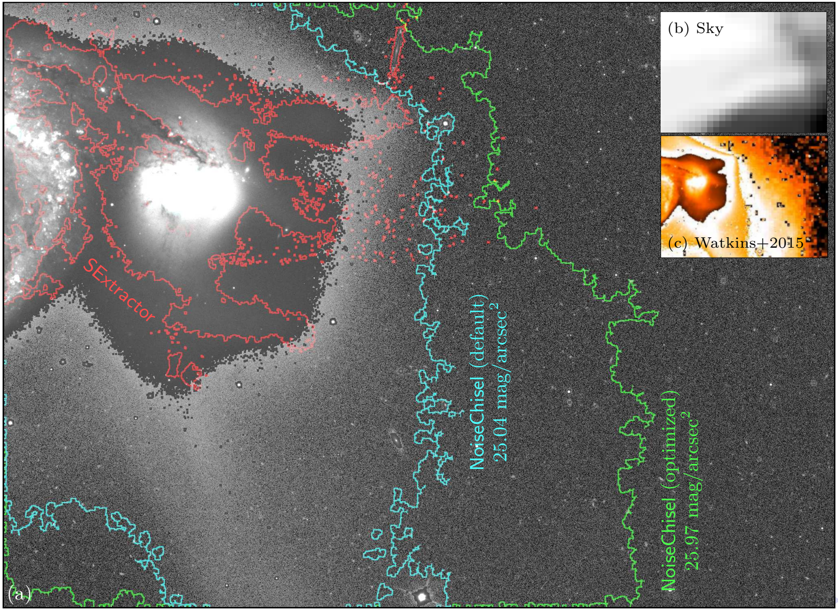
New growth demo: growth quantile: 0.60

M51 with 12-inch telescope (10hr): https://i.redd.it/jfqgpqg0hfk11.jpg

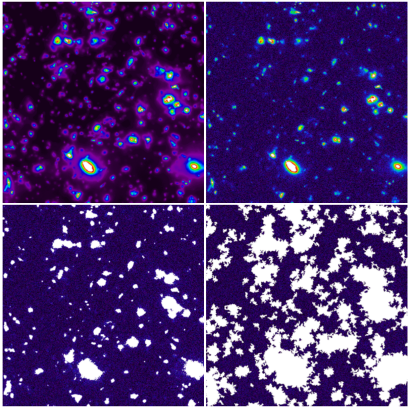
M51 (single exposure SDSS image: $\sim$ 1min, 2.5m telescope)

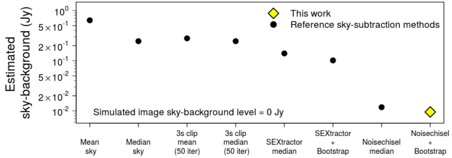
M51 (flux truncated to see the outskirts)

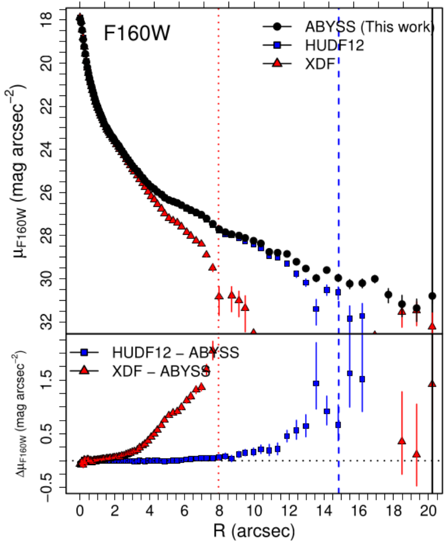
M51 detected pixels (for more, see arXiv:1909.11230)



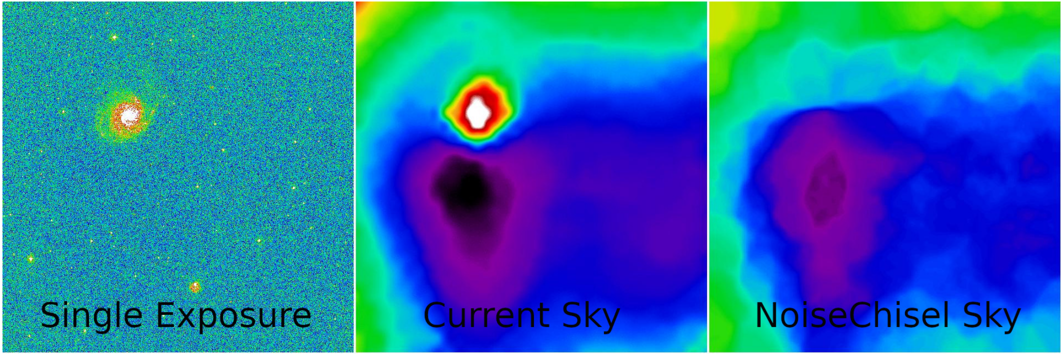
Re-processing HST images with NoiseChisel (Borlaff et al. 2019, arXiv:1810.00002)







This problem is also present in J-PLUS DR2



GTC/OSIRIS spectroscopy of a $z = 6.5$ protocluster (Calvi et al. 2019: arXiv:1908.01827)

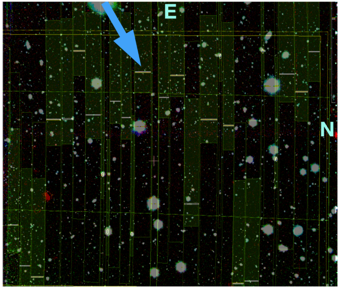
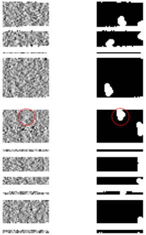


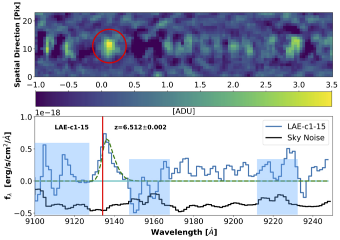
Figure 2. The mask used for the MOS in the OSIRIS/GTC observations. The mask includes 20 science objects and 6 guiding stars. The blue arrow indicates the slit associated with C1-01, one of the confirmed LAEs from Ouchi et al. (2010), included for comparison.



(a) Sky mask on the noise image

(b) Clumps detection

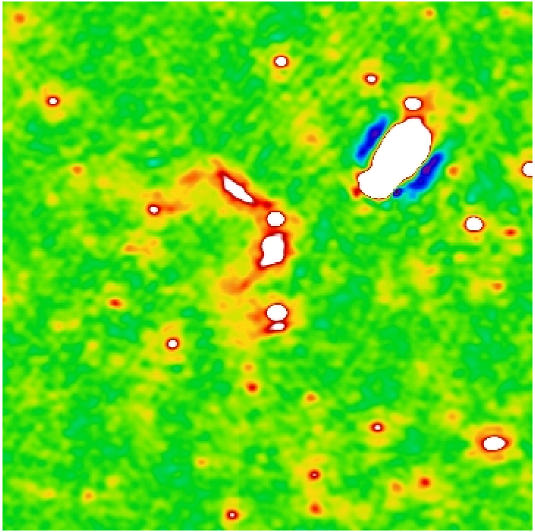
Figure 4. Example of *Noise Chisel* detection process on LAE-c1-15. **Left:** a part of the 2D spectrum with the sky emission lines masked. **Right:** the detected clumps with S/N larger than the preset threshold. The grey-scale shows higher ADU counts in lighter region. The red circles indicate the detected source.



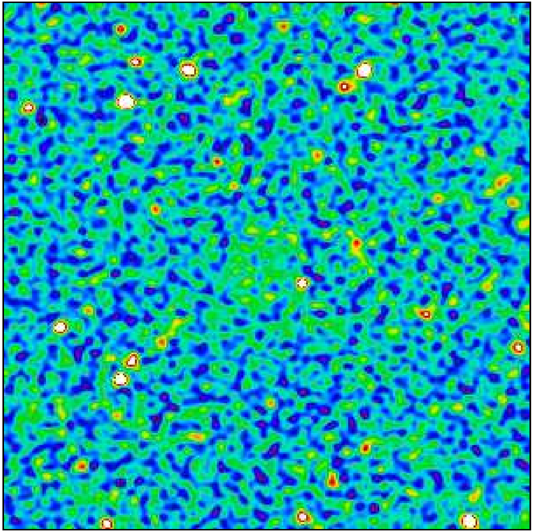
(b) C1-15

EMU examples

Abell S1136 (Thanks to Peter Macgregor)

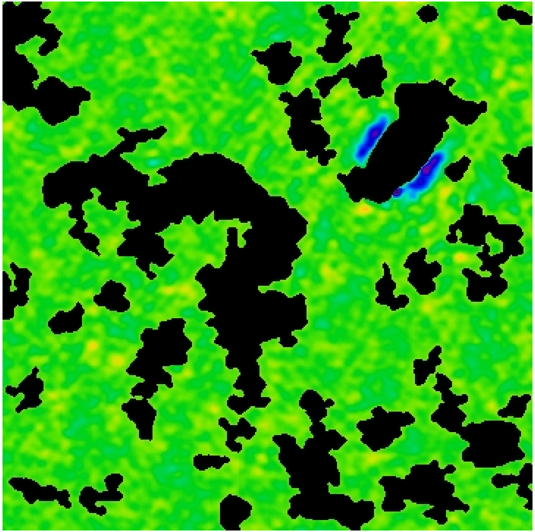


Diffuse blob (Thanks to Rami Alsaberi)

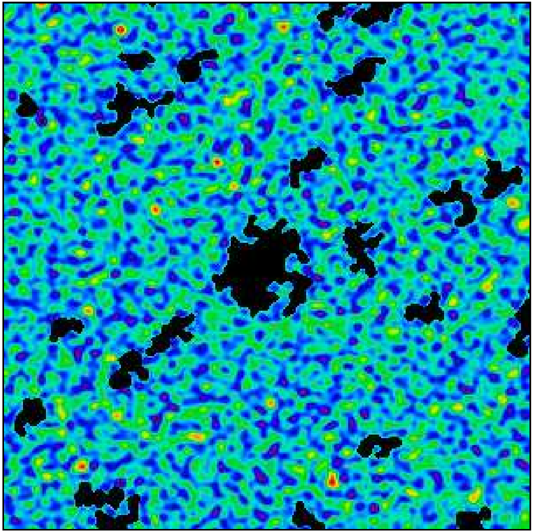


EMU examples (NoiseChisel detections masked)

Abell S1136

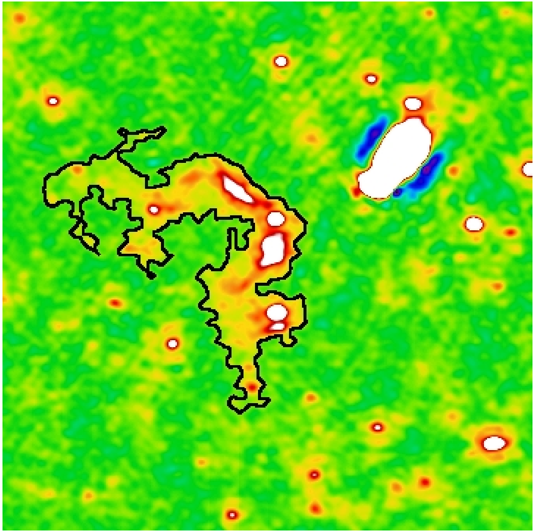


Diffuse blob

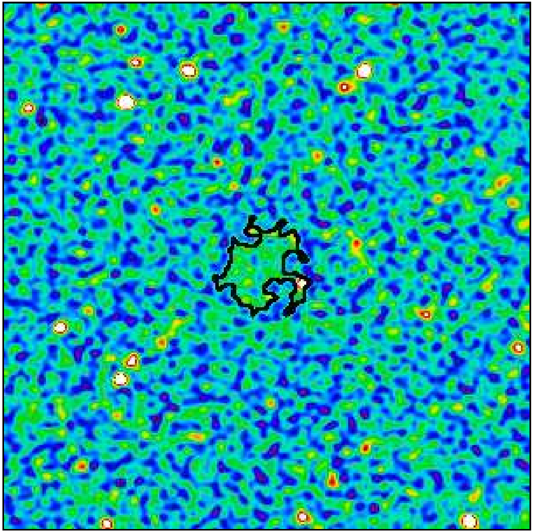


EMU examples (outer edges)

Abell S1136 (outer SB: 6.16×10^{-5} Jy/beam)



Diffuse blob



Detection of diffuse comet tails (from arXiv:2208.02781)

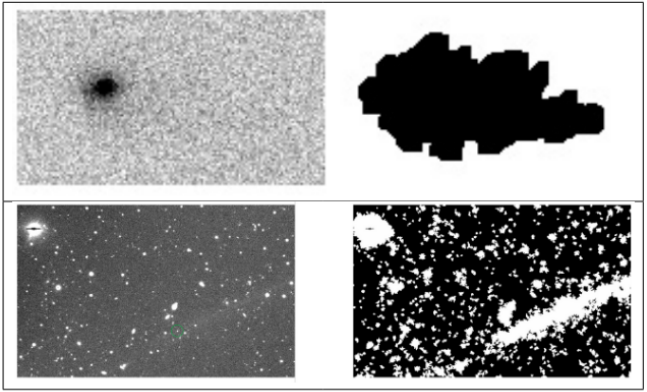
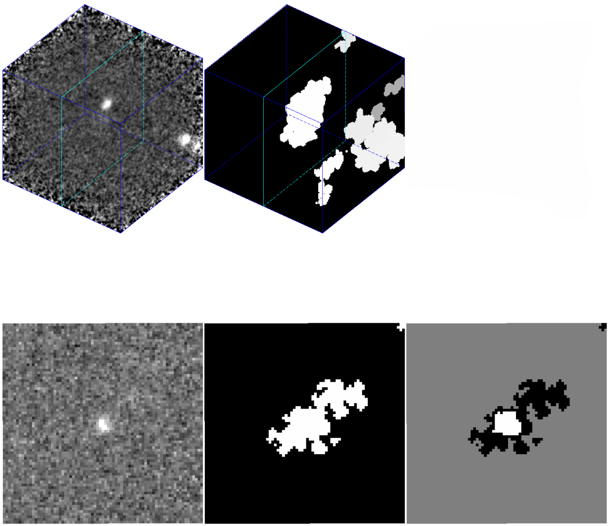
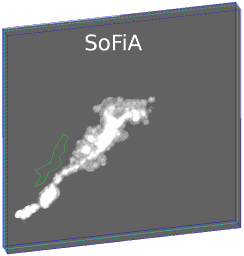


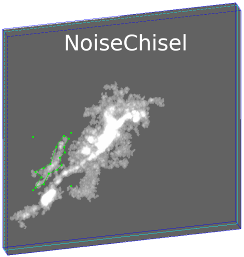
Figure 6. Example results from the tool NoiseChisel (Akhlaghi & Ichikawa 2015; Akhlaghi 2019). Top: Comet 358P/Pan-STARRS with a tail and coma that are difficult to identify in the original image (left), but the resulting NoiseChisel output (right) provides clear evidence of activity. Image courtesy Mohammad Akhlaghi and Henry Hsieh. Bottom: The original image (left) of comet 67P/Churyumov–Gerasimenko (green circle) with a faint tail extending roughly ENE. The NoiseChisel output (right) provides much more contrast, important for manual analysis by humans and automated searches by computers. Image courtesy Agata Rożek.

3D NoiseChisel (a Lyman-α emitter in the MUSE UDF10)

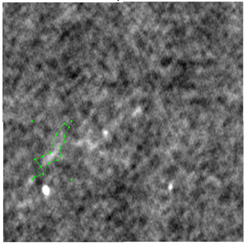


3D NoiseChisel (HI content of HCG 16 from Jones+2019, arXiv:1910.03420)

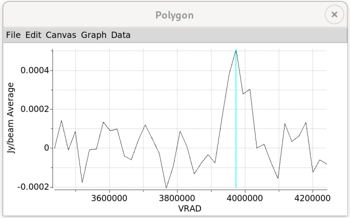




Slice of input cube

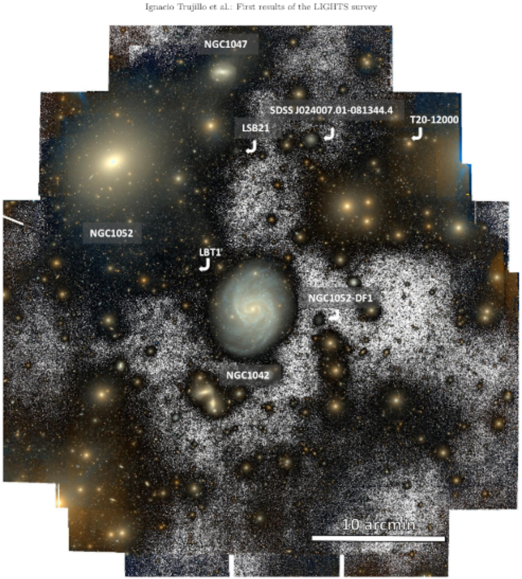


Spectrum of selected region

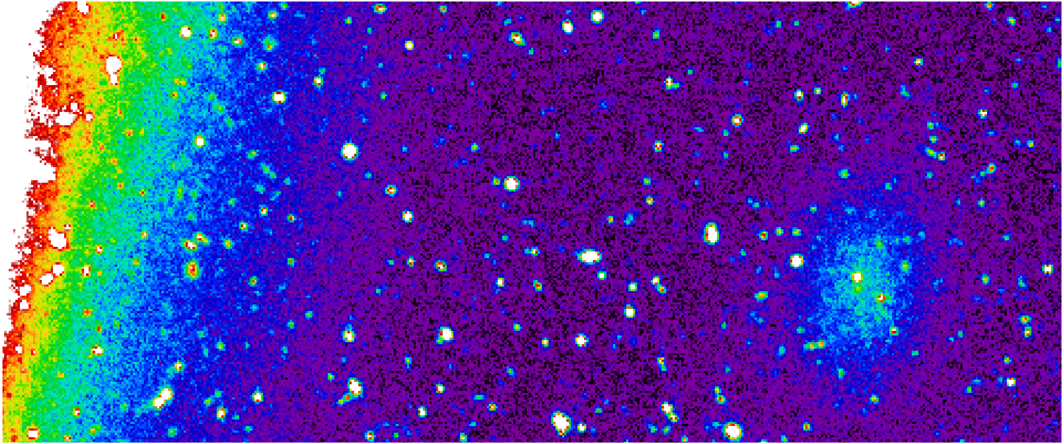


The problem of Segmentation

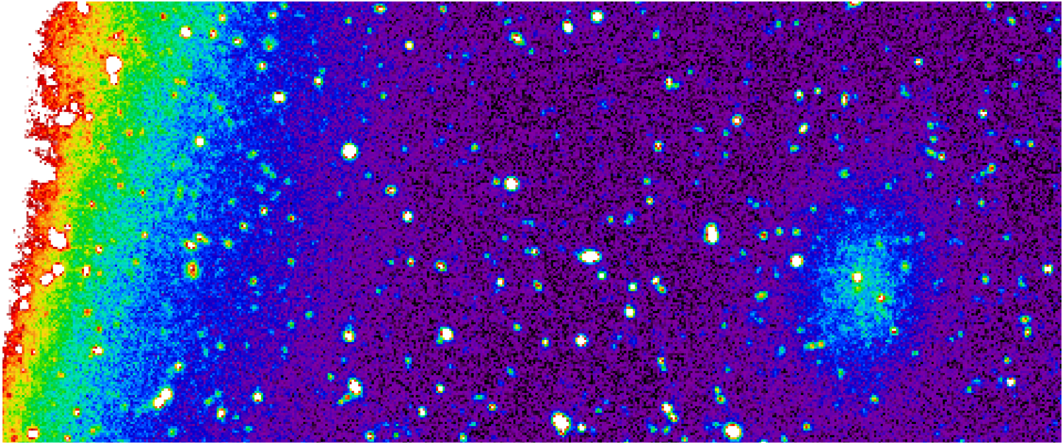
The beautiful view of NGC1042 in LIGHTS



LIGHTS view of NGC1042 (zoom-in: left is NGC1042, right: DF1)

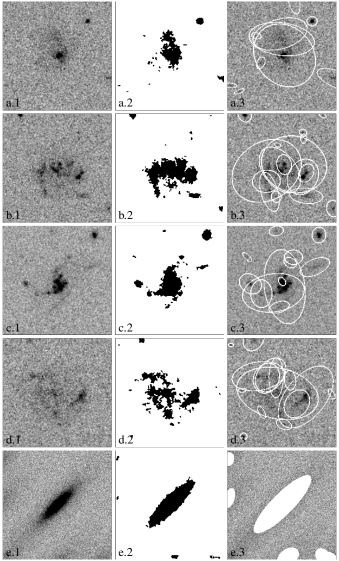


LIGHTS view of NGC1042 (zoom-in: left is NGC1042, right: DF1)



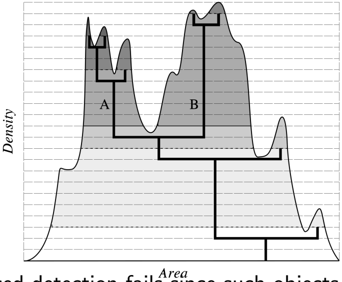
The smaller objects may be globular clusters, star forming regions, faint milky way stars, low mass or high redshift galaxies (all imporatant for science).

Just as a reminder... (of the old method, shown before)



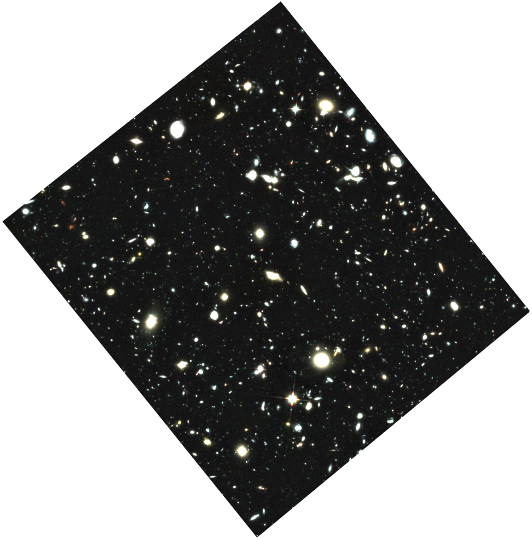
Real observed galaxies:

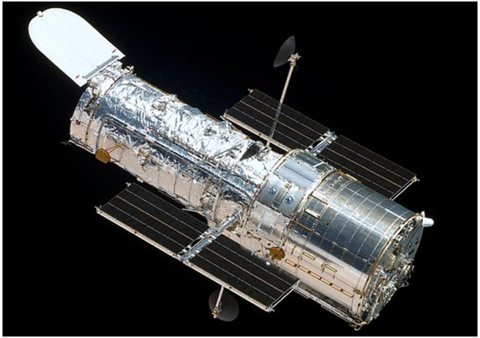
- Are not a clean ellipse.
    - Can be clumpy.
    - Can be diffuse.
    - Can have spiral arms.
    - Can be on the edge of the image.
- SExtractor's deblending uses layers and the parent is used to identify true peaks (systematic biases):



Signal-based detection fails since such objects do not satisfy its *a priori* assumptions.

eXtreme Deep Field (XDF) survey: deepest image of the universe by Hubble Space Telescope (Covering an area 175 times smaller than the Moon!)





Examples from Bacon et al (2017)

Rafelski+2015 (arXiv:1505.01160) use muliple runs of SExtractor on UDF.

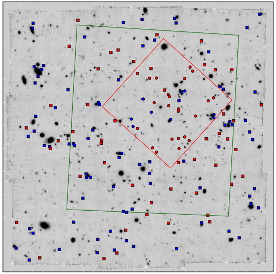
Parameter	Deep	Shallow	Deep Deblend[a]	Shallow Deblend[a]	NUV	NUV Deblend[a]
detect_thresh	1.1	3.5	1.1	3.5	1.0	1.0
analysis_thresh	1.1	3.5	1.1	3.5	1.0	1.0
deblend_nthresh	32	32	8	8	32	8
deblend_mincont	0.01	0.01	0.3	0.3	0.01	0.3
detect_minarea	9	9	9	9	9	9
back_size	128	128	128	128	128	128
back_filtersize	5	5	5	5	5	5
back_photo_thick	26	26	26	26	26	26

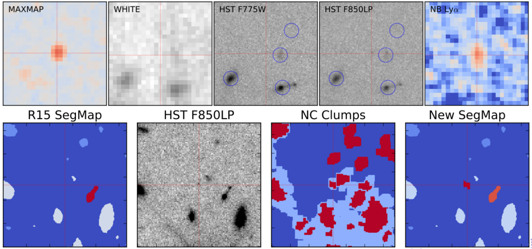
Examples from Bacon et al (2017)

Rafelski+2015 (arXiv:1505.01160) use muliple runs of SExtractor on UDF.

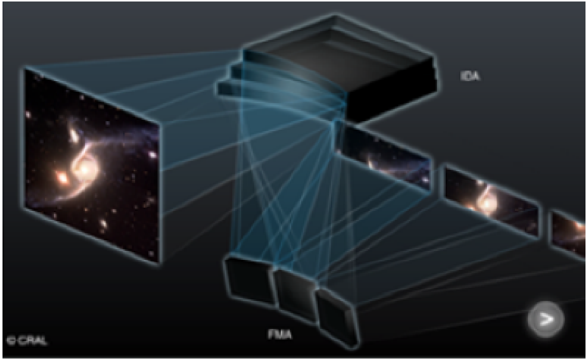
Parameter	Deep	Shallow	Deep Deblend[a]	Shallow Deblend[a]	NUV	NUV Deblend[a]
detect_thresh	1.1	3.5	1.1	3.5	1.0	1.0
analysis_thresh	1.1	3.5	1.1	3.5	1.0	1.0
deblend_nthresh	32	32	8	8	32	8
deblend_mincont	0.01	0.01	0.3	0.3	0.01	0.3
detect_minarea	9	9	9	9	9	9
back_size	128	128	128	128	128	128
back_filtersize	5	5	5	5	5	5
back_photo_thick	26	26	26	26	26	26

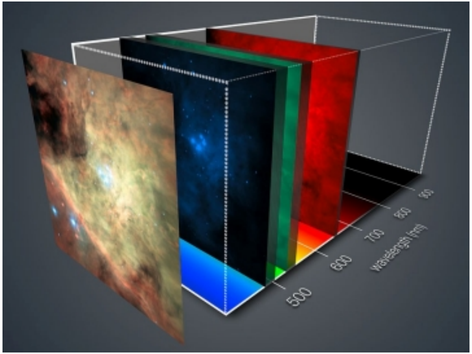
Bacon+2017 (arXiv:1710.03002) found 160 MUSE emission line objects not in Rafelski+2015. 88 (55%) had $> 5\sigma$ flux in a fixed aperture, and 57 (35.62%) were covered by SExtractor's segmentation maps (deblending problem).





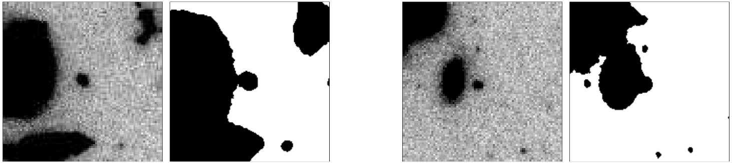
What is MUSE? Its an Integral Field Spectrometer:
Takes a spectrum of each image pixel, *or* takes an image at each wavelength





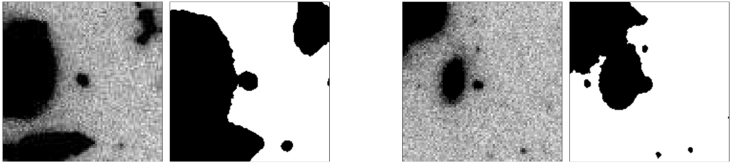
In the segmentation map, but not in the catalog...

Failure to deblend near bright objects (75.44%).

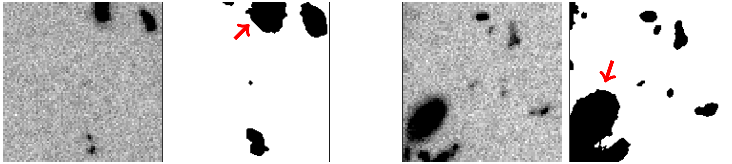


In the segmentation map, but not in the catalog...

Failure to deblend near bright objects (75.44%).

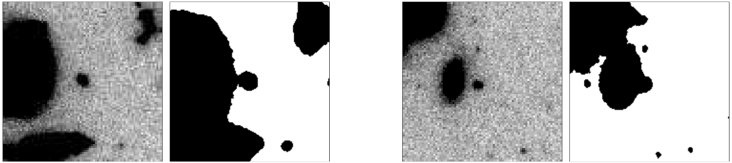


Identify distant objects as one (8.77%).

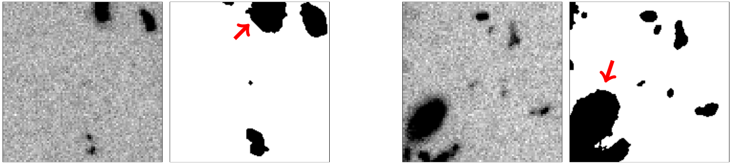


In the segmentation map, but not in the catalog...

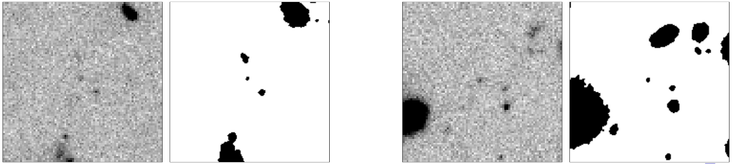
Failure to deblend near bright objects (75.44%).



Identify distant objects as one (8.77%).

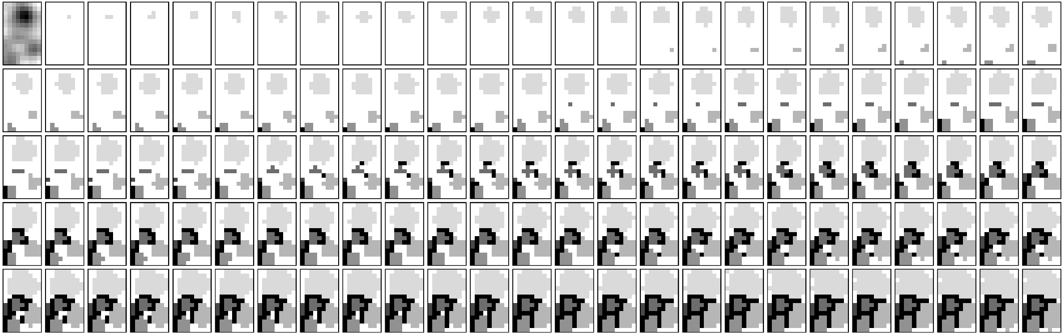


Manually removed from catalog (present in segmentation map. 15.79%).



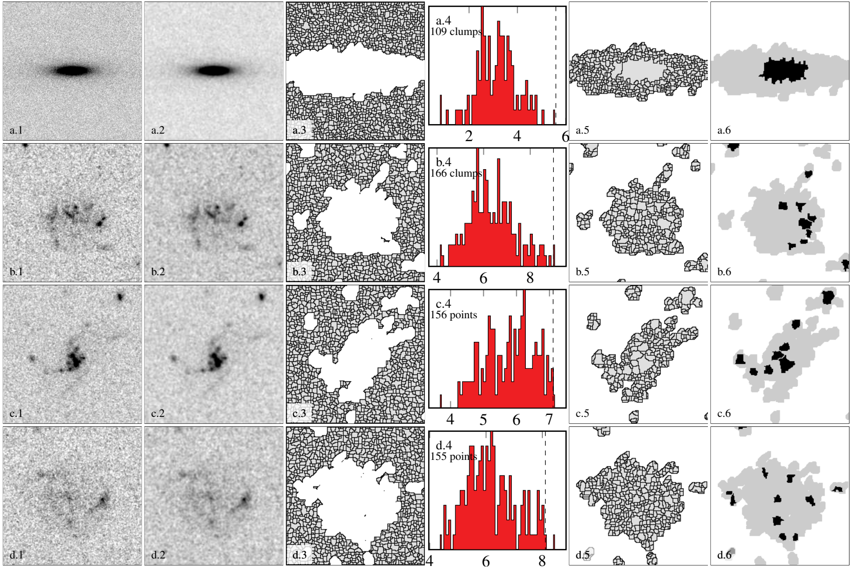
Segment: clumps through the watershed algorithm

A clump is found using the maximum resolution of the convolved image:



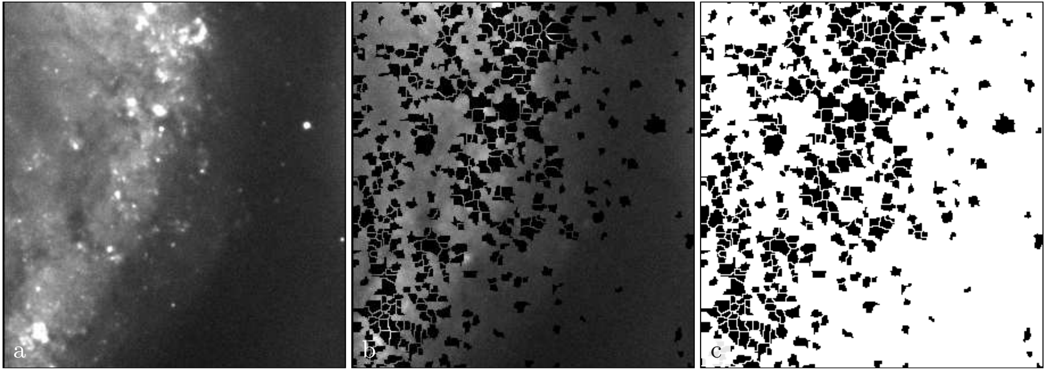
No more layers

Segment: True clumps using ambient noise as reference

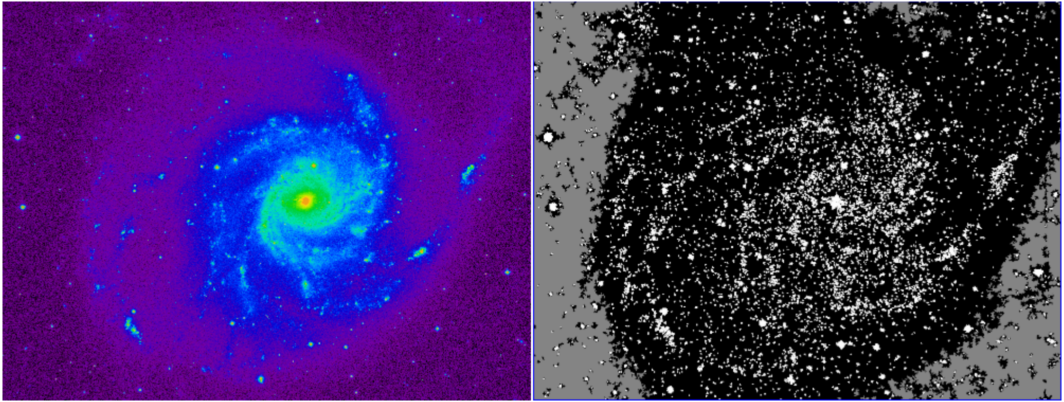


True clumps are found independent of user input.

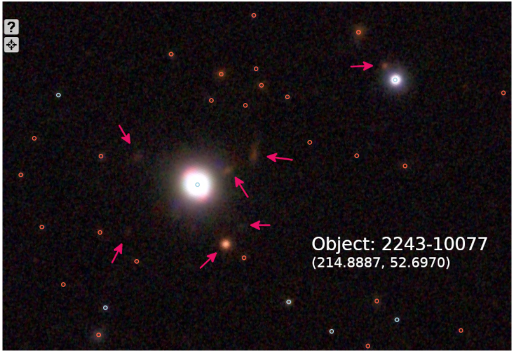
M51 clumps (for more, see arXiv:1909.11230)

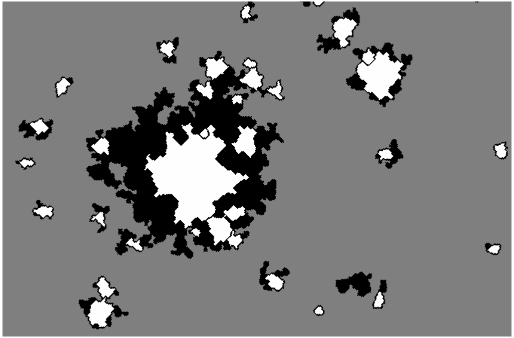


M101 star forming regions labeled as “clumps” on J-PLUS DR2

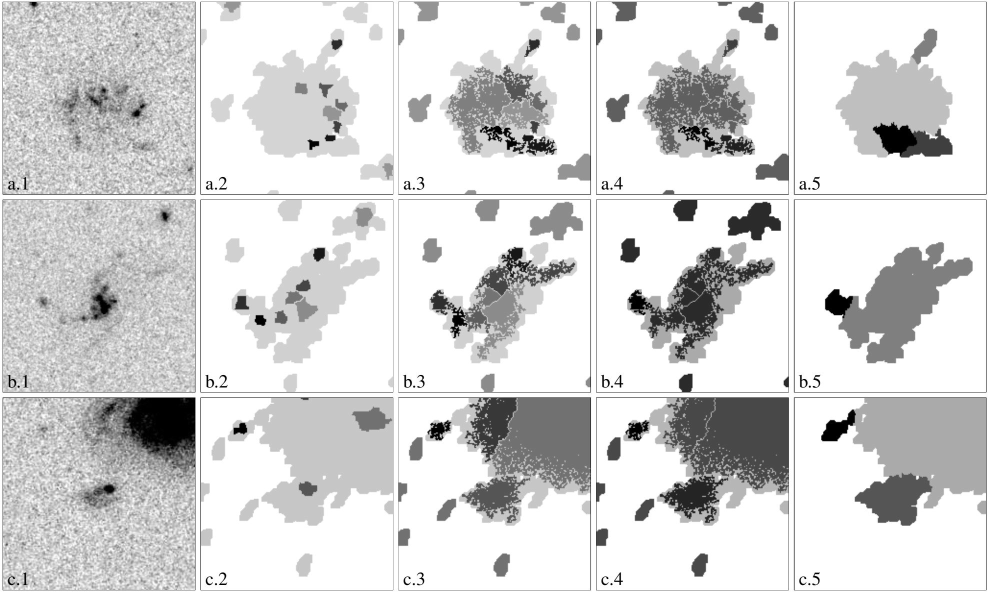


MiniJ-PAS deblending issues

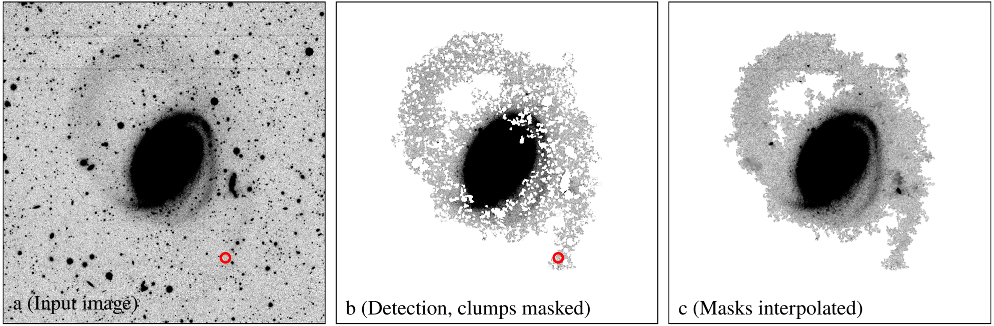




Segment: objects (better done in multiple colors)

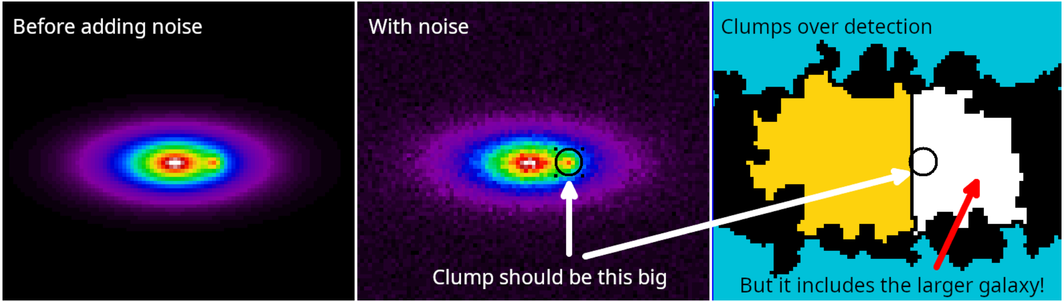


Masking clumps to remove foreground or background objects over diffuse regions (figure from 2023A&A...671A.141M)

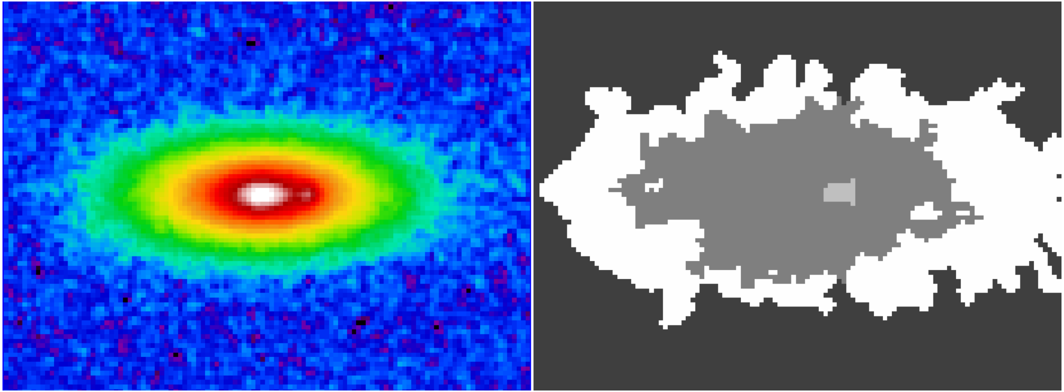


Segment is still being improved (showing bug 62702)

When a small clump falls on the edge of a strong gradient, it steals the flux:



Work is ongoing (have found algorithm to fix problem; currently being tested):

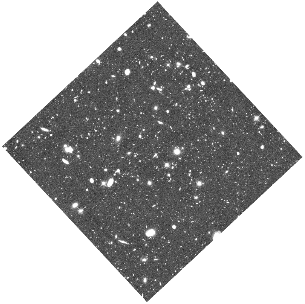


Note how the clump's labels are not confied, and don't cover the wings of the galaxy.

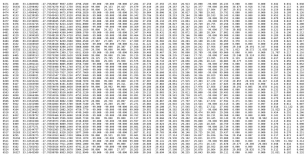
But this is still under development, and evolving!

The problem of Measurements (catalogs)

Data flow in other detection software



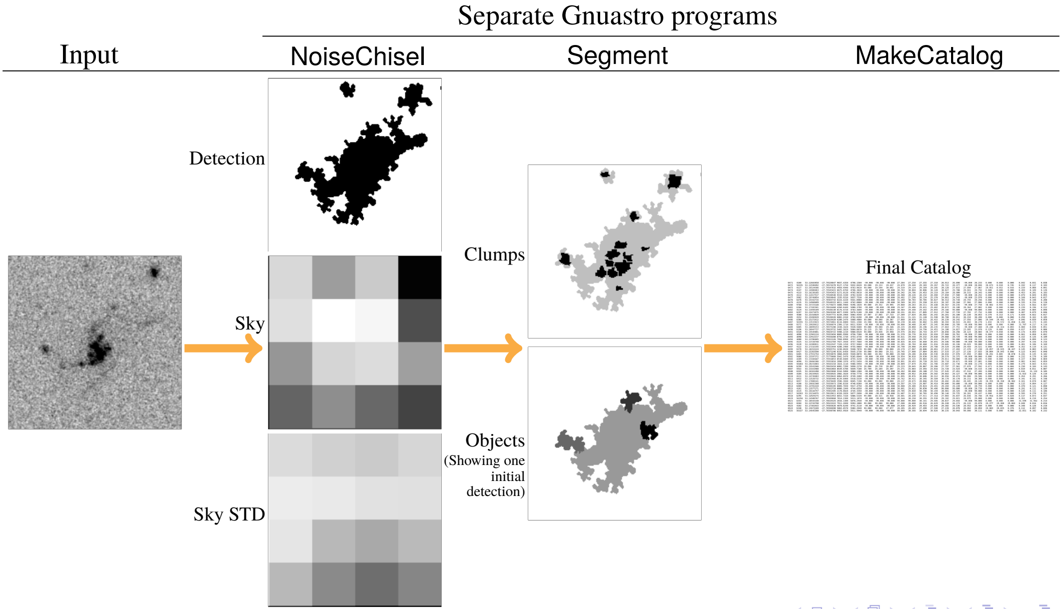
→ Software/Pipeline →



As a result:

- Catalog production is computationally expensive.
- Decreases modularity, or creativity, and thus scientific objectivity.

Data flow: we adopted a *modular* data flow.



MakeCatalog's operation:

Initial objects segmentation map (can also be clump map) ...

MakeCatalog's operation:

Initial objects segmentation map (can also be clump map) ...

... The pixels on each label are measured in parallel (on various threads)...

MakeCatalog's operation:

Initial objects segmentation map (can also be clump map) ...

... The pixels on each label are measured in parallel (on various threads)...

... But not over the whole image, only in the smallest box that covers them.

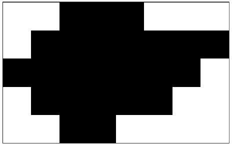
MakeCatalog's operation:

Initial objects segmentation map (can also be clump map) ...

... The pixels on each label are measured in parallel (on various threads)...

... But not over the whole image, only in the smallest box that covers them.

(Future work) These are binary images, for example:



$\rightarrow$

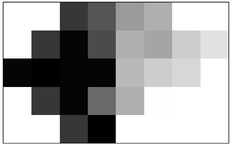
0	0	1	1	1	0	0	0
0	1	1	1	1	1	1	1
1	1	1	1	1	1	1	0
0	1	1	1	1	1	0	0
0	0	1	1	0	0	0	0

So, only one bit is necessary for each pixel.

Deblending and matching

Deblending (when necessary) can be defined as
1-byte (256 layers) or 2-byte (65536 layers) integers.

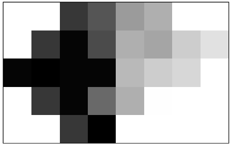
0	0	200	170	100	80	0	0
0	200	250	180	80	90	50	30
250	255	250	250	70	50	40	0
0	200	250	150	80	1	0	0
0	0	200	255	0	0	0	0



Deblending and matching

Deblending (when necessary) can be defined as
1-byte (256 layers) or 2-byte (65536 layers) integers.

0	0	200	170	100	80	0	0
0	200	250	180	80	90	50	30
250	255	250	250	70	50	40	0
0	200	250	150	80	1	0	0
0	0	200	255	0	0	0	0

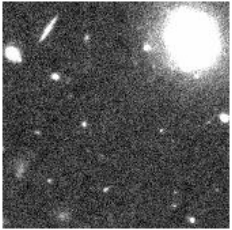


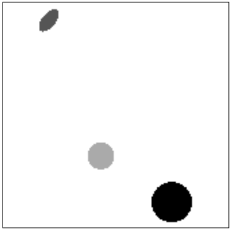
In a similar way, the pixel maps can be warped and/or convolved to match images with other pixel or spatial resolutions (from other surveys).

Aperture photometry

Aperture photometry only needs detection for the Sky and σ_{Sky}.

In Gnuastro, MakeProfiles is in charge of building profiles (apertures in this case) on an image.





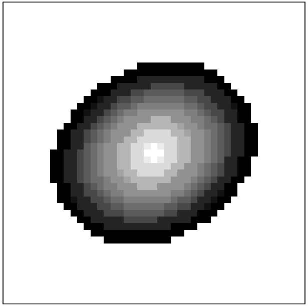
Curve of growth or radial profile

Curves of growth (or radial profile) segmentation maps (elliptical annuli) can also be created easily with MakeProfiles.

MakeProfiles input (to define each annulus):

```
1 22 22 5 1 0 30 0.8 1 1
2 22 22 5 2 0 30 0.8 2 1
3 22 22 5 3 0 30 0.8 3 1
```

(Columns are: ID, X, Y, function, radius, func. param, PA, axis ratio, value, truncation)



Parent software: GNU Astronomy Utilities

GNU Astronomy Utilities (Gnuastro): NoiseChisel's parent software

- Gnuastro is a large collection of programs and libraries for astronomical data manipulation and analysis.
- Programs are run directly on the command-line with no mini-environment (unlike Python or IRAF).
- They are thus fast and easy to combine with other command-line programs. For example:

```
$ astnoisechisel image.fits
$ asttable binary-table.fits | awk '$4>10'
$ asttable binary-table.fits --range=SN,10,inf
```

- The Gnuastro experience is thus very familiar and similar to basic Unix-like command-line tools (e.g., `ls` and `cat`).
- Gnuastro has a complete and up-to-date manual (like many GNU software).

> GNU (+35 years old) is one of the oldest free or open-source software communities. For the GNU label, a software has to be refereed by the GNU Evaluation Committee, and has to abide by the time-tested GNU Coding Standards.

Current list of Gnuastro programs (sorted alphabetically)

- Arithmetic: arithmetic operations on multiple datasets (images).
- BuildProgram: Compile, link and run C/C++ code with Gnuastro's library.
- ConvertType: FITS images to and from text, JPEG, TIFF, EPS or PDF.
- Convolve: Convolve data with a given kernel.
- CosmicCalculator: Cosmological calculations.
- Crop: Crop region(s) from an image and stitch several images if necessary.
- Fits: View and manipulate FITS file extensions and header keywords.
- MakeCatalog: Make catalog of labeled images, see arXiv:1611.06387.
- MakeProfiles: Make mock 2D profiles (e.g., Sérsic, Gaussian, Moffat).
- Match: Match two given catalogs in 1D or 2D within an aperture.
- NoiseChisel: Detect signal in noise, see arXiv:1505.01664 & arXiv:1909.11230.
- Query: Query online data bases, for example VizieR, Gaia, NED and etc.
- Segment: Segment detections, see arXiv:1909.11230
- Statistics: Statistical calculations on the input dataset.
- Table: Read/write FITS (binary or ASCII) or plain text tables.
- Warp: Warp image to new pixel grid.

Summary

- *NoiseChisel* is a program to detect signal very deep into the noise.
- *Segment* is a program to segment the detections were necesary (into *objects* and *clumps*).
- *MakeCatalog* is a program to generate a catalog from the output of NoiseChisel or Segment.
- GNU Astronomy Utilities (Gnuastro) is a highly robust and refereed set of tools containing the programs above (along with many other useful programs and libraries) that is guaranteed to be free to use for the future.
- Gnuastro's webpage: `https://www.gnu.org/software/gnuastro`
- Video tutorial @ADASS2021: `https://www.youtube.com/watch?v=iukkBV-EBbM`
- Matrix Chat room: `#gnuastro:openastronomy.org`
- These slides are available at: `http://akhlaghi.org/pdf/noisechisel.pdf`
- Contact for questions or bugs: `help-gnuastro@gnu.org` and `bug-gnuastro@gnu.org`.