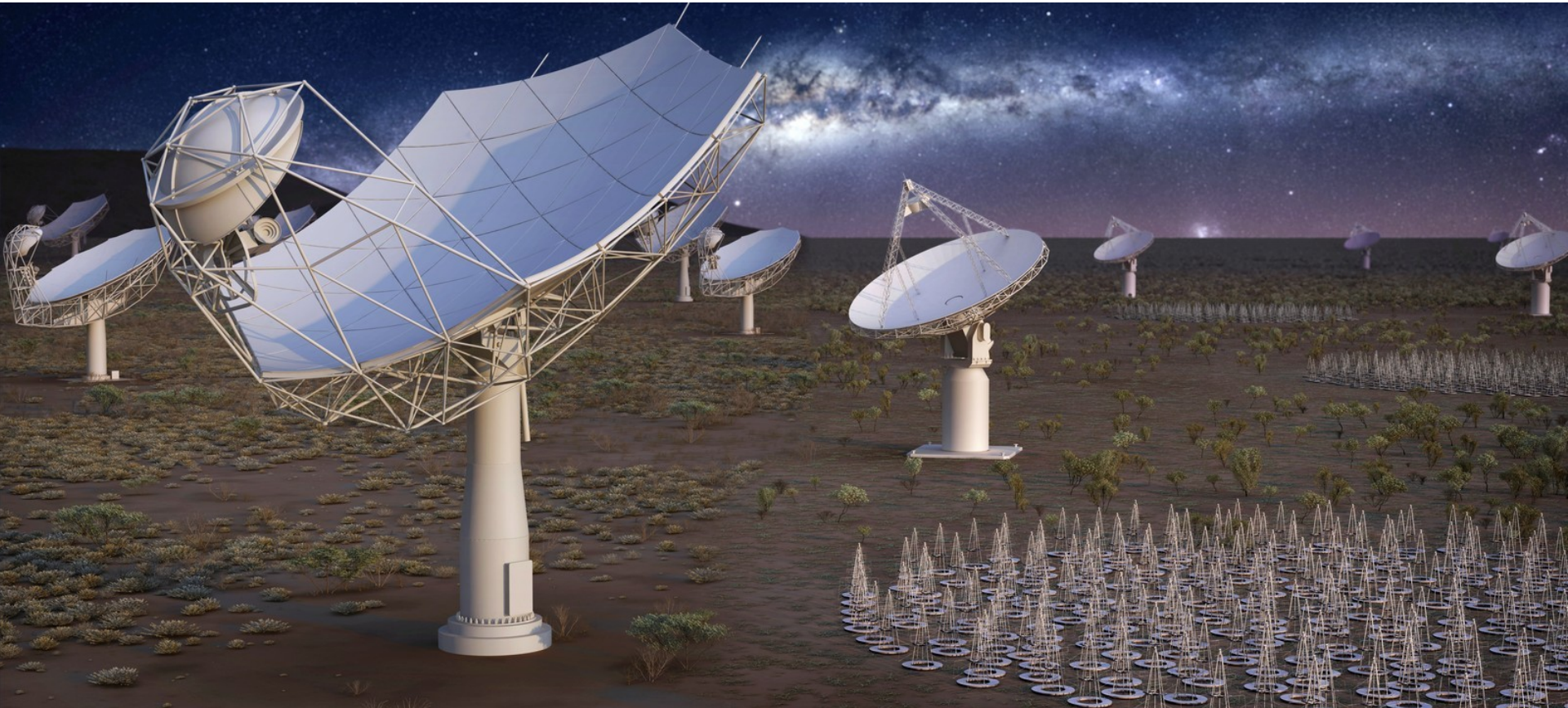


Can Crowdsourcing be Replaced by GPUs?

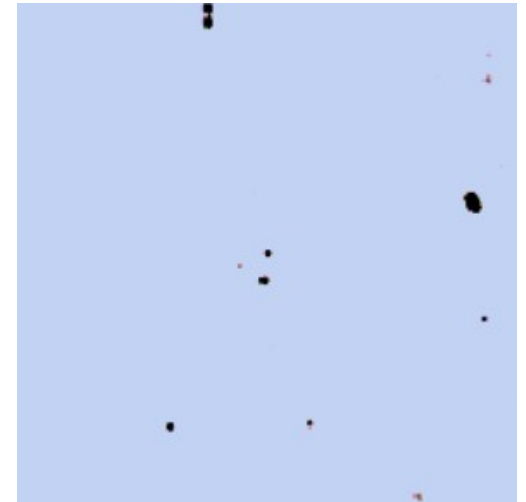
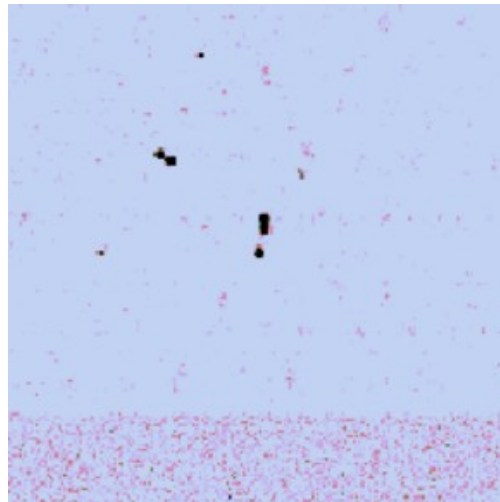
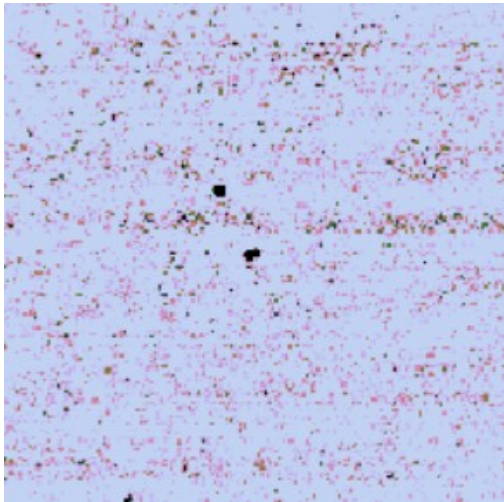
Erica Hopkins

The Future: Upcoming Radio Surveys

- Square Kilometer Array (SKA)

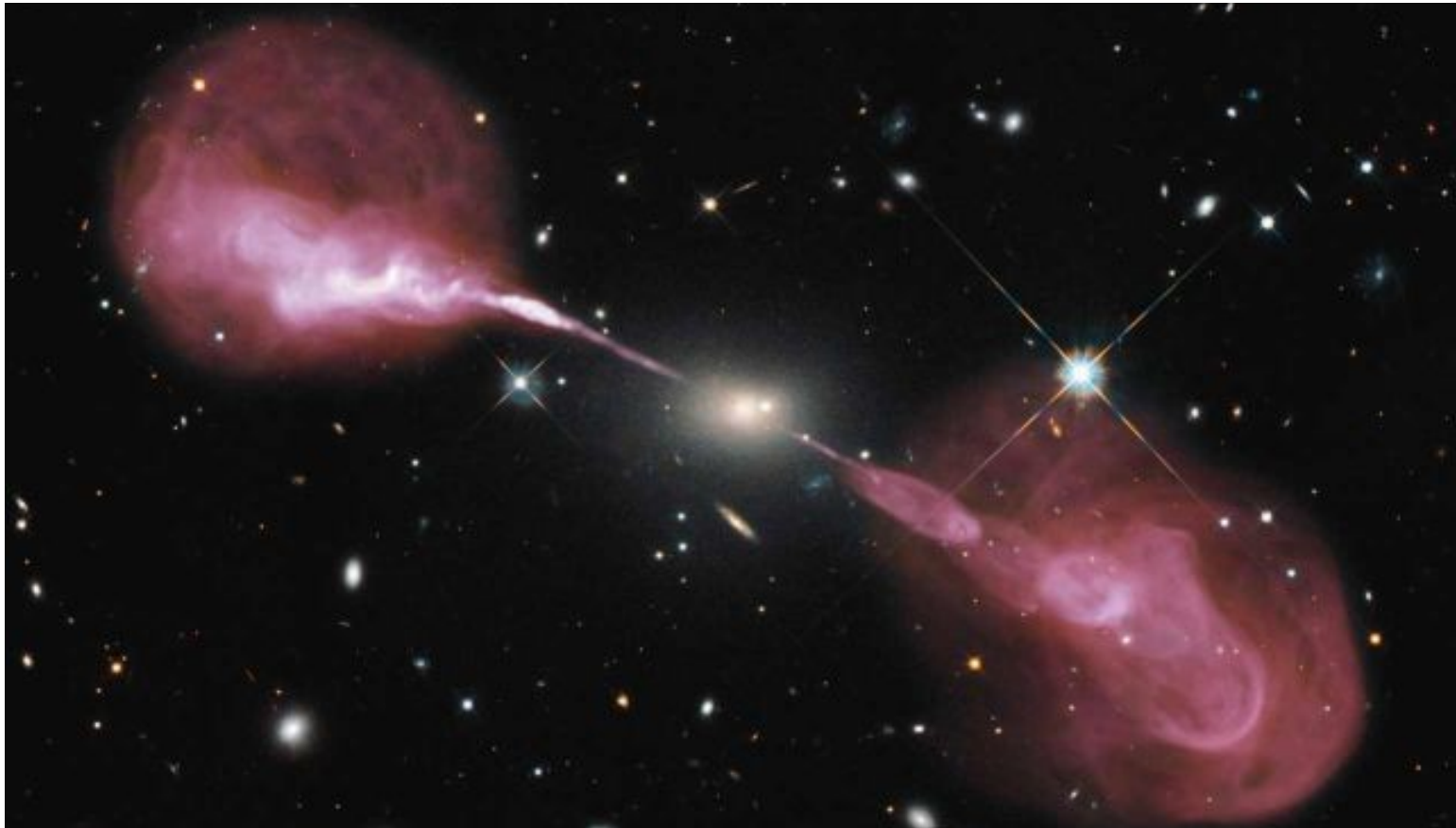


Are these radio sources part of the same galaxy?



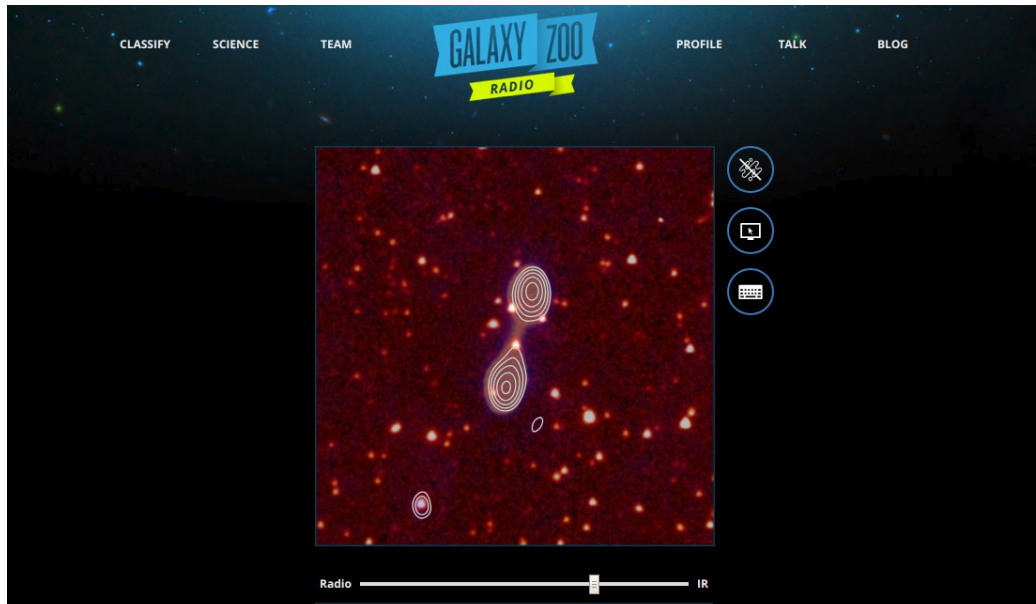
How can we answer this question?

- Radio emission is extended
- Central part of the galaxy detectable in the IR



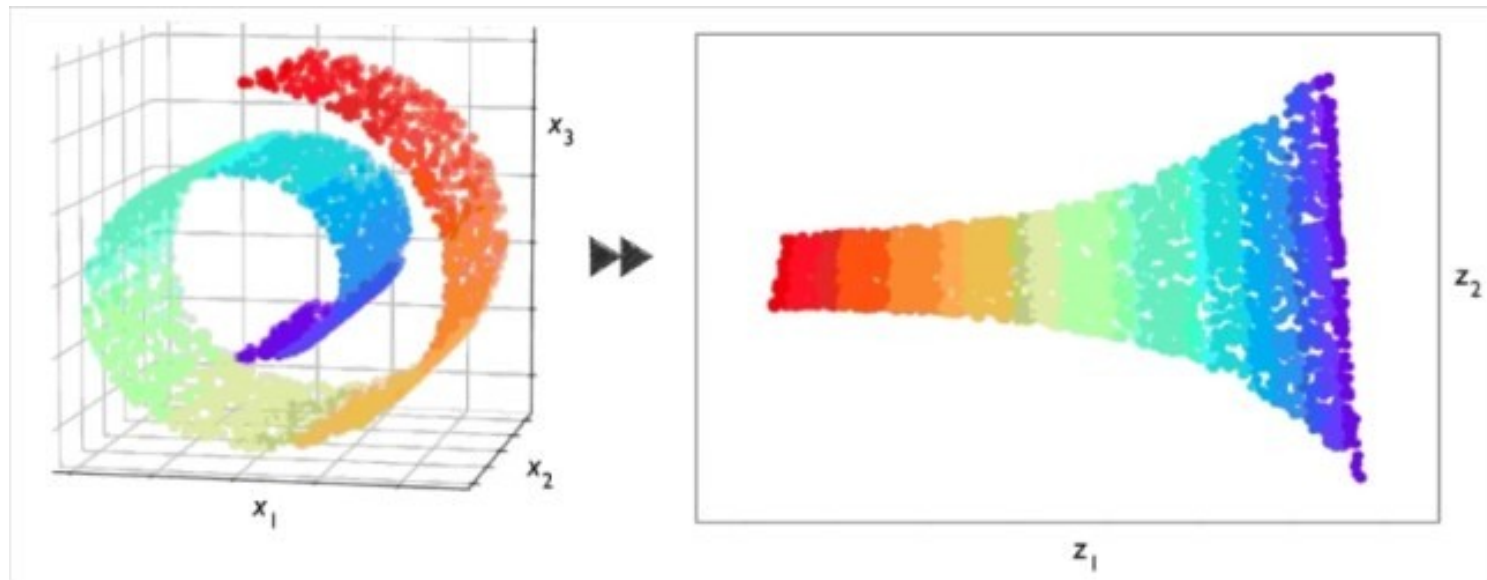
Why do we need machine learning?

- Technically... we don't
- Radio GalaxyZoo classified over 170k sources
 - ... over 4 years
 - ... thanks to a team of scientists
 - ... and over 10,000 volunteers

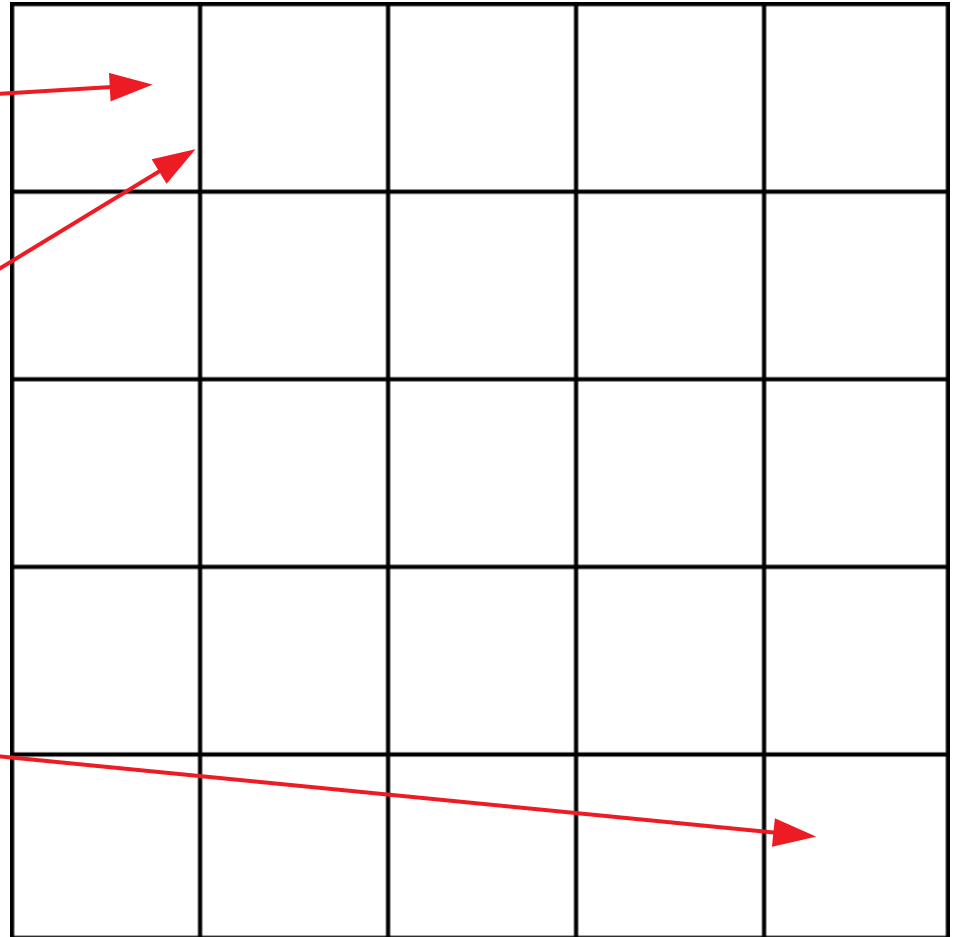
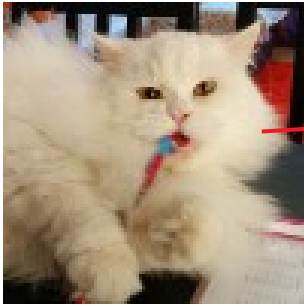


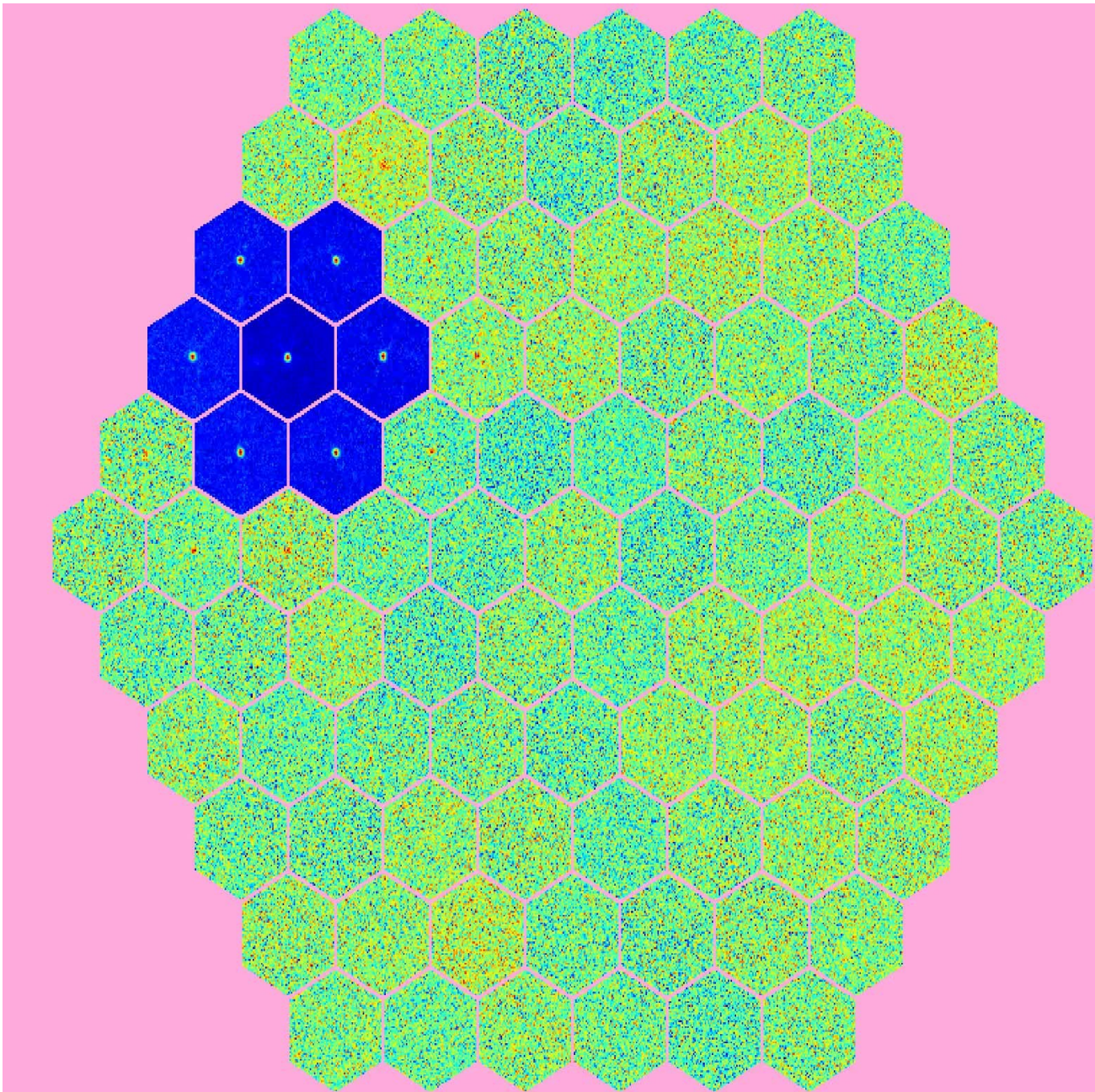
How can this be done more efficiently?

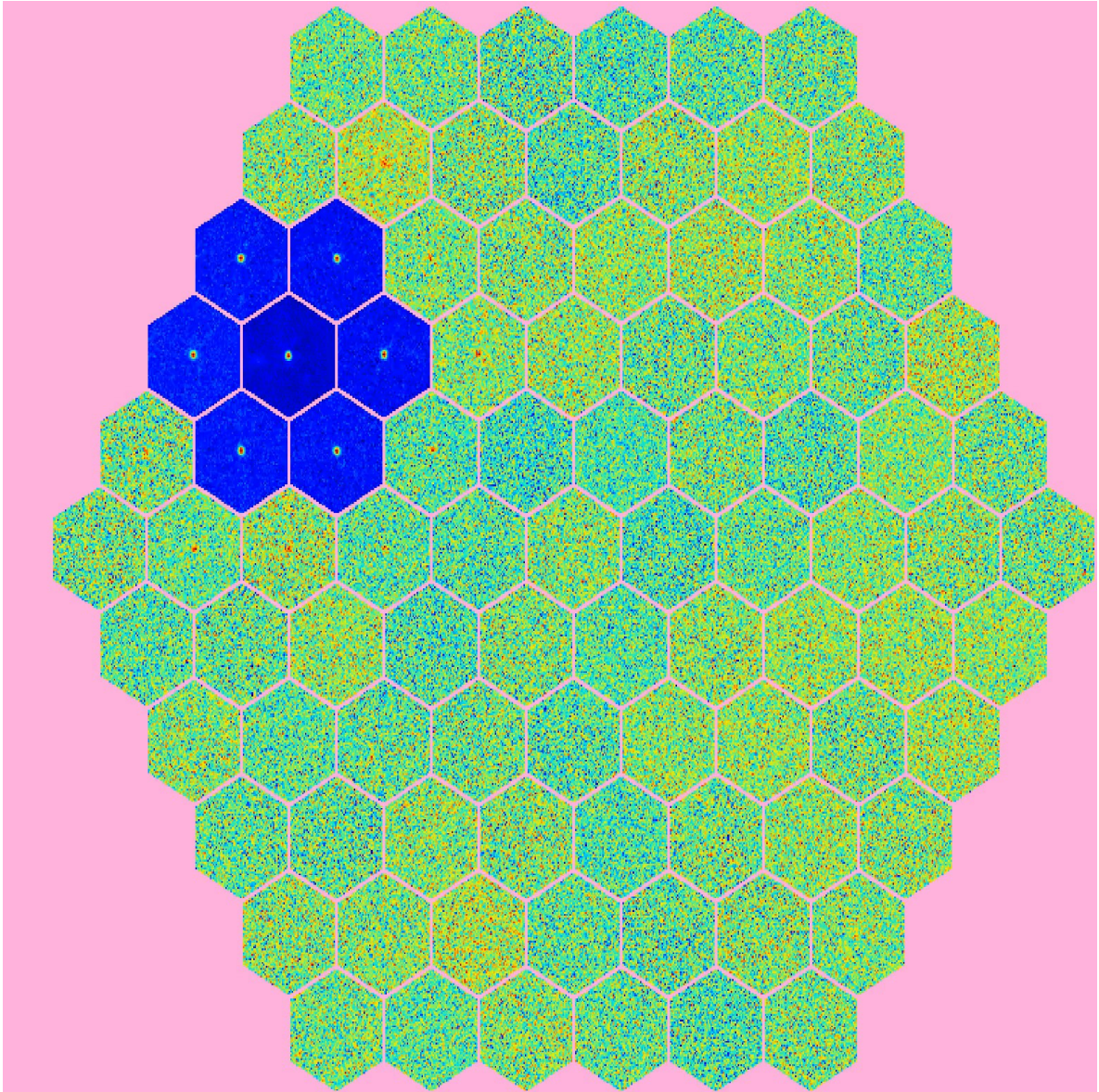
- Use dimensionality reduction technique
 - Project from high-dimensional (8,100d) to low-dimensional space (2d)
 - Preserve similarity relation
 - place similar objects next to each other



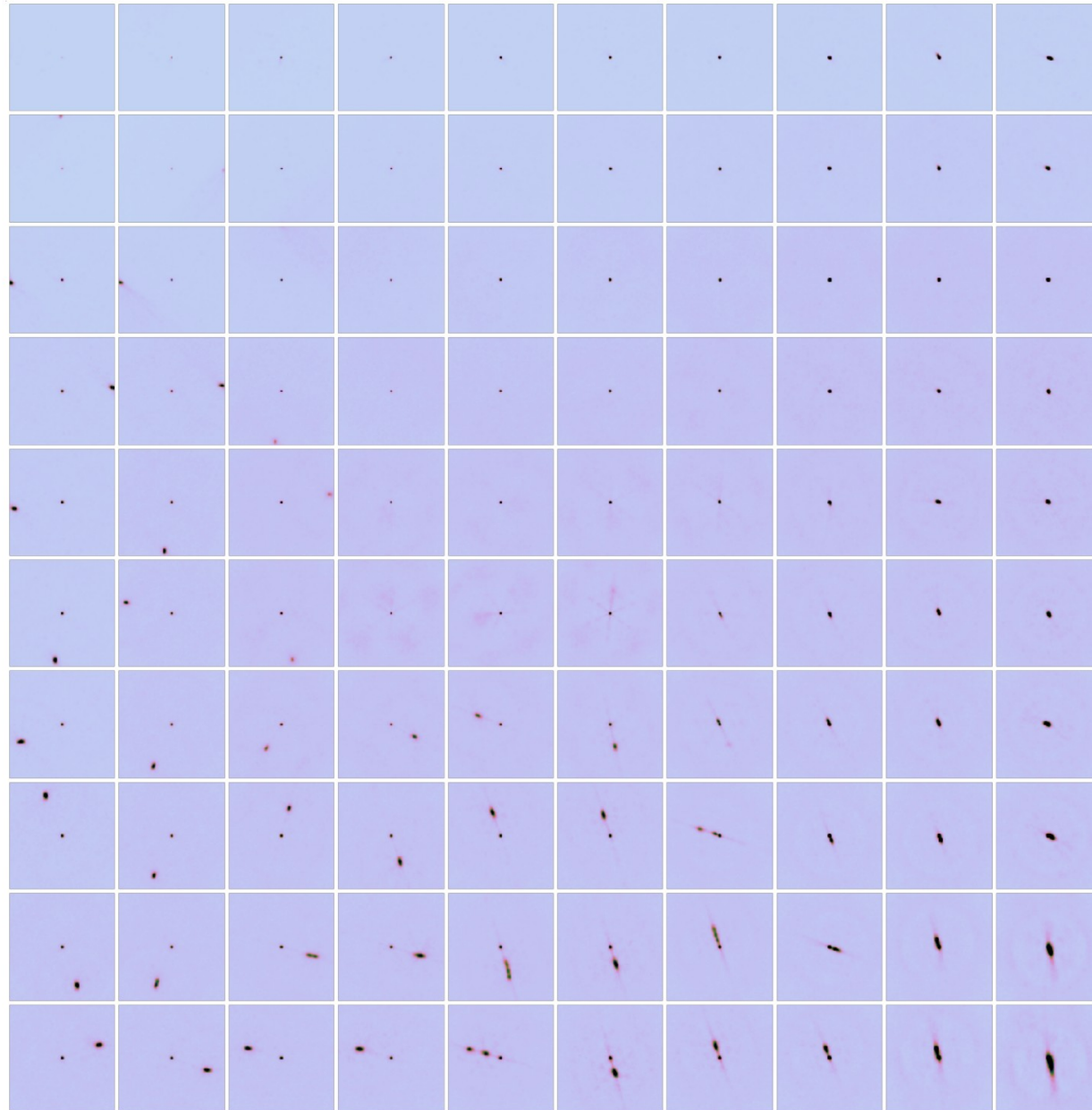
How Morphologies are Projected



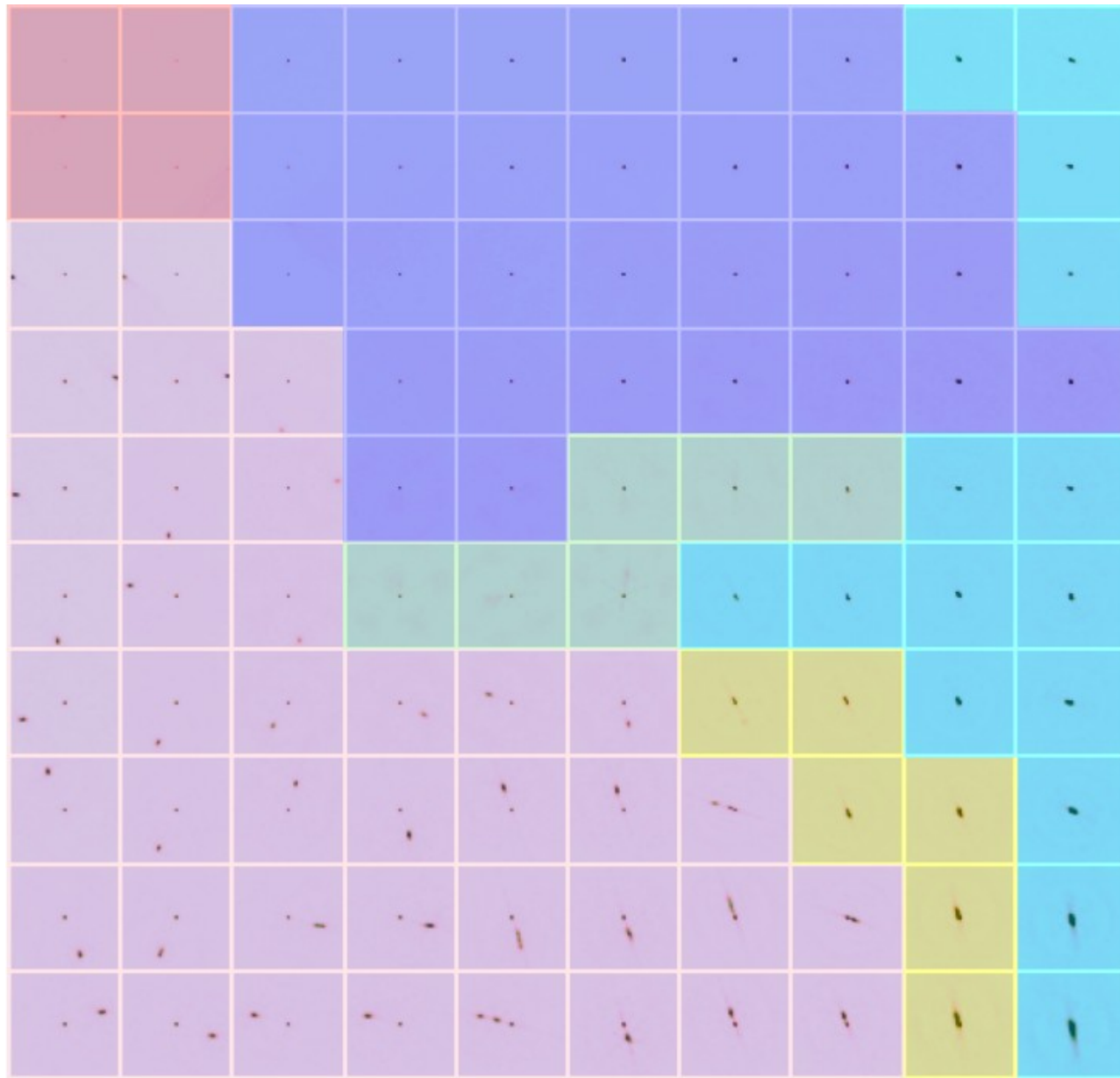




Radio Results



Radio Results



dim noisy point extended concatenated detached

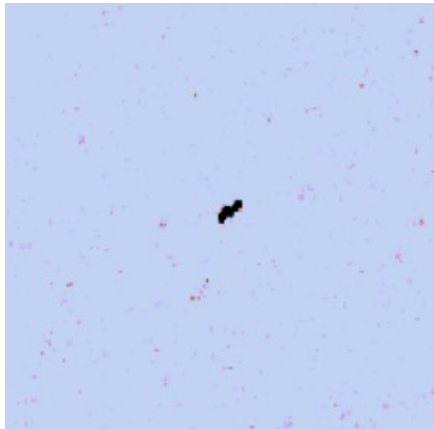


Transferring Expert Knowledge

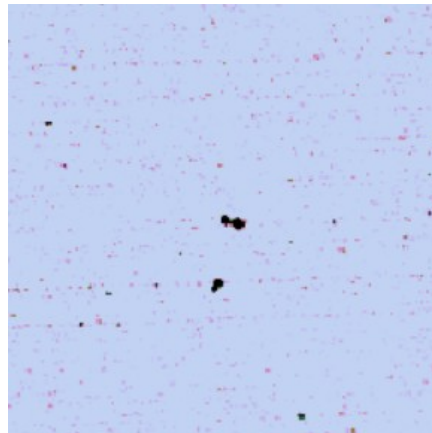
- First approach:
 - Use the best matching prototype only
 - Poor performance
- Second approach:
 - Pseudo-probabilistic method to determine class
 - Normalized by class distribution



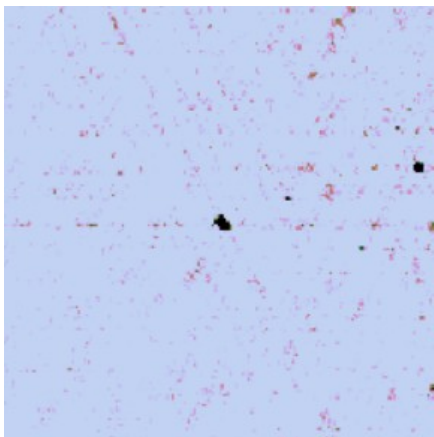
Catalog



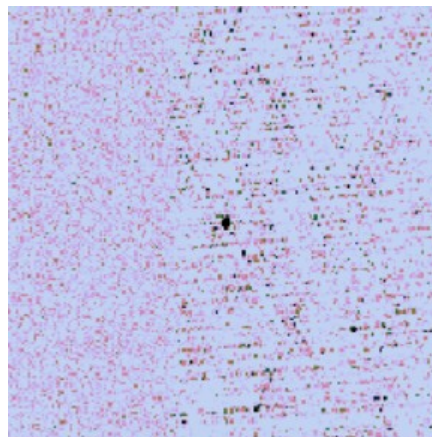
Dim: 0%
Noisy: 0%
Point: 5%
Extended: 80%
Concatenated: 15%
Detached: 0%
IR source: 72%
Outlier Score: 9



Dim: 0%
Noisy: 2%
Point: 4%
Extended: 5%
Concatenated: 70%
Detached: 19%
IR source: N/A
Outlier Score: 31



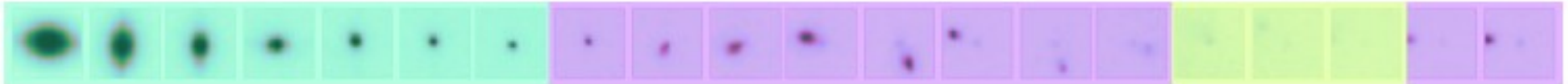
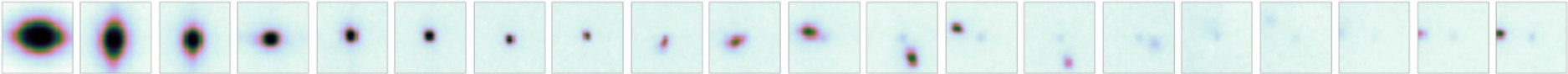
Dim: 0%
Noisy: 12%
Point: 40%
Extended: 46%
Concatenated: 1%
Detached: 1%
IR source: N/A
Outlier Score: 27



Dim: 2%
Noisy: 31%
Point: 50%
Extended: 6%
Concatenated: 5%
Detached: 7%
IR source: 100%
Outlier Score: 11

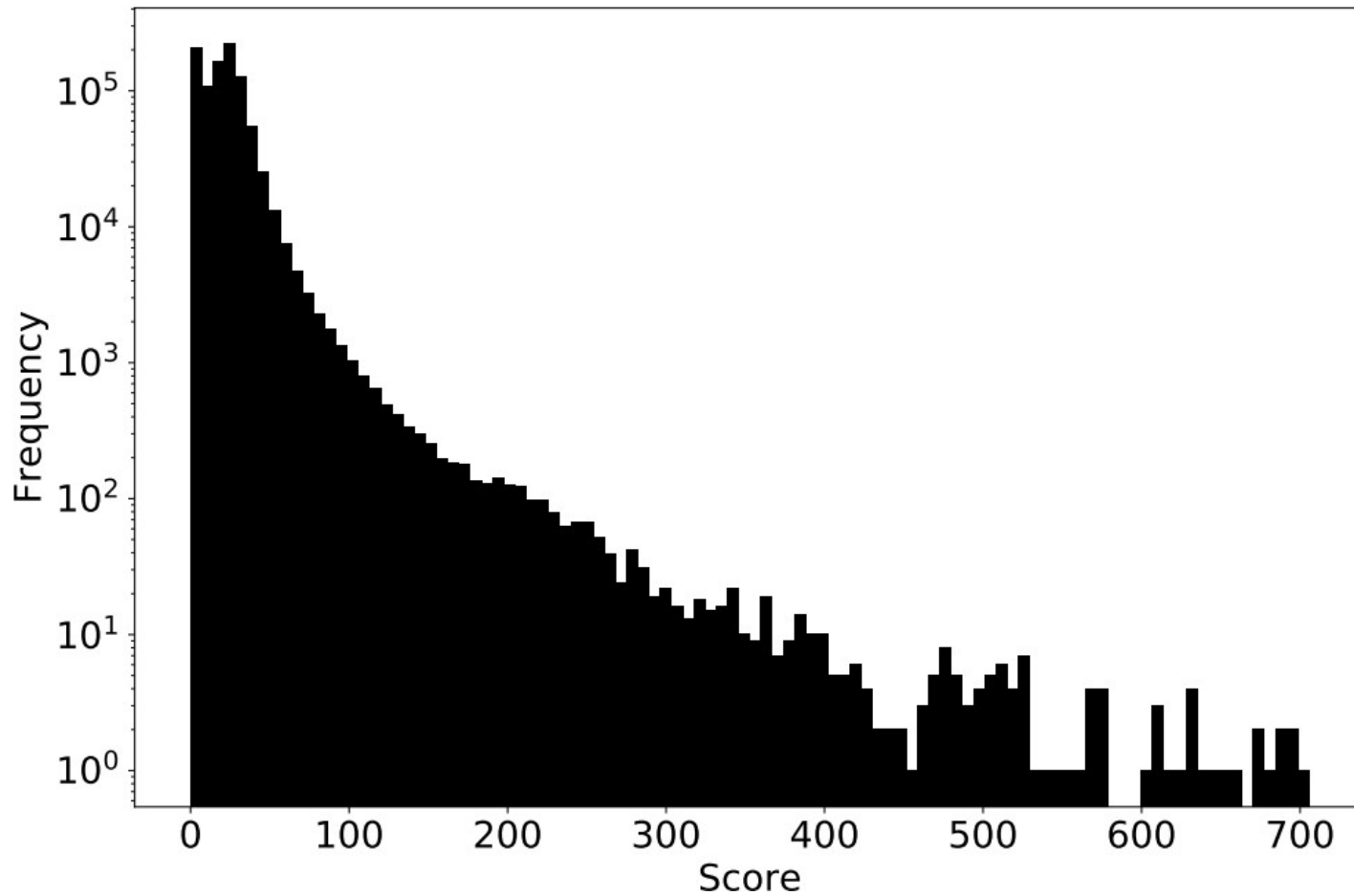


IR Source

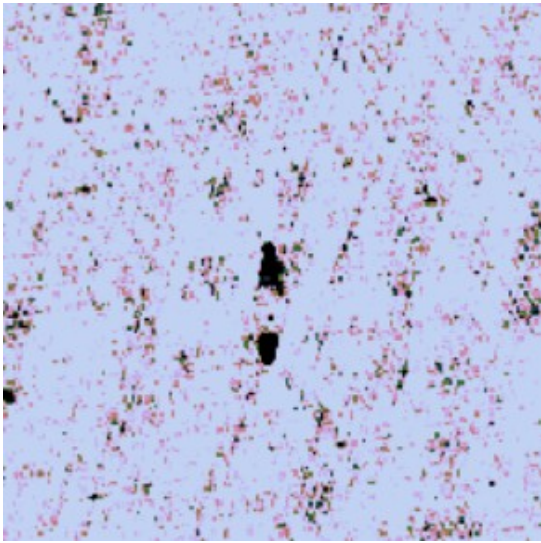
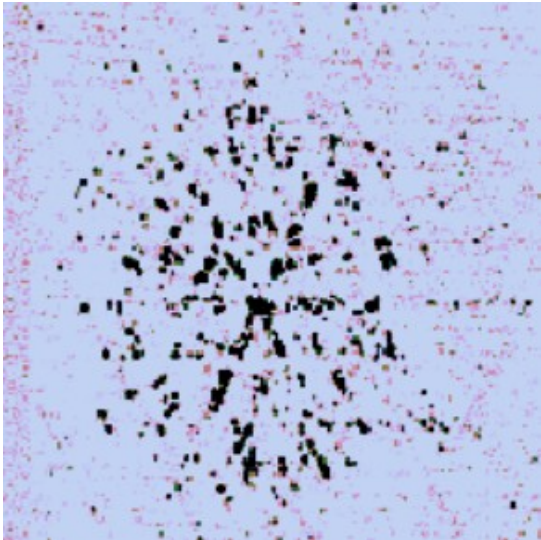
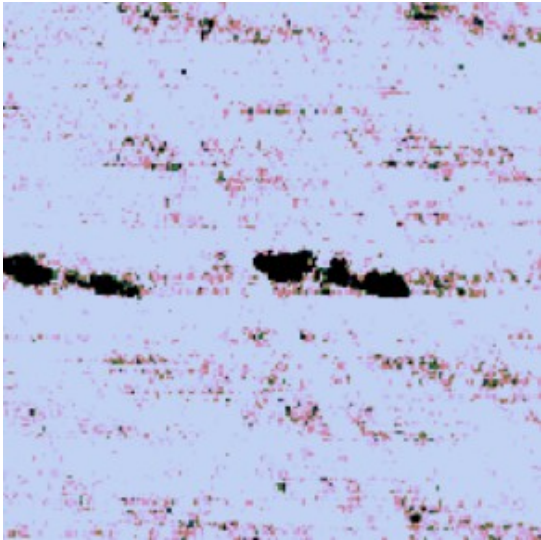
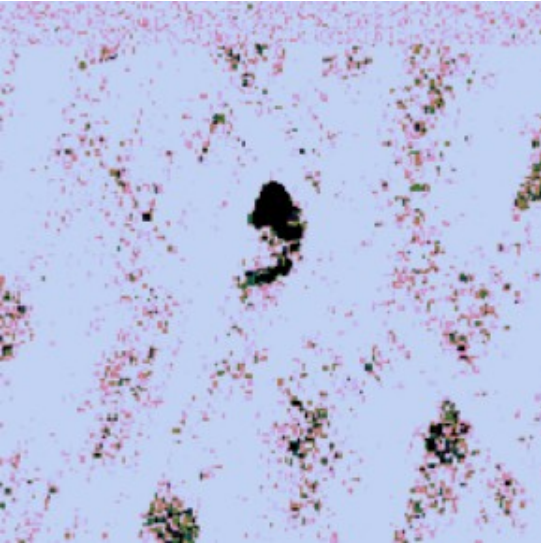


centered off center no source

Finding Outliers

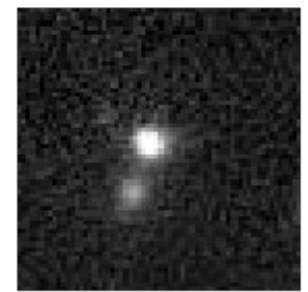
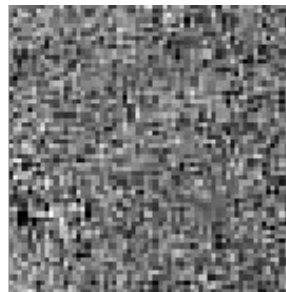
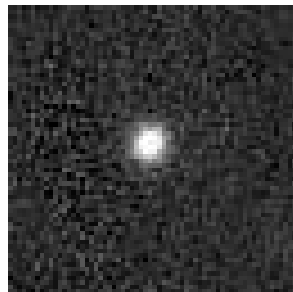
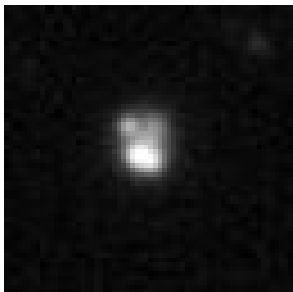
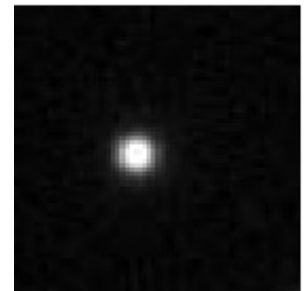
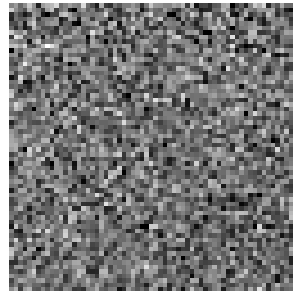
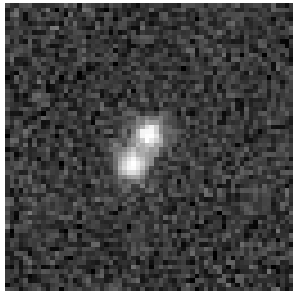


Outliers



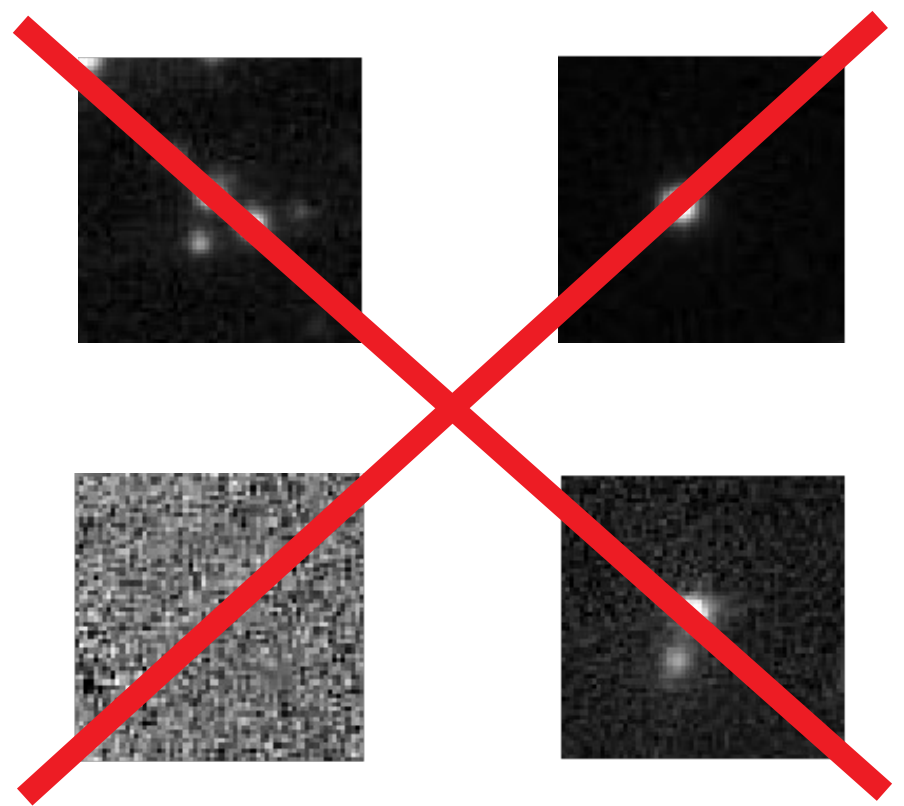
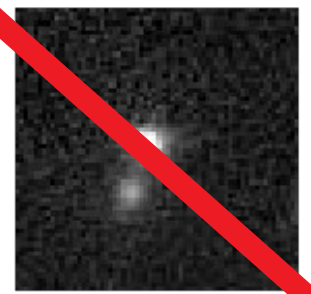
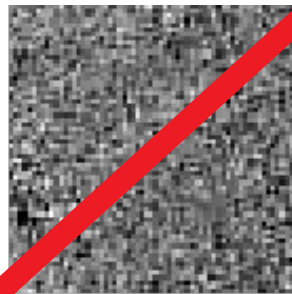
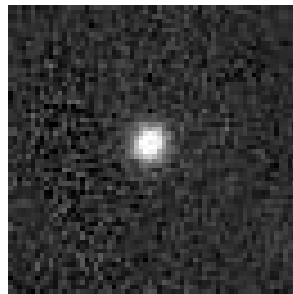
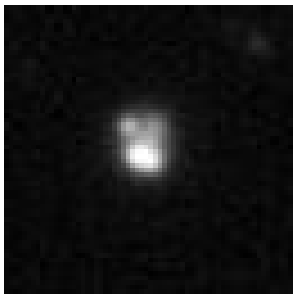
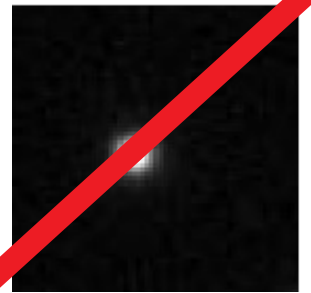
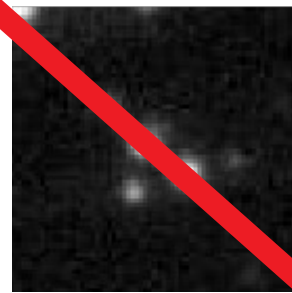
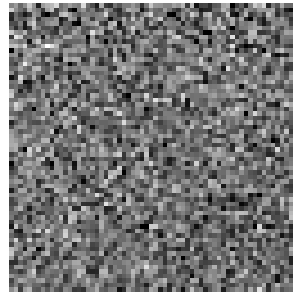
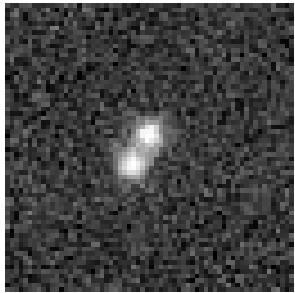
Finding the Needle in the Haystack

- Find strong gravitational lenses in Pan-STARRS



Finding the Needle in the Haystack

- Find strong gravitational lenses in Pan-STARRS



Zooniverse as an Internal Tool



Lenses

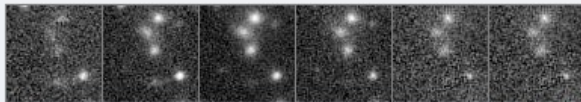
ABOUT

CLASSIFY

TALK

COLLECT

RECENTS



TASK

TUTORIAL

Can you recognize this as a lens?

0: Not recognizable as lens (noise)

1: Slight lens characteristics (single point visible)

2: Possibly a lens (1-2 points visible)

3: Probably a lens (2+ points clearly visible)

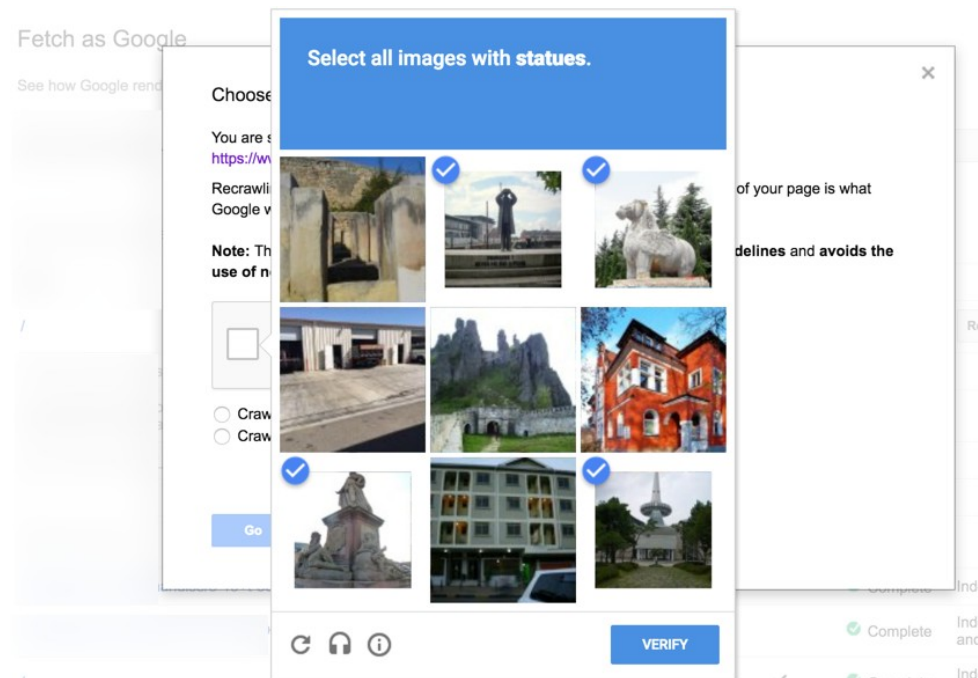
4: Clearly a lens (3+ points clearly visible)

Done



Can crowdsourcing be replaced by GPUs?

- Each has their place
- ML is needed to deal with massive amounts of data
- Crowdsourcing should be thoughtfully applied to the most important data



Acknowledgments



Bernd Doser



Joachim
Wambsganß



Kai Polsterer

Any Questions?

