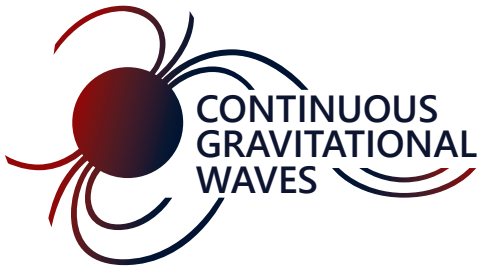


Clustering Observational Data Using Deep Learning Network

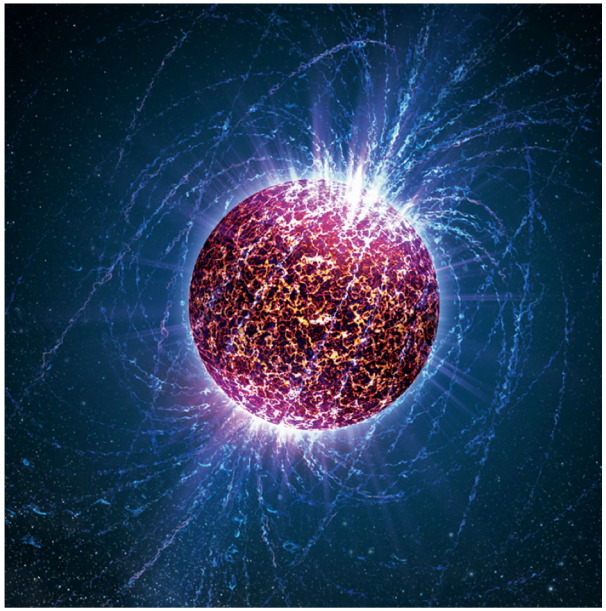
Banafsheh Beheshtipour
Maria Alessandra Papa
Max-Planck Institute for Gravitational Physics
Albert Einstein Institute
06/25/2019

Outline:

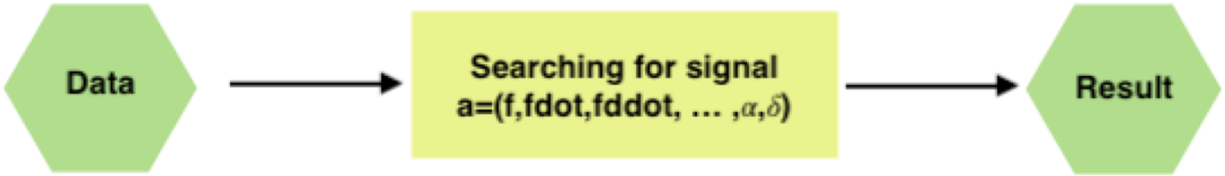
- Continuous gravitational wave searches
  - Clustering
- Deep learning
  - Training set
  - Image segmentation

Continuous Gravitational Wave

- Rapidly rotating neutron stars with non-axisymmetric deformations.
- Has not yet been detected due to their weakness.

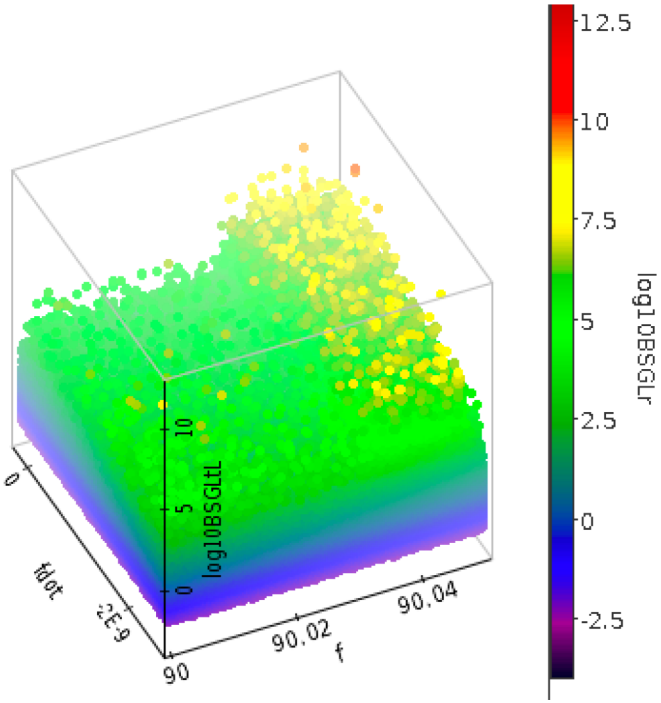


Sensitive detectors & powerful and sophisticated data analysis technique



- All-sky searches for unknown neutron stars.
- Einstein@Home computing power is used for semi-coherent searches.
- LIGO O1 all-sky search (20-100 Hz) (Abbott+ 2017).

- Detection statistic: a score based on match filtering (2F), likelihood of having a GW specific signal.
  - Improved to be Robust against lines and transients, BSGLtL
- A loud signal or a disturbance can produce high values of the detection statistic and over-density of increased values in a region of waveform parameter space.

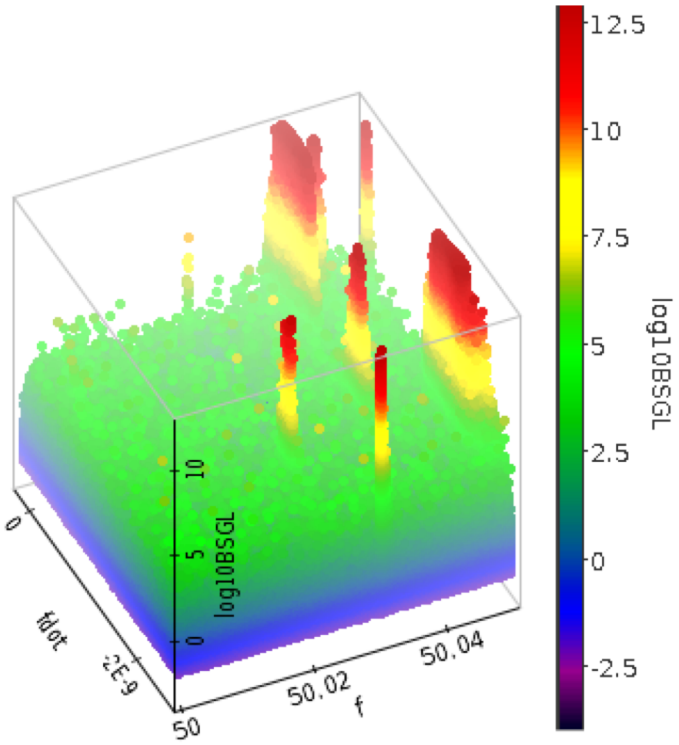


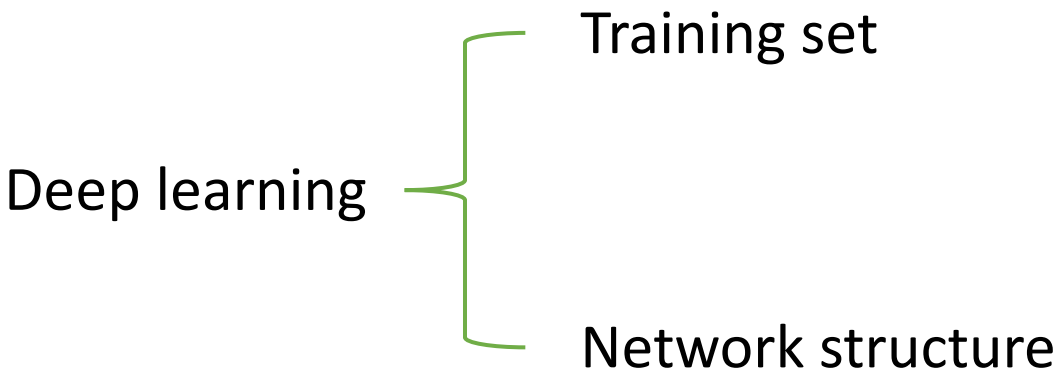
Result for 50 mHz band at 90.000 Hz frequency band

Clustering:

- Bundles together candidates due to the same root cause.
- Reduce the cost of follow-up searches ➔ increases the search sensitivity.

- No exact model to define the cluster volume, e.g.:
  - Based on the average clustering properties of signals (Papa+ 2016, LVC 2017).
  - Adaptive clustering procedure (Singh+ 2017).
  - Deep learning?!

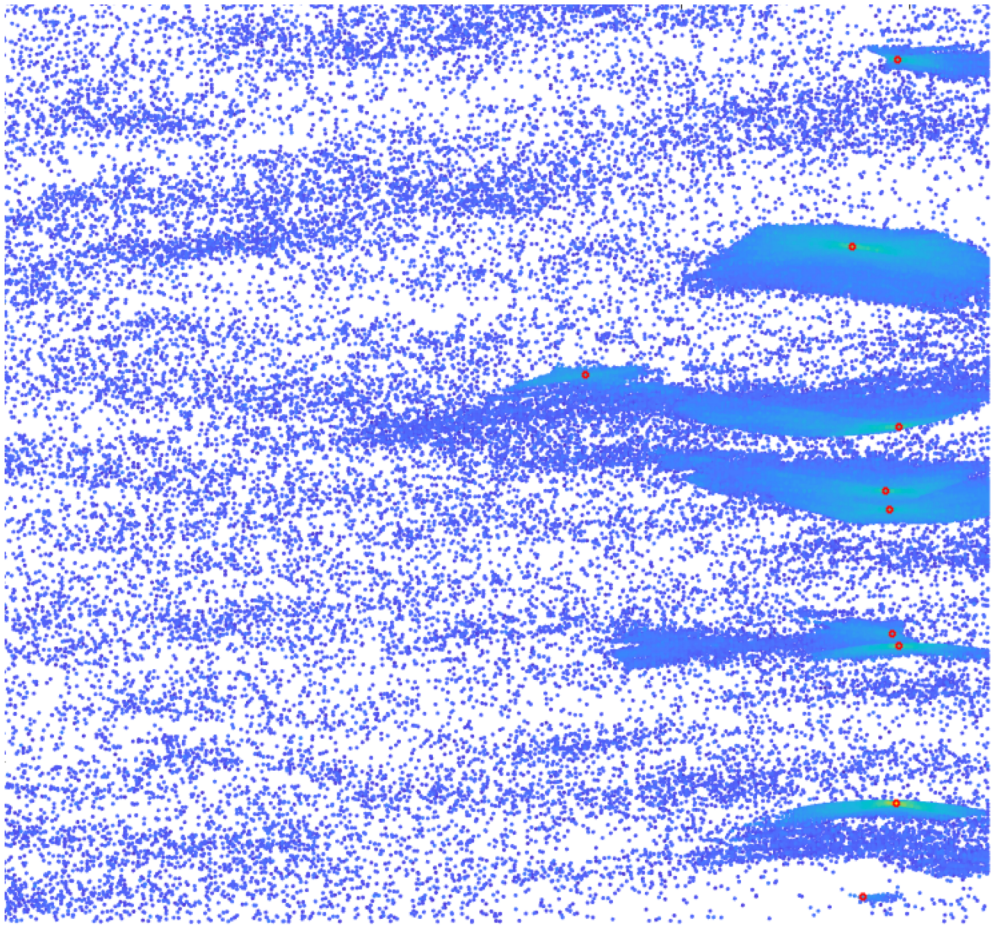



Deep learning
Training set
Network structure

Challenges

- Scale down 4-D data (f, fdot, and sky location) ➔ two sets of 2-D data.

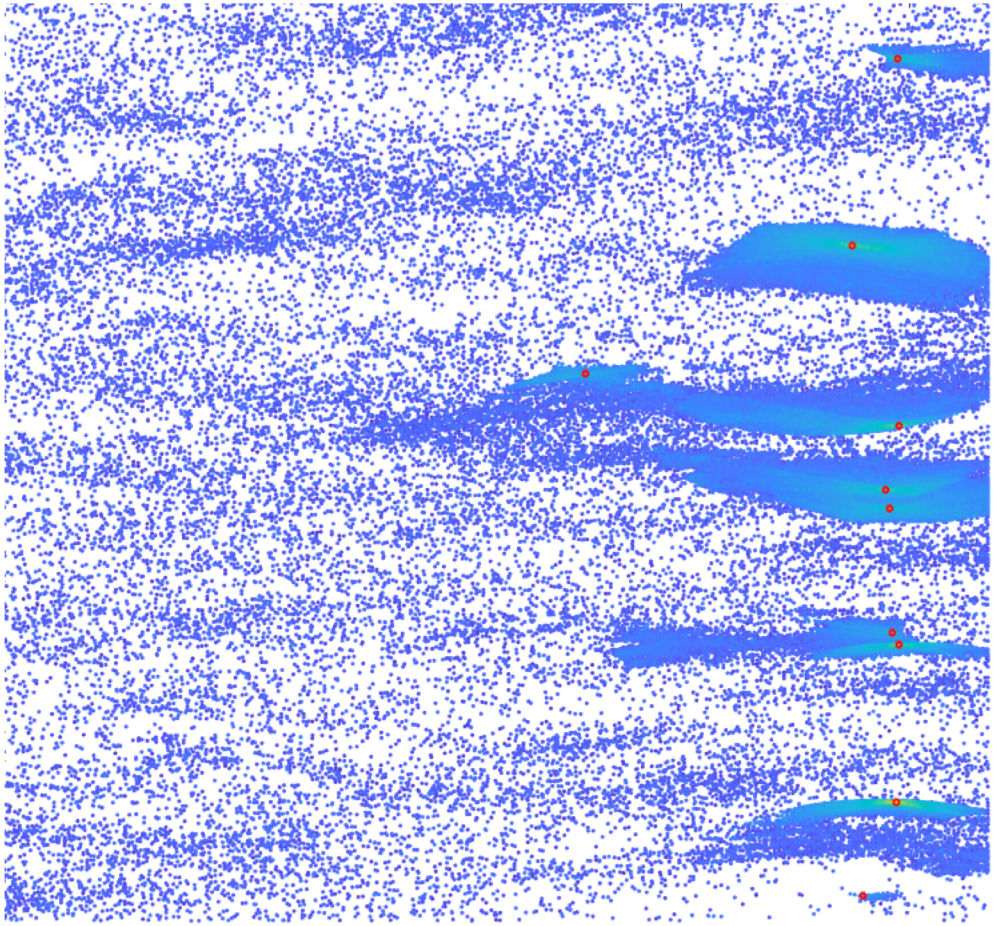
- This problem can be approached as an image processing network
- object detection + segmentation ➔ instance segmentation



50 mHz band at 20.65 Hz

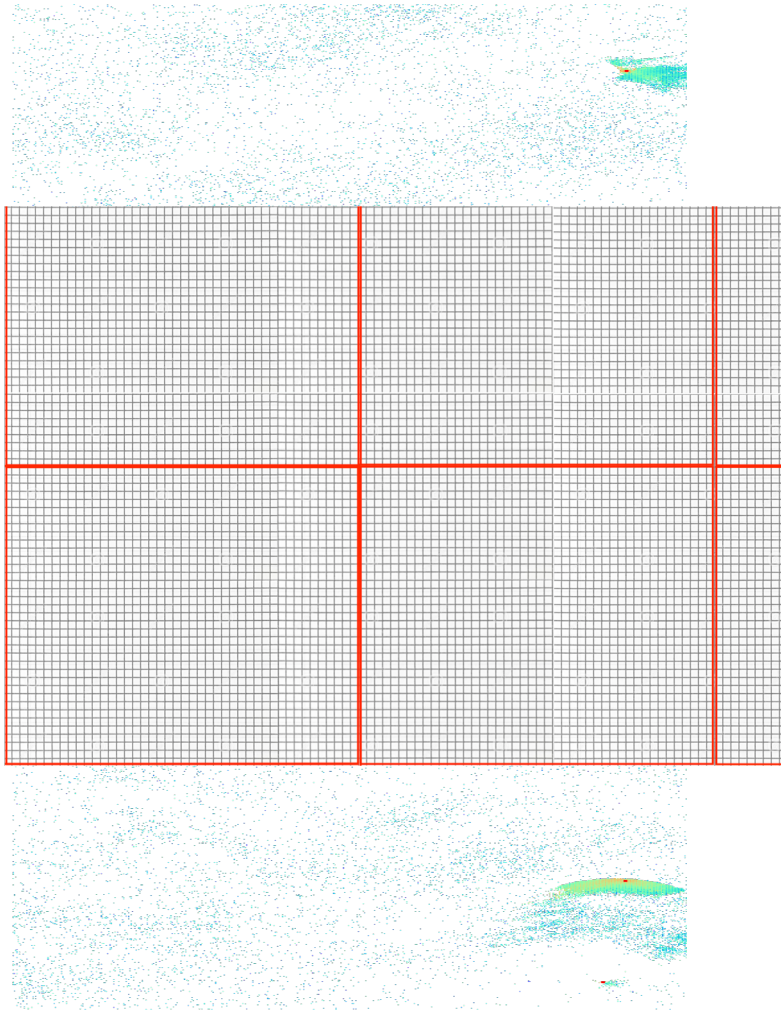
Challenges

- Large number of points, in a typical search around 10^9.
- Data point are sparse.



50 mHz band at 20.65 Hz

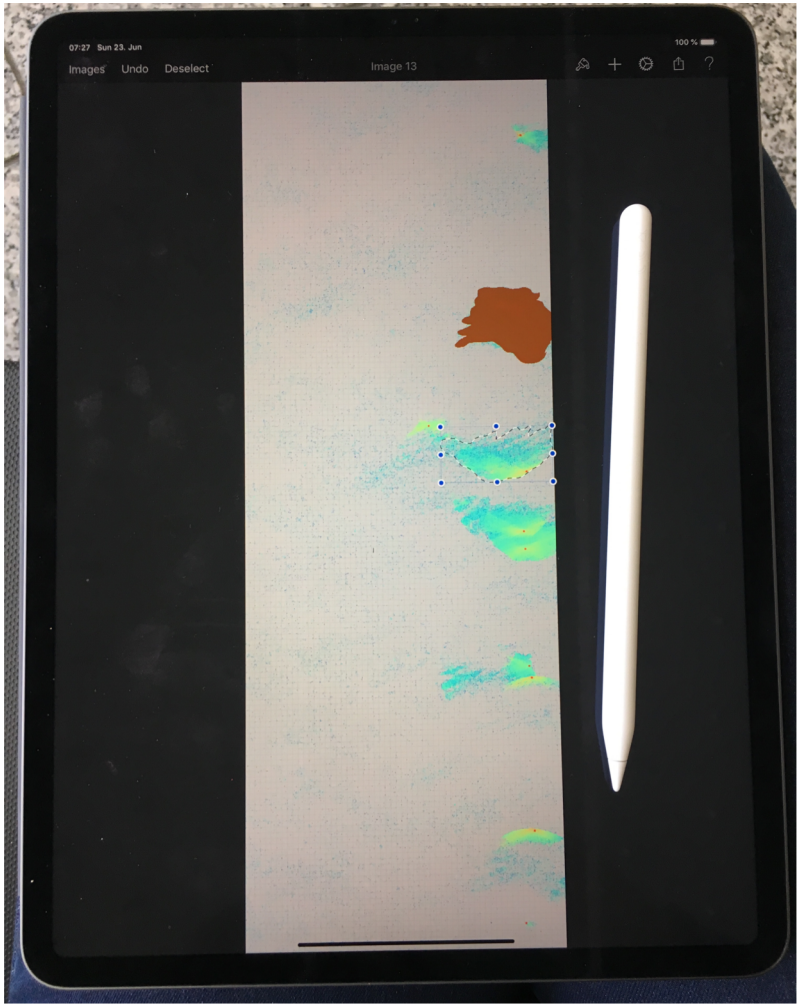
- Reconstruct the image to a lower resolution image
  - divide the original image small size tiles
  - Calculate the density of the point for each tile
  - Assign each tile to one pixel of the reconstructed image
- Reduce the dimensionality by 3-4 orders of magnitude.
- Still too large for the network ➔ tile up the image and give each tile (~200*200) the network.
- Segmentation networks are proven to work in tiles and find the connections between the tiles.



Low resolution 50 mHz band at 20.65 Hz

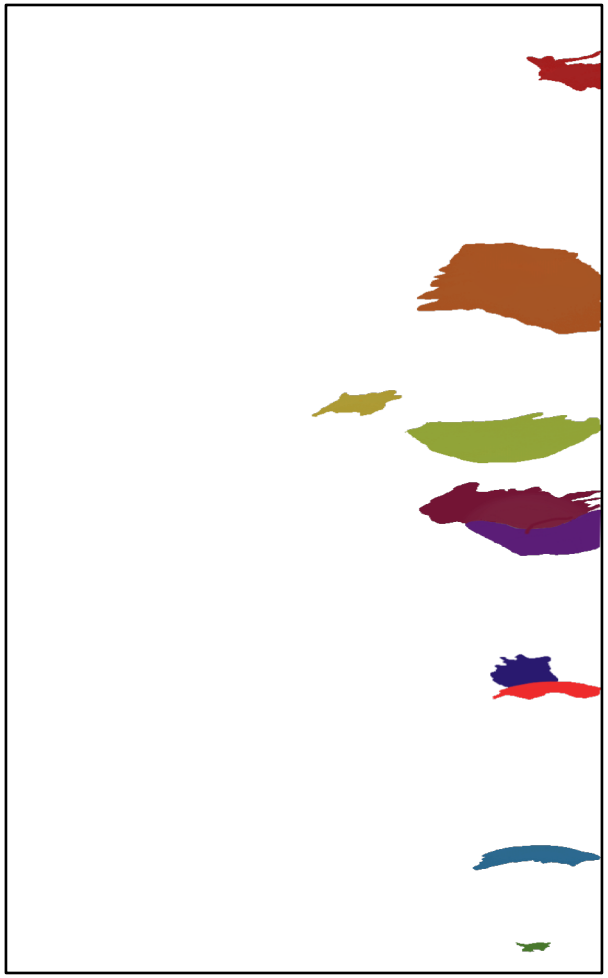
Preparing the mask

- A graphic editor software is used to mark the region covered by each cluster.
- iPad & apple pencil is used



Preparing the mask

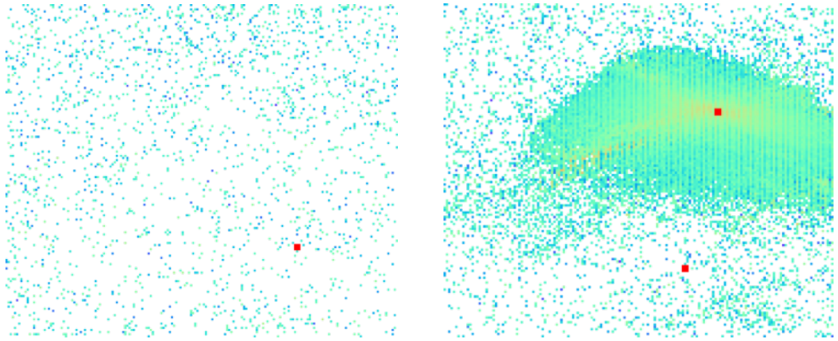
- We compare each image with the back ground noise to make sure of the edges of each cluster.



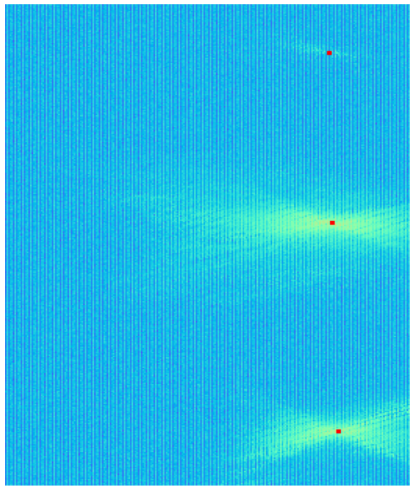
Mask of clusters in 50 mHz band at 20.65 Hz

Important Strategic Choices:

- Separating the classes into large/small clusters (visible/invisible cluster)
- Focusing on a certain frequencies for the input
- 4D/2D data?



20.65 Hz



92.40 Hz

Conclusion

- Clustering is an important step in searching for continuous gravitational waves.
- Using a deep learning network , we aim to automate the clustering algorithm.
- Instance segmentation network, Mask-RCNN, seems to be a good approached for this problem.
- We are preparing the training set and will start training the network once it is ready.