Unsupervised Machine Learning For Identifying Signatures Of Past Mergers In Galaxy Light Profiles

Joyo Smit (he/they)

Saint Mary's University

Background

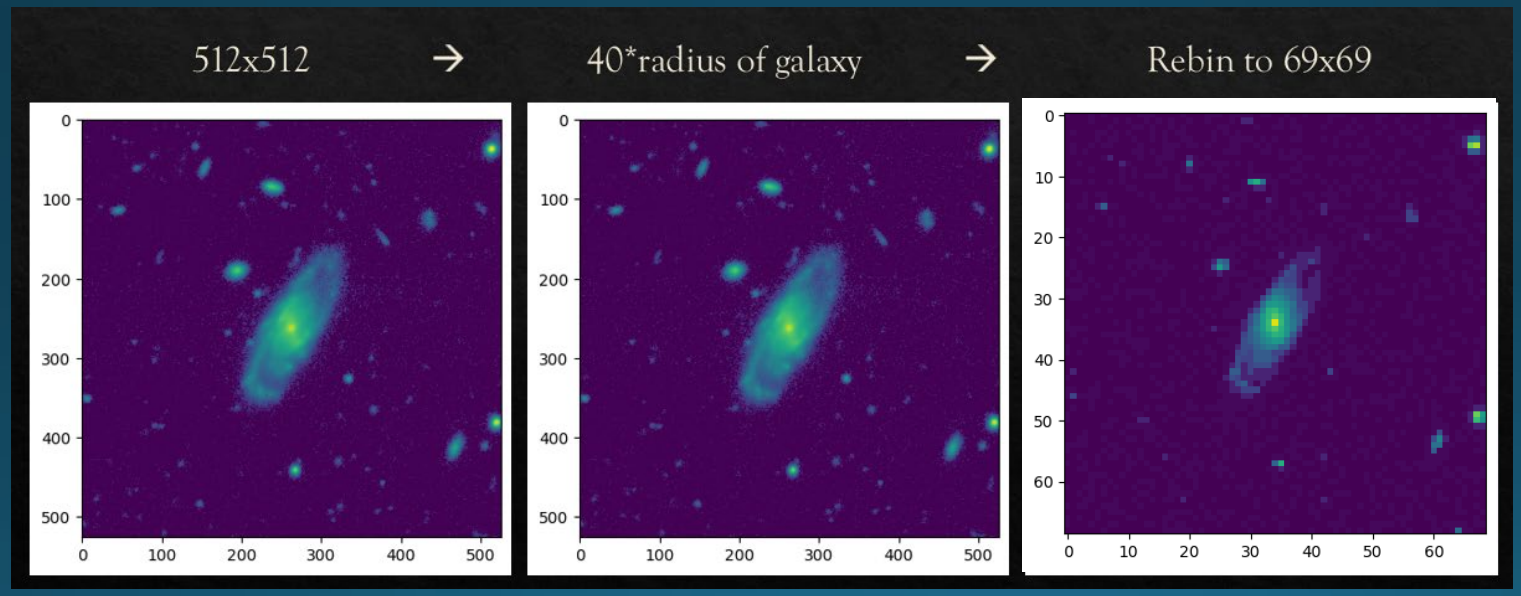
- Galaxy interactions → stars and gas being pulled into different shapes (tidal features)
  - Streams & shells
- Difficult to detect by eye
- Solution: Machine learning!
  - Previous research: not enough labels for supervised → try unsupervised

Fig. 1. Two galaxies interacting, producing tidal streams. Photo credit: HSC-SSP

Methodology

1. Scale images

Image Dataset: Hyper Suprime-Cam Subaru Strategic Program PDR-3 512x512 pixels in g,r,i bands with labels for tidal features (for method testing)



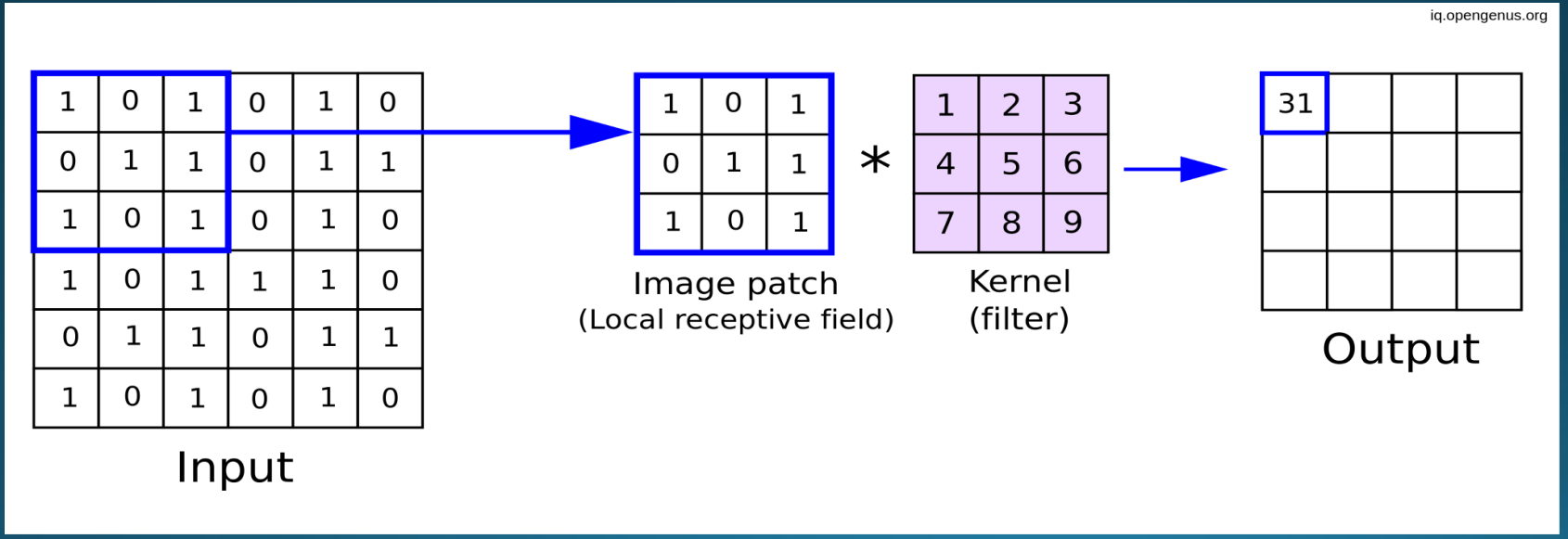
*https://hsc-release.mtk.nao.ac.jp/doc/

Methodology

2. Augmentations of scaled images

- Maintain tidal features
- Result: image pair
  - one image scaled and augmented, one scaled and non-augmented

3. Convolution Step



Methodology

3. Compute similarity (loss)

- Between augmented and non-augmented image pair

$$\text{Cosine Similarity} = \frac{\sum_{i=1}^{n} x_i y_i}{\sqrt{\sum_{i=1}^{n} x_i^2}\sqrt{\sum_{i=1}^{n} y_i^2}}$$

Methodology

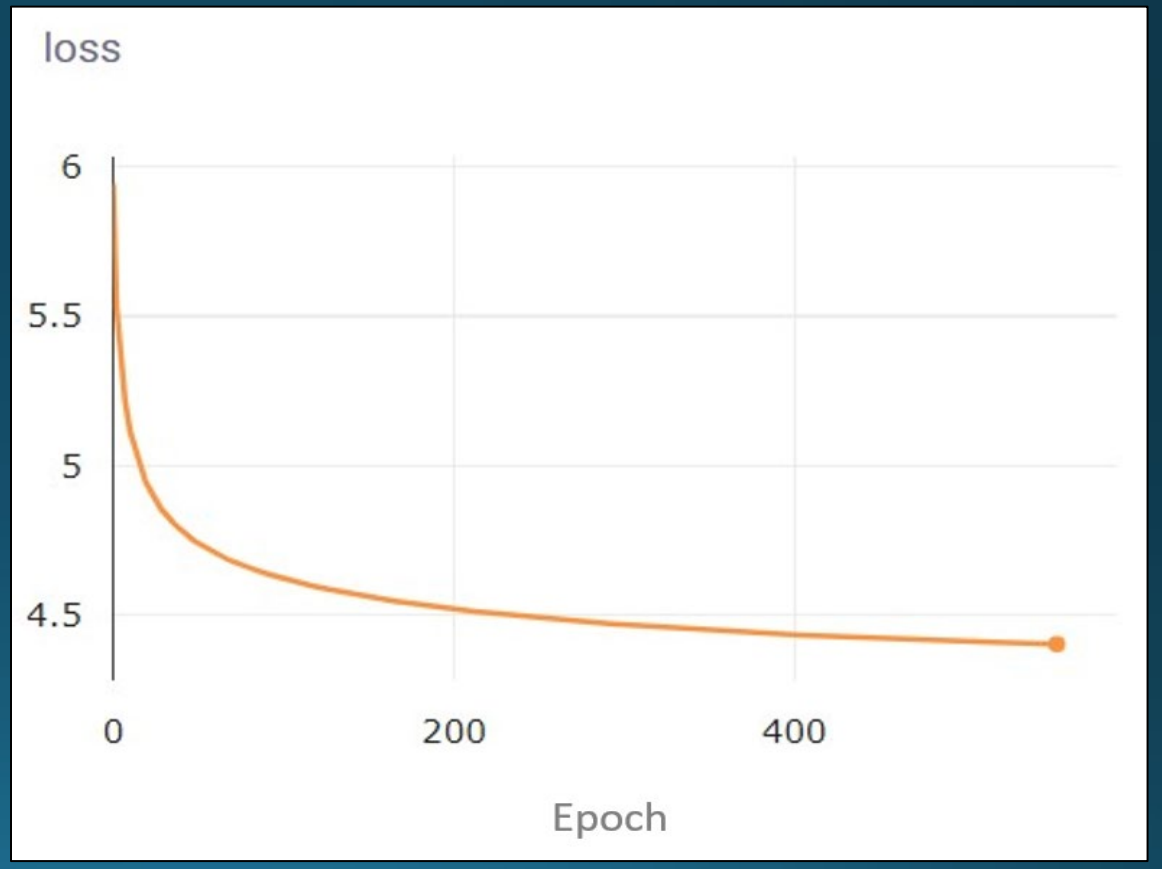
5. Adjust weights through iteration

Goal: minimize loss function between images within pairs and maximize it between different pairs

Output

multidimensional representation of dataset

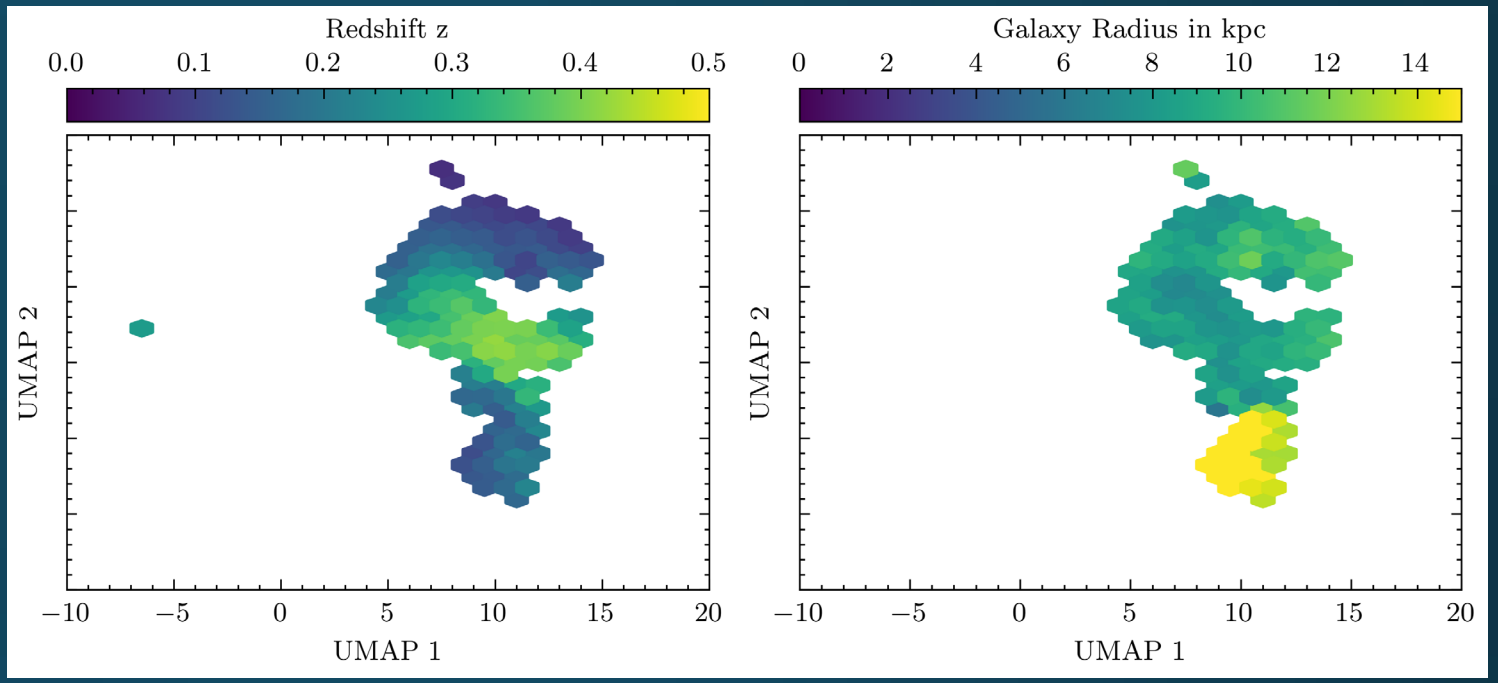
- Similar images grouped
- Scale to 2D
- Plot parameters

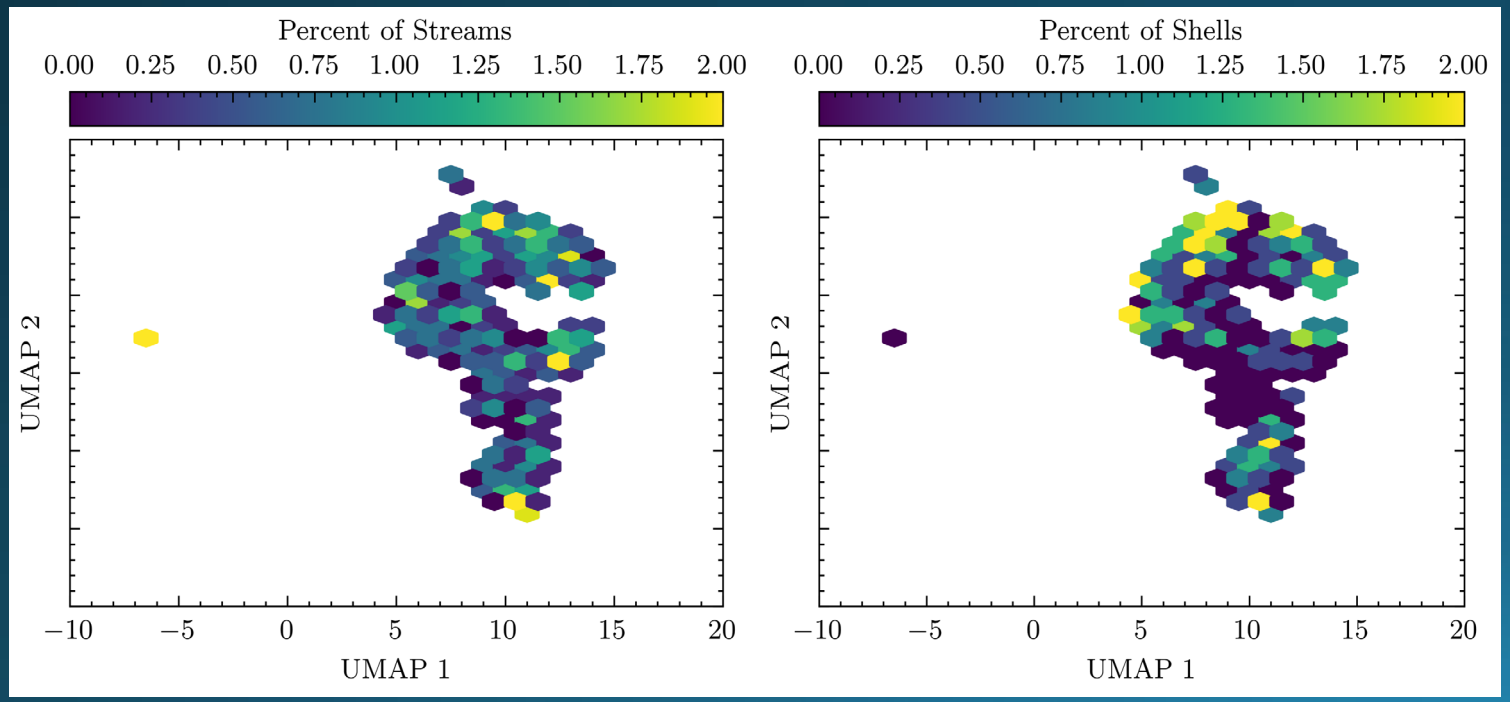


Results

Top Plots

- Clear trends in galaxy redshift and galaxy size





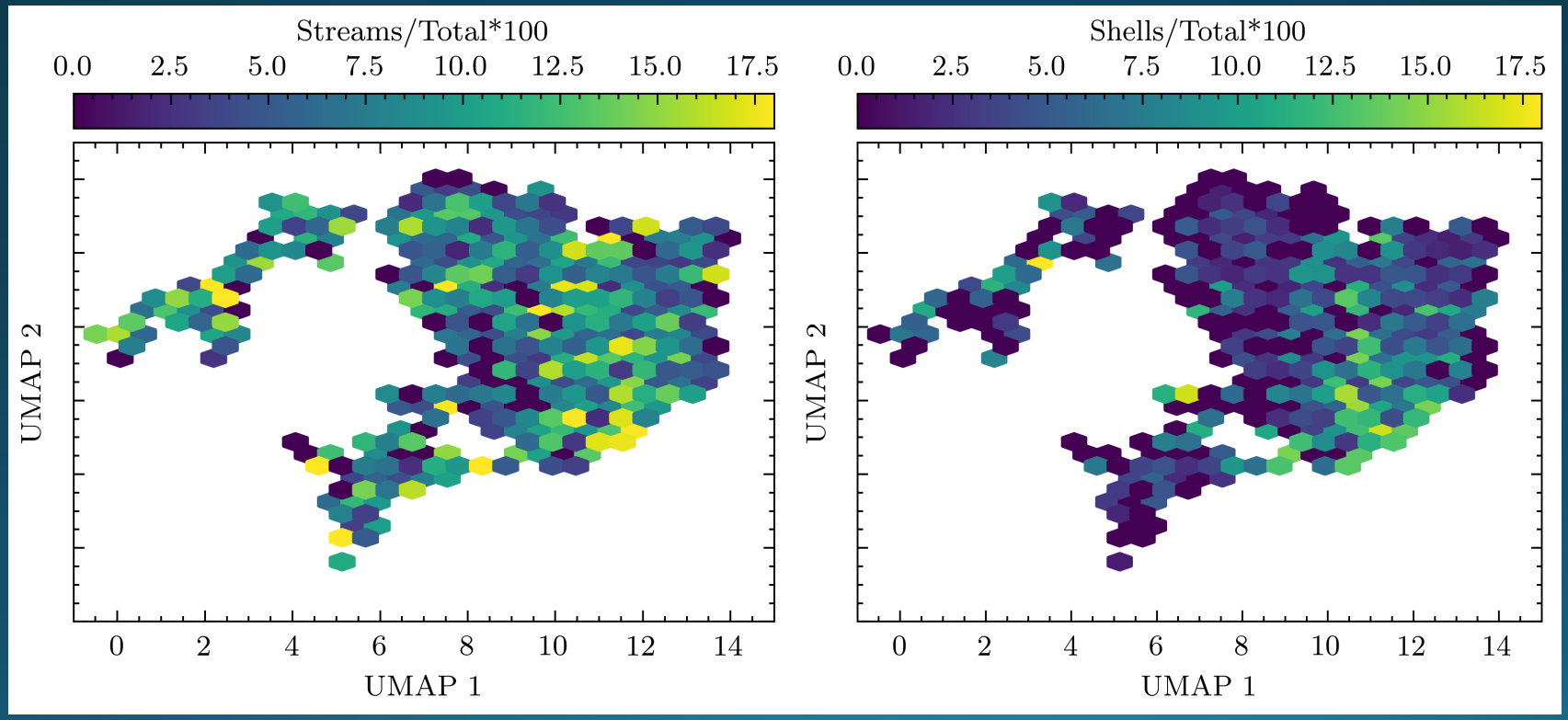
Bottom plots

- Faint trend in tidal shell plot
- No trend in tidal streams

Results

Run again

- Each image has its reflection subtracted
- Emphasis on asymmetric data like tidal features
- Improved results



Conclusion

- Unsupervised machine learning does separate galaxies based on their parameters, but not the parameters we want
- To continue improving the model, we need a larger set of galaxy images and to experiment with augmentations.

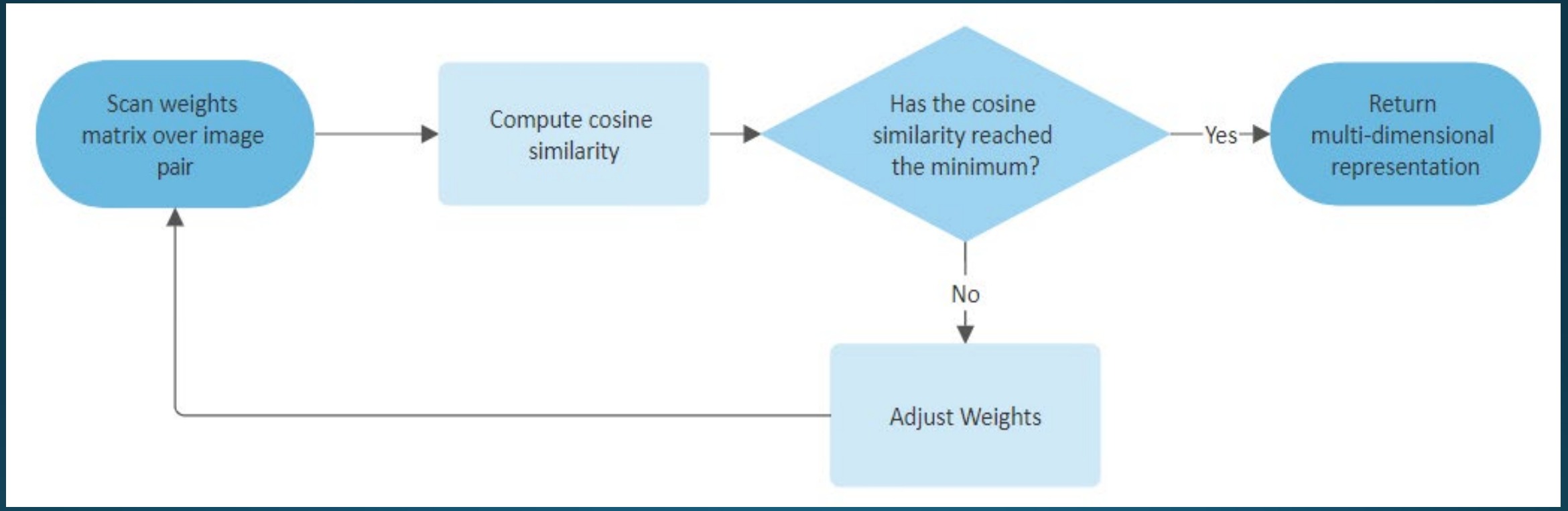
NTCO

Saint Mary's University

Centro de Estudios de Física del Cosmos de Aragón
CEFCA

Thank you!

Universidad de Valladolid


Scan weights matrix over image pair
Compute cosine similarity
Has the cosine similarity reached the minimum?
Yes
Return multi-dimensional representation
No
Adjust Weights

Streams

Shells