

Quantifying HI morphologies in the MGCLS-HI sample : From 2D Analysis Towards 3D and Machine Learning Approaches

ABSTRACT

Neutral hydrogen (HI) asymmetries are vital indicators of gas accretion and removal processes that drive galaxy evolution. However, current classification methods - both parametric (e.g tilted-ring models) and non-parametric often fail to reliably identify morphologically disturbed galaxies. This project aims to address these challenges by developing a scalable Machine Learning framework for automated classification on 3D data.

DATA SAMPLE



Figure 1: A schematic showing the sample selection criteria applied to MGCLS HI data.

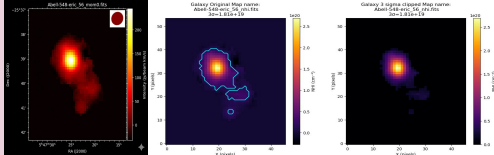


Figure 2: The three images show moment 0 maps being converted into HI Column Density (NH) maps, then 3 σ clipping threshold was applied to the resulting maps.

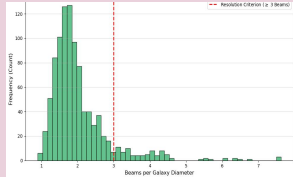


Figure 3: The figure shows the distribution of beams per galaxy diameters in the MGCLS-HI sample. The vertical red line indicates the borderline between resolved and unresolved galaxies used in this study.

HI MORPHOMETRICS

From the initial **967 detections**, a final sample of **88 resolved galaxies** was obtained. The non-parametric morphometrics were employed to objectively measure the asymmetries properties of these galaxies, specifically focusing on the **Asymmetry (A)**, **Concentration (C)**, **Smoothness (S)**, **Gini Coefficient (G)**, and the **M(20)** moment of light.

Below are the equations that were used to calculate the asymmetry of the 88 resolved galaxies ;

1. Concentration (C)

$$C = 5 \cdot \log_{10} \frac{r_{50}}{r_{20}}$$

2. Asymmetry (A_{mod})

$$A_{mod} = \frac{1}{N} \frac{\sum |I_{ij} - I_{rot,ij}|}{\sum |I_{ij}| + \sum |I_{rot,ij}|}$$

Where N is the total number of pixels, I_{ij} is the original pixel value, and $I_{rot,ij}$ is the corresponding pixel value in the rotated version.

3. Smoothness (S)

$$S = \frac{\sum |I_{ij} - I_{s,ij}|}{\sum |I_{ij}|}$$

Where N is the total number of pixels and $I_{s,ij}$ is the smoothed (blurred) version of the original image.

4. Gini Coefficient (G)

Where X_i are the pixel values sorted in increasing order, and n is the total number of pixels.

$$G = \frac{1}{N} \sum_{i=1}^n \frac{(2i-n-1)}{n-1} |X_i|$$

Where X_i are the pixel values sorted in increasing order, and n is the total number of pixels.

5. Moment of Light (M_{20})

$$M_{20} = \log_{10} \frac{\sum M_{ij}}{M_{tot}}, \text{ for } \sum f_i < 0.2 f_{tot}$$

Where $M_{ij} = f_i [(x_i - x_0)^2 + (y_i - y_0)^2]$ represents the second-order moment of a pixel.

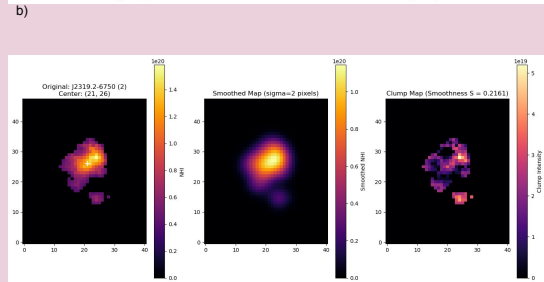
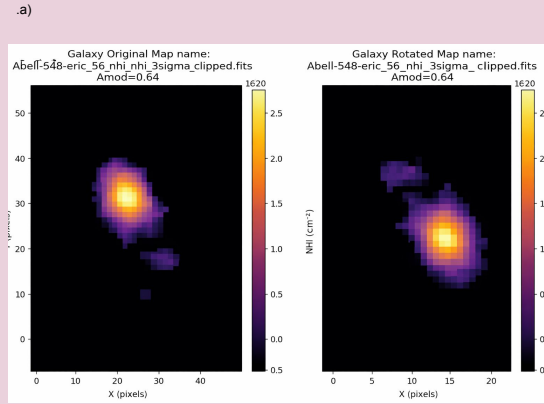


Figure 3: The images in panel a) show side-by-side maps of a galaxy and its 180° rotated version these are steps done during Amod calculations. For this galaxy the asymmetry was 0.64 after subtracting the rotated image from the original one. Images in panel b) show some steps taken during smoothness calculation. The final **Smoothness S = 0.2161** is derived from the ratio of the flux in this residual map to the total flux in the original map

RESULTS

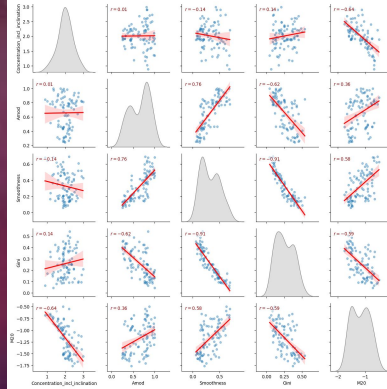


Figure 4: This graph is a **Correlation Matrix (Pair Plot)** visualizing the relationships between different morphological indices for the sample of galaxies. There is a **Strong Positive Correlation** ($r = 0.76$) between **Asymmetry A(mod)** and **Smoothness**. This makes sense because galaxies that are clumpy and "knotty" are rarely perfectly symmetrical..

CONCLUSION AND FUTURE WORK

We will move beyond spatial analysis by utilizing **3D HI data cubes**, enabling the integration of velocity information to distinguish between internal feedback and external environmental stripping and then incorporate **machine learning techniques** for automated classification of HI morphologies.