Exploring the Clustering and Diversity of Low Surface Brightness Galaxies in Dark Energy Survey Data

2025-02-20 @ PAiP-2025

Michal Vrabel*, Hareesh Thuruthipilly*, Junais*✤, Agnieszka Pollo*, Katarzyna Małek*

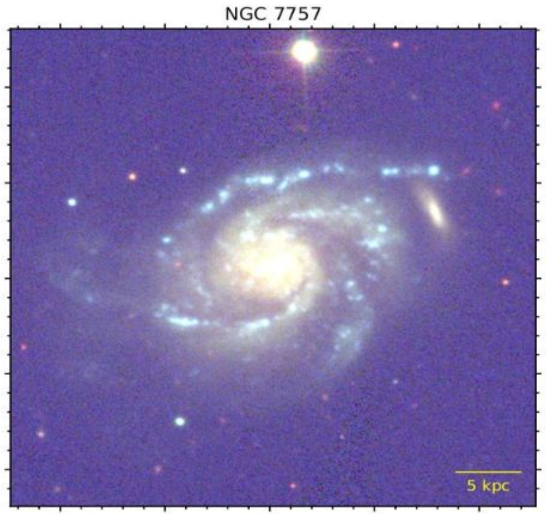
*National Centre for Nuclear Research (NCBJ), Warsaw, Poland
✤Instituto de Astrofísica de Canarias (IAC), Tenerife, Spain

NATIONAL SCIENCE CENTRE POLAND

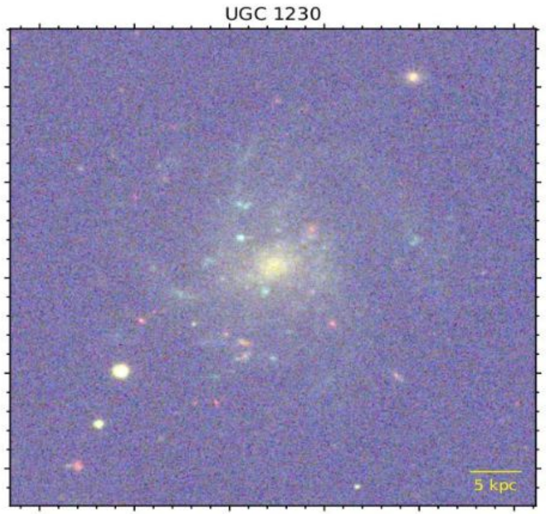
NATIONAL CENTRE FOR NUCLEAR RESEARCH ŚWIERK

What is a low surface brightness galaxy

Low surface brightness galaxy = Galaxy that emit much less light per unit area than "normal" galaxies.



High surface brightness galaxy (HSB)



Low surface brightness galaxy (LSB)

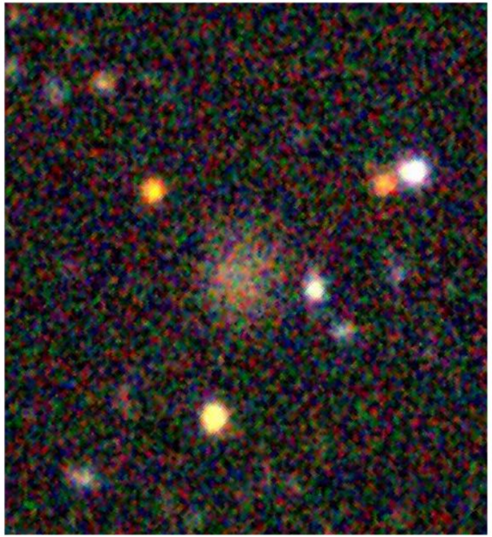
An example: Two galaxies which are about the same distance and size, but very different in their light.

Image credit: McGaugh 2021

Same LSB as seen by different instruments

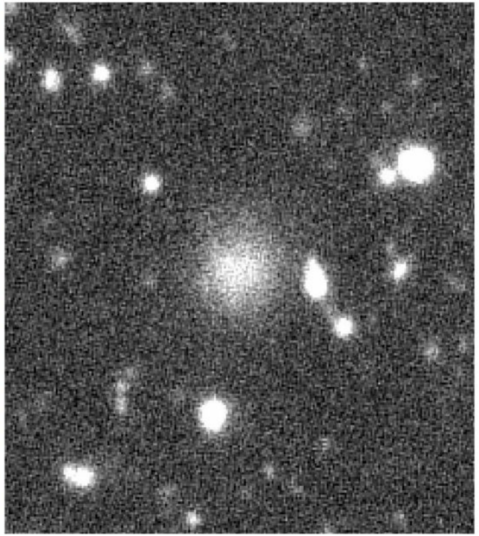
Pan-STARRS survey (2016)

1.8m telescope; 350 seconds exposure



DES survey (2018)

4m telescope; 900 seconds exposure



HSC data (2024)

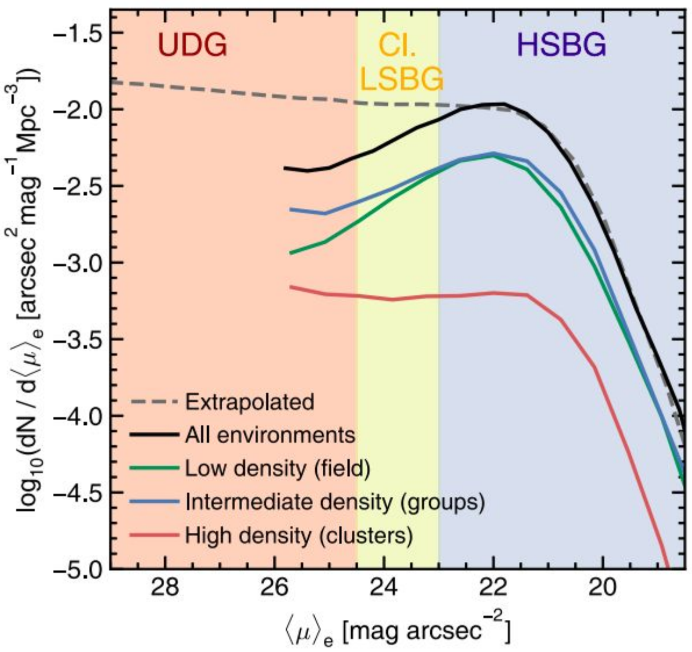
8.2m telescope; 4500 seconds exposure

Slide: K. Małek

Importance of LSBs

- LSBs are generally defined as a galaxies with an **average surface brightness** $\mu_{e,r}$>23 mag arcsec^{-2} ,
- LSBs may account up to
  - **30-60% of the galaxy population**, and
  - **15-20% of the total dynamical mass** contained in galaxies,
- Only **limited studies** on LSBs due to their faintness,
- Vast discovery space for LSBs with **new deep surveys**, like for example **LSST** or **Euclid**!

Faint ← Bright



Martin et al. 2019

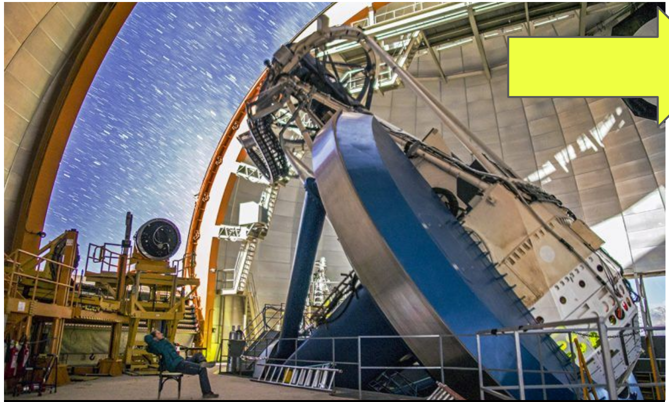
Work of our group on the topic of LSBs

NATIONAL CENTRE FOR NUCLEAR RESEARCH ŚWIERK

Astrophysics division **BP4**

- **Study of LSB galaxy evolution**
  - Dust content (Małek, Junais, Pollo +2024; Junais, KM +2023)
  - Metalicity (Junais et al. 2024)
  - Molecular gas (Galaz, Junais et al. 2024)
- Finding **statistically large sample of LSBs** using Transformer models (Thuruthipilly et al. 2024)
  - LSB searches in other datasets such as KiDS, HSC, etc. (Thuruthipilly in prep.)
  - Followed by spectroscopic analysis of selected samples (Vanzanella in prep.)
  - Study of spatial clustering strength in relation to surface brightness (Sureshkumar in prep.)
  - **Categorization of LSB galaxies (Vrábel in prep.)**
    - Organization and clustering by morphological and photometric properties

LSBs in Dark Energy Survey DR1 catalog



Dark Energy Survey (DES)

Observing period: 2013–2019 (6 years)
Coverage: ~5000 deg² (southern Galactic cap)
Wavelengths: Optical & near-infr
Bands: grizY
Instrument: **Dark Energy Camera (DECam)**
Telescope: Blanco 4m Telescope (CTIO, Chile)
Abbott et al. 2018, 2021; darkenergysurvey.org

DES Y3 Gold (DR1)

Data release: First 3 years of DES (Y3)
Catalogue: **DES Y3_gold_2_2.1**
Detection: **SourceExtractor**
Processing: Image depth & quality cuts, calibration, deblending

DES LSBG Catalogue (Tanoglidis et al. 2021b)

Method: **SVM + Sérsic fitting + visual inspection**
SVM training: ~8000 labeled objects (640 LSBGs)
Artefact removal: 20,000+ rejected
Final steps: Sérsic fitting, extinction correction, selection cuts
Catalogue size: **23,790 LSBGs**

LSBG Identification with Transformers (Thuruthipilly et al. 2024)

Method: Transformers + Sérsic fitting + visual inspection.
Training: Multiband cutouts of 23,103 LSBs and 18,474 artefacts
Results: 4,083 new LSBGs (+17%),
1310 cluster LSBGs, 317 ultradiffuse galaxies
Combined catalogue size: 27,873 LSBs

+ additional catalogue of **50,000 rejected candidates**

This work

This work

Motivation

- Find useful (non-obvious) **subclasses** of LSBs
- Find **relations** between LSB properties
- **Extend** LSBs sample through refined fits and less strict selection criteria

The presented results are preliminary

Method

- **Dimensionality reduction and clustering methods** applied on the LSB sample
- Utilizing **morphological** and **photometric** parameters
- Relate to **additional galaxy properties**
  - category, nucleation, cluster association, etc.
- **Evaluating nucleation** by comparing quality of morphological fits
  - Sérsic profile fitting using Galfit,
  - based on Yagi 2016 method

Yagi, Masafumi, et al. “Catalog of Ultra-Diffuse Galaxies in the Coma Cluster from Subaru Imaging Data.” *The Astrophysical Journal Supplement Series*, vol. 225, no. 1, 2016, p. 11, https://doi.org/10.3847/0067-0049/225/1/11

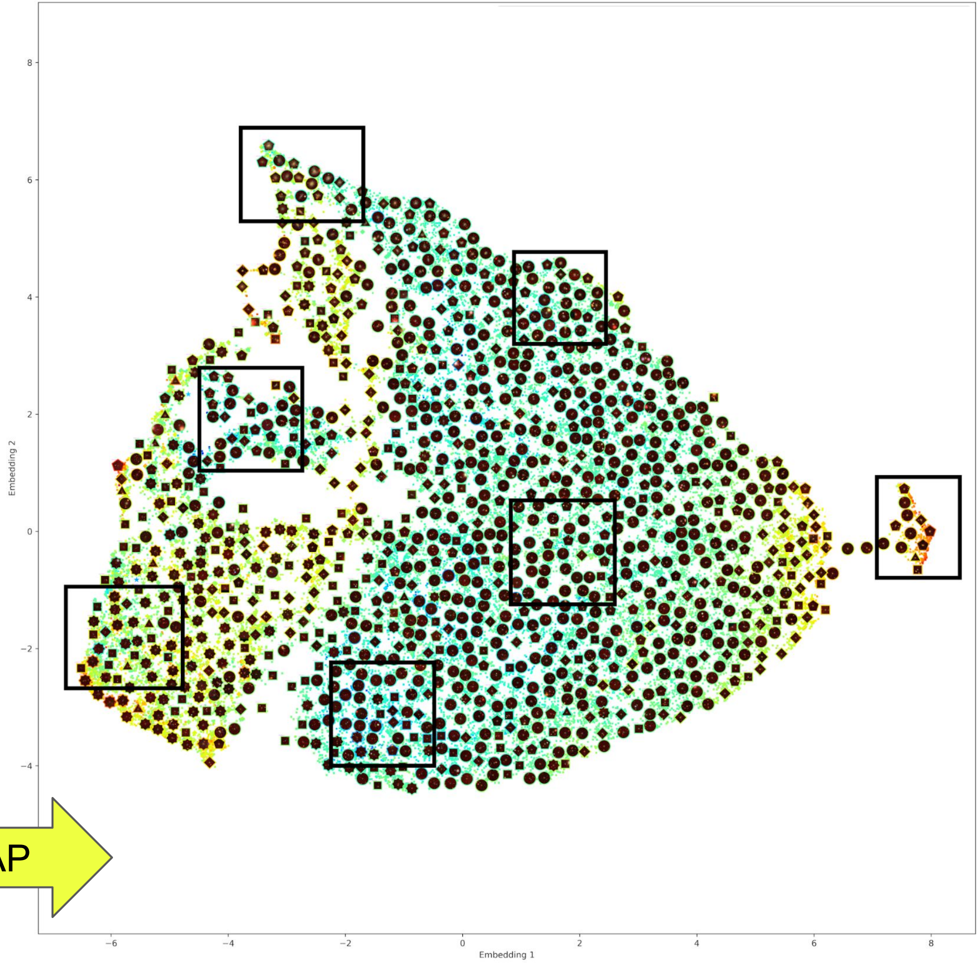
Embedding of the combined Tanglitis and Thuruthipilly catalogs

morphological and **photometric** parameters transformed into 2D dimensions

(UMAP method - nonlinear embedding)

Magnitude (g/i, Extinction Corrected)
Mean Surface Brightness (g/i, Extinction Corrected)
Effective Radius (arcsec, g/i)
Sérsic Index (g-band)
Axis Ratio (g-band)
Color (g - i)

UMAP



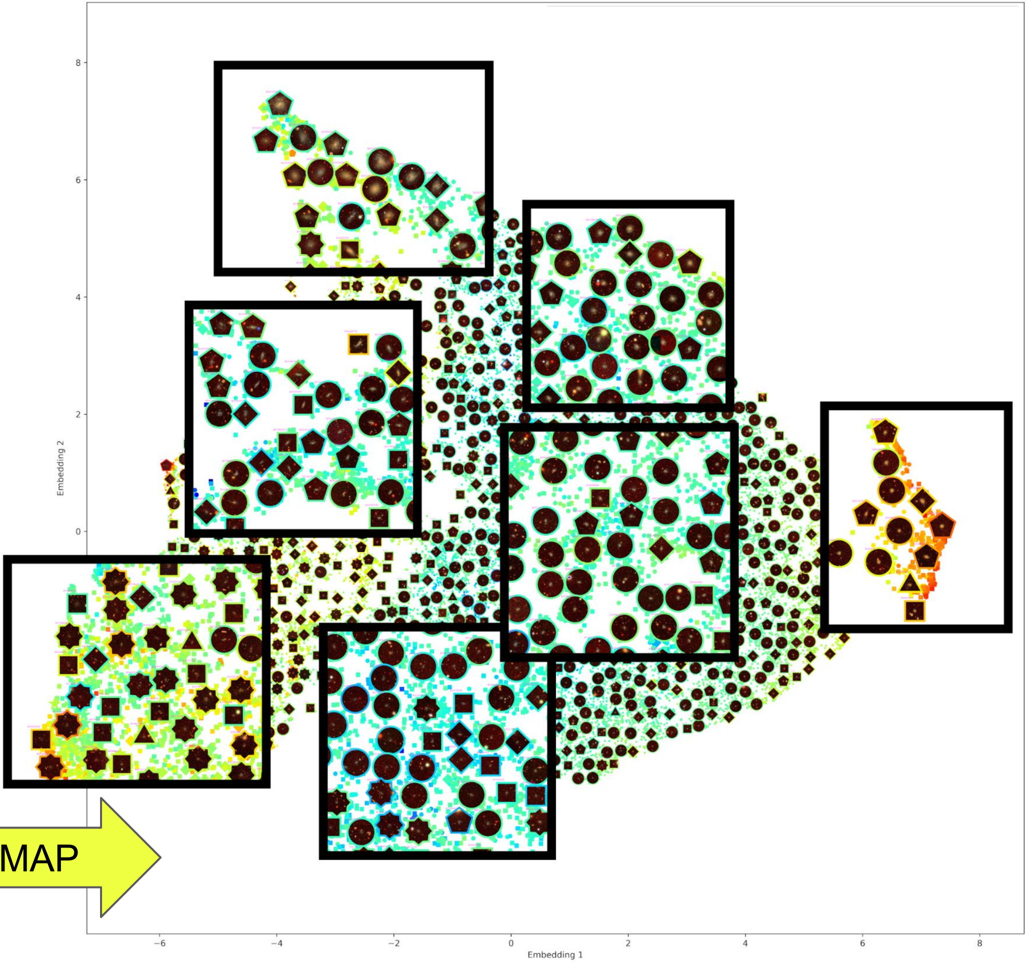
Embedding of the combined Tanglitis and Thuruthipilly catalogs

morphological and **photometric** parameters transformed into 2D dimensions

(UMAP method - nonlinear embedding)

Magnitude (g/i, Extinction Corrected)
Mean Surface Brightness (g/i, Extinction Corrected)
Effective Radius (arcsec, g/i)
Sérsic Index (g-band)
Axis Ratio (g-band)
Color (g - i)

UMAP

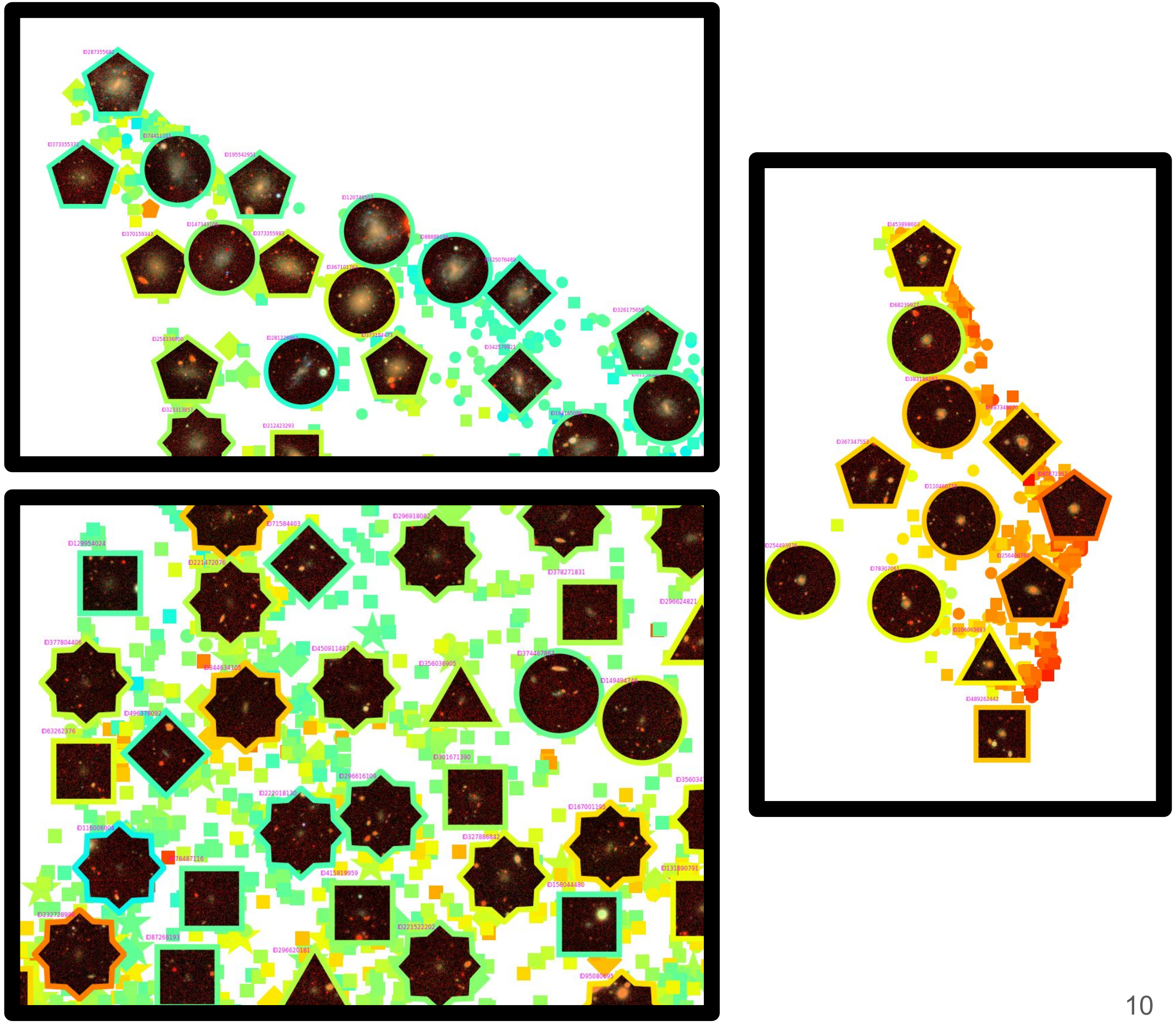


Embedding of the combined Tanglitis and Thuruthipilly catalogs

morphological and **photometric** parameters transformed into 2D dimensions

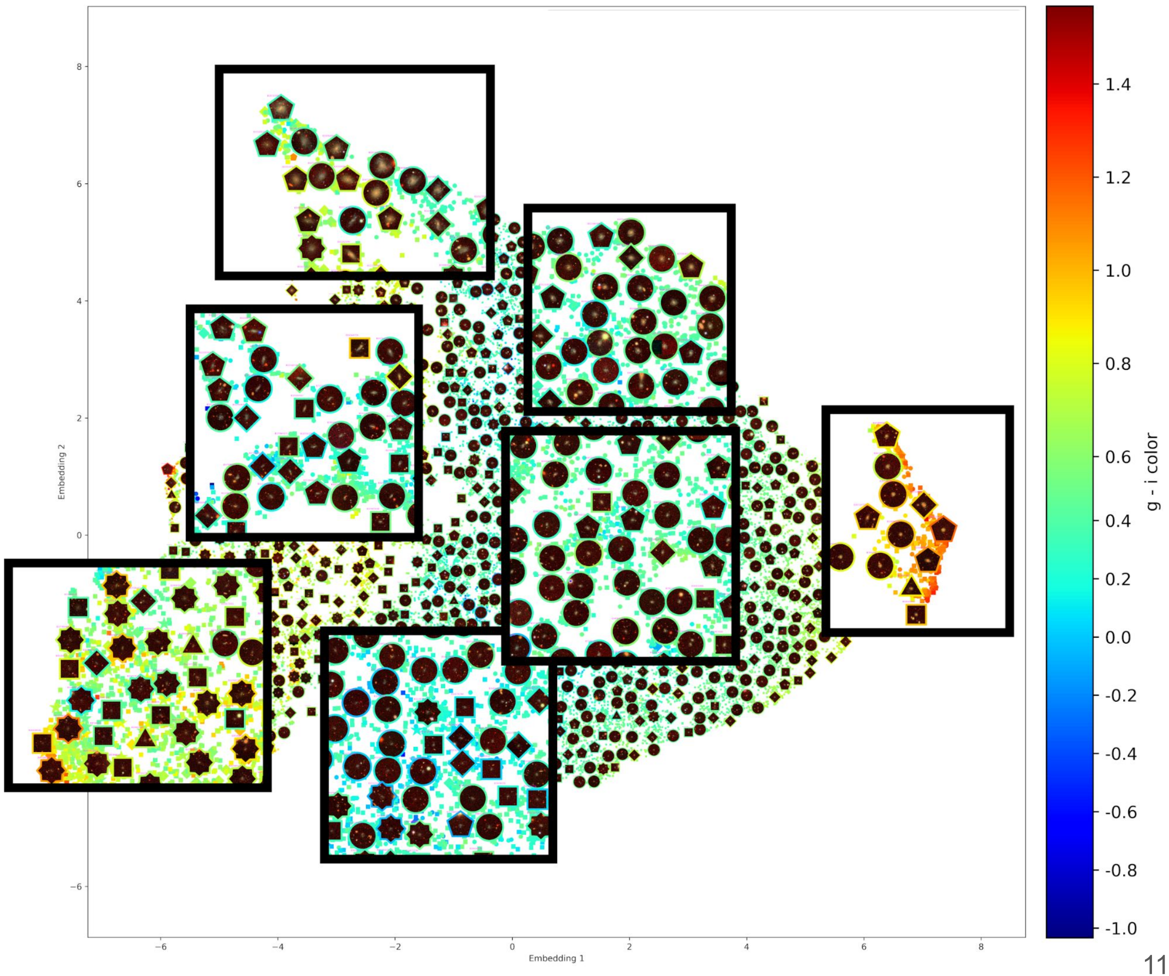
(UMAP method - nonlinear embedding)

Magnitude (g/i, Extinction Corrected)
Mean Surface Brightness (g/i, Extinction Corrected)
Effective Radius (arcsec, g/i)
Sérsic Index (g-band)
Axis Ratio (g-band)
Color (g - i)



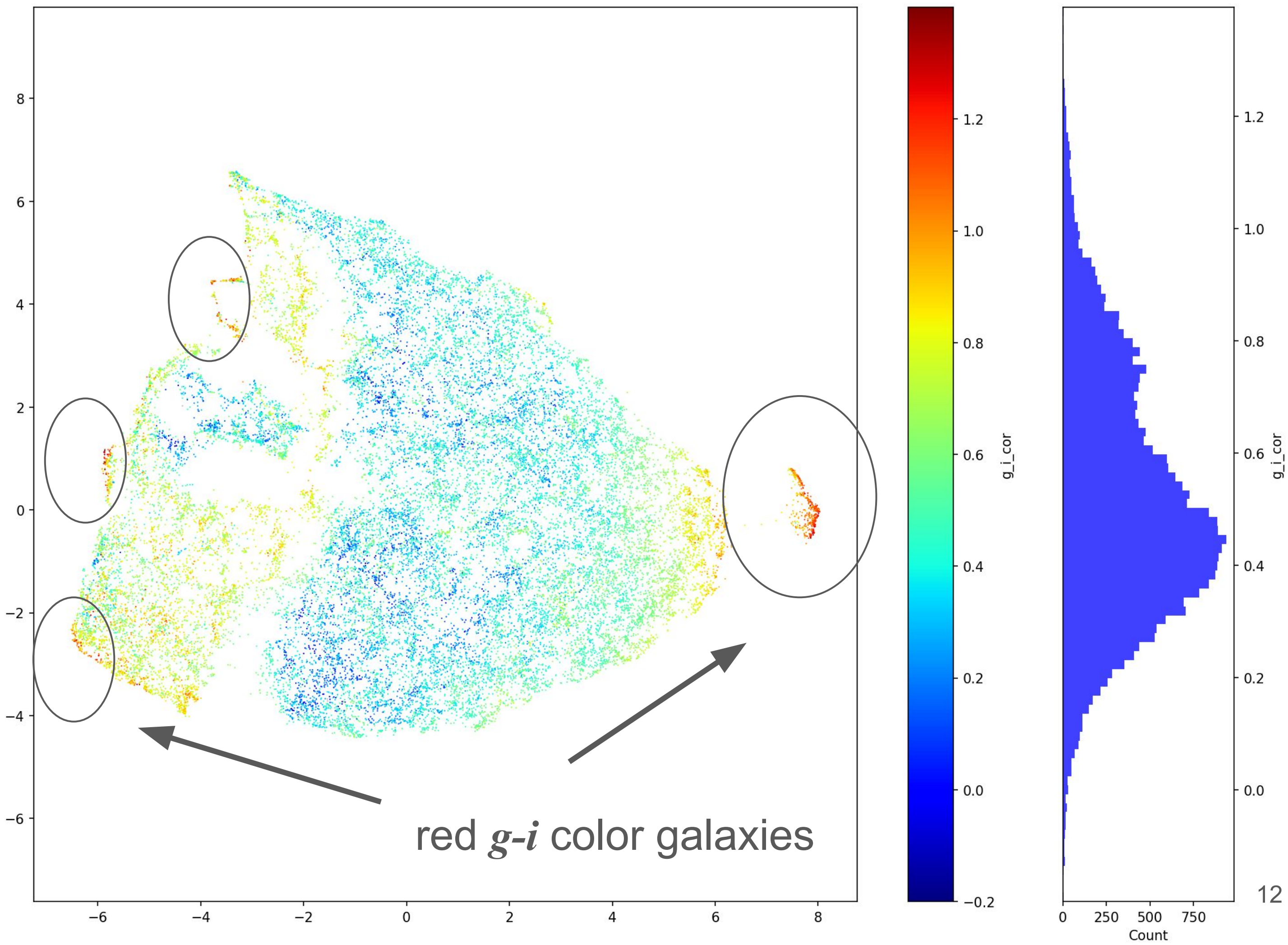
Embedding of the combined Tanglitis and Thuruthipilly catalogs

Visualizing **morphological** and **photometric** parameters using color scale

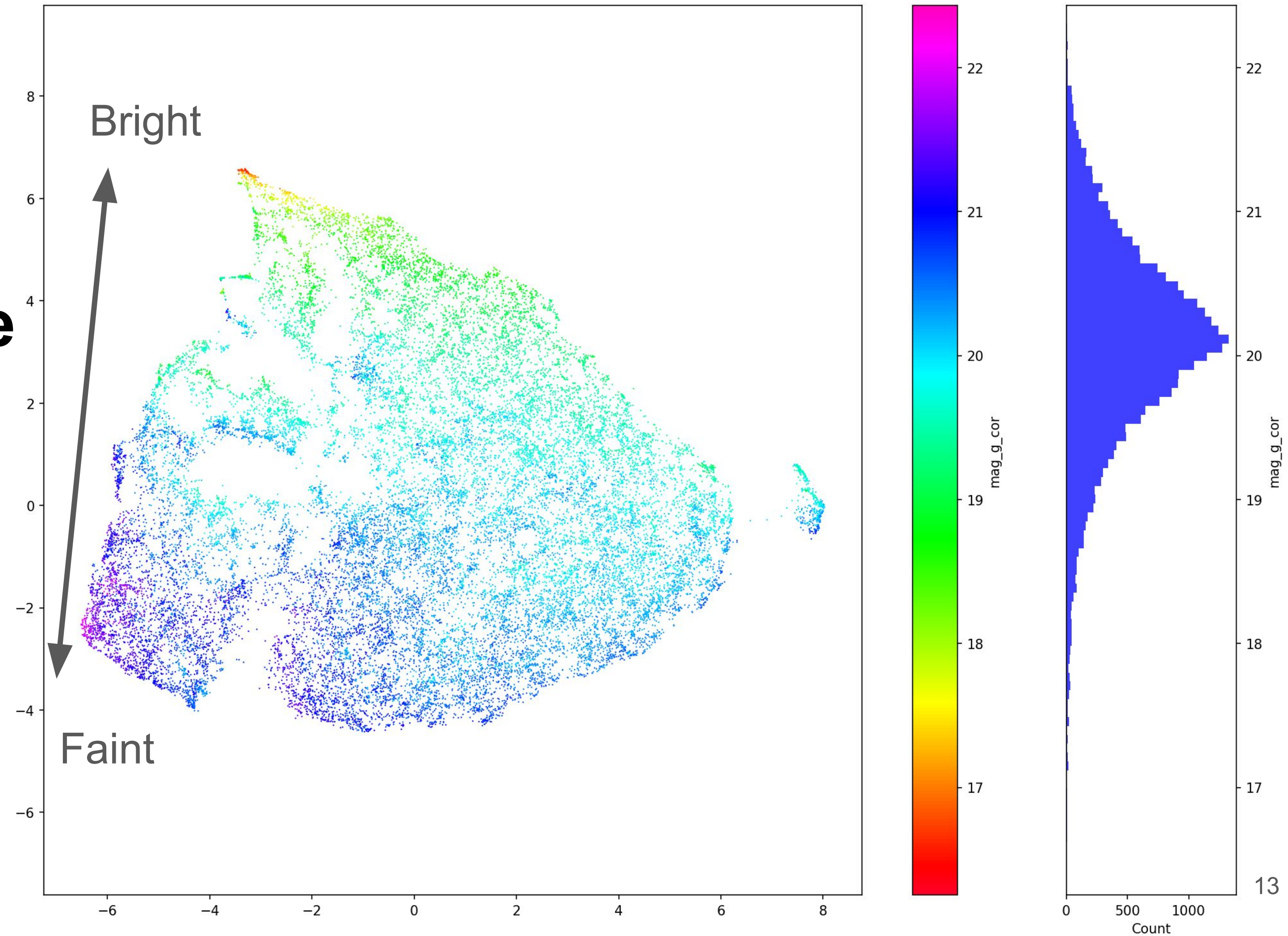


Visualizing properties:

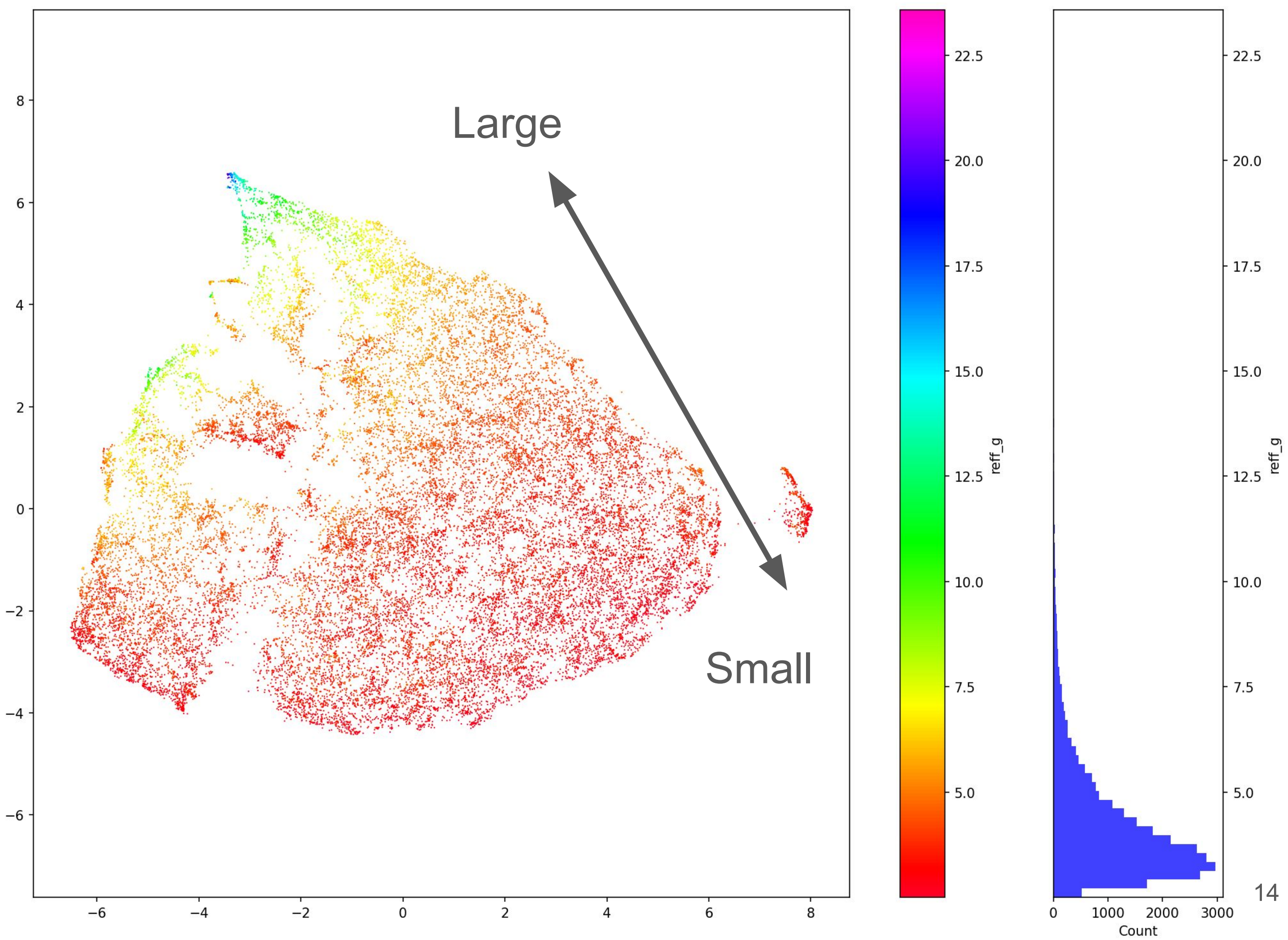
g-i color



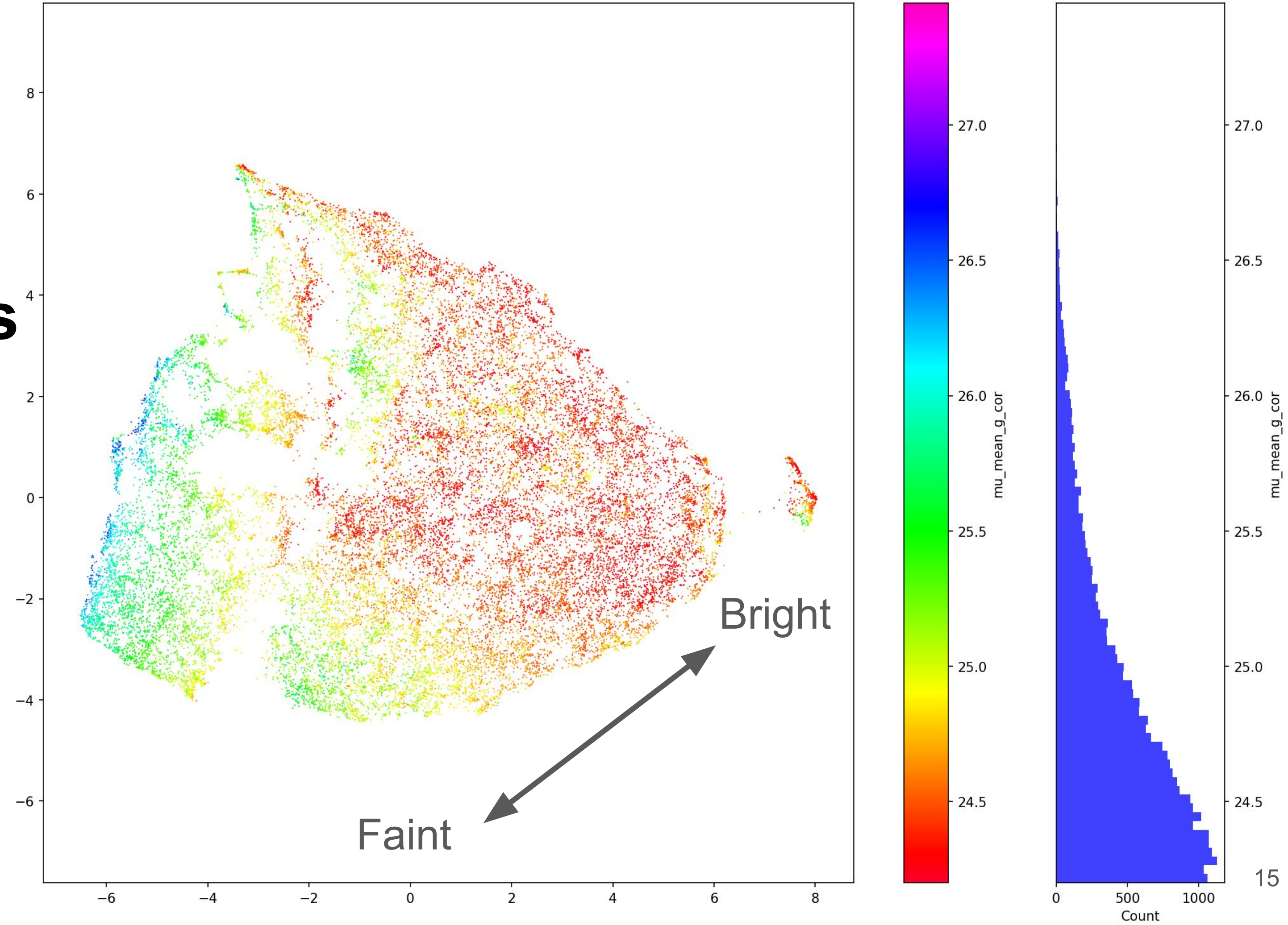
Visualizing properties: **g-band magnitude**



Visualizing properties: **half-light radius**

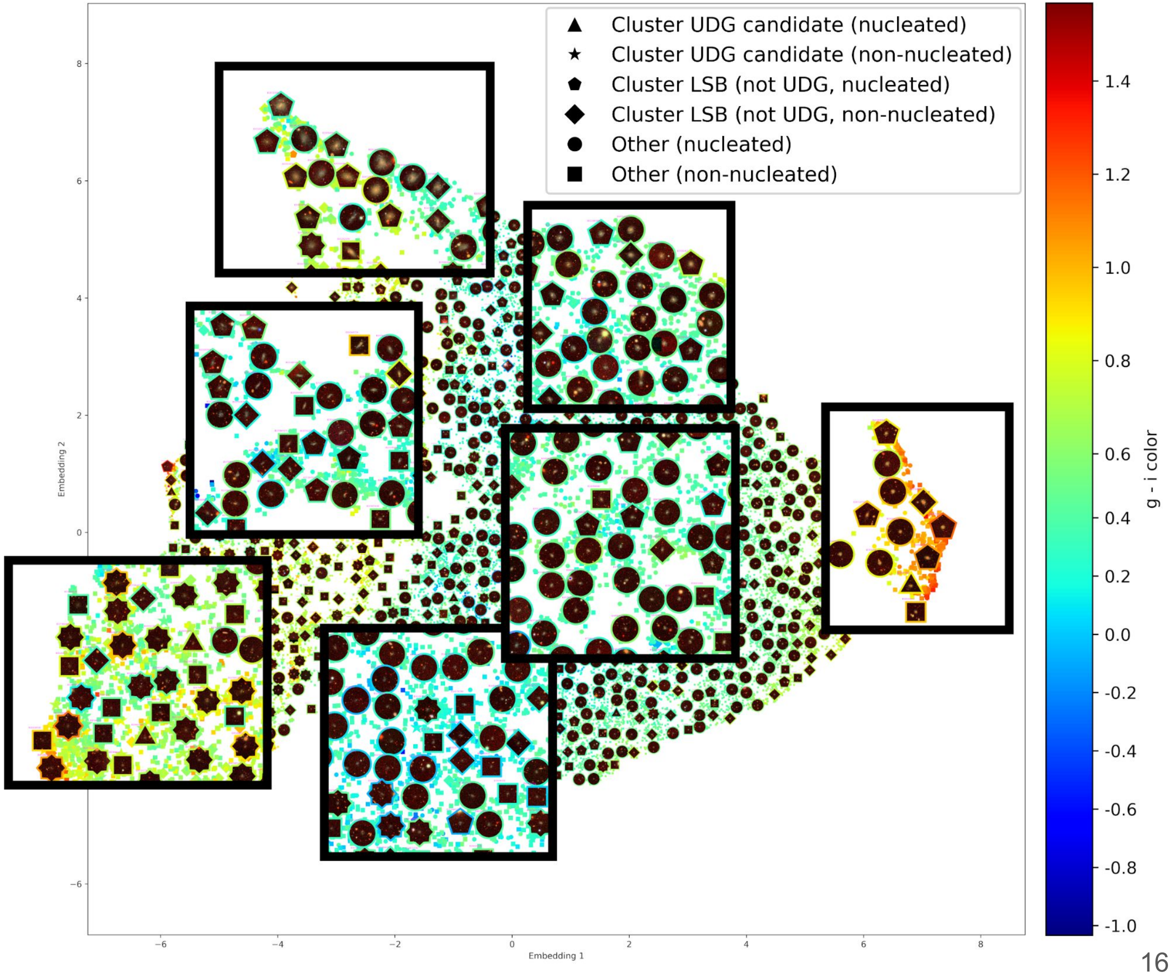


Visualizing properties: **Surface brightness**



Embedding of the combined Tanglitis and Thuruthipilly catalogs: **Relating galaxis with additional properties**

- Highlighting **ultra diffuse galaxies** (UDG)
- Highlighting presence of **nucleus**

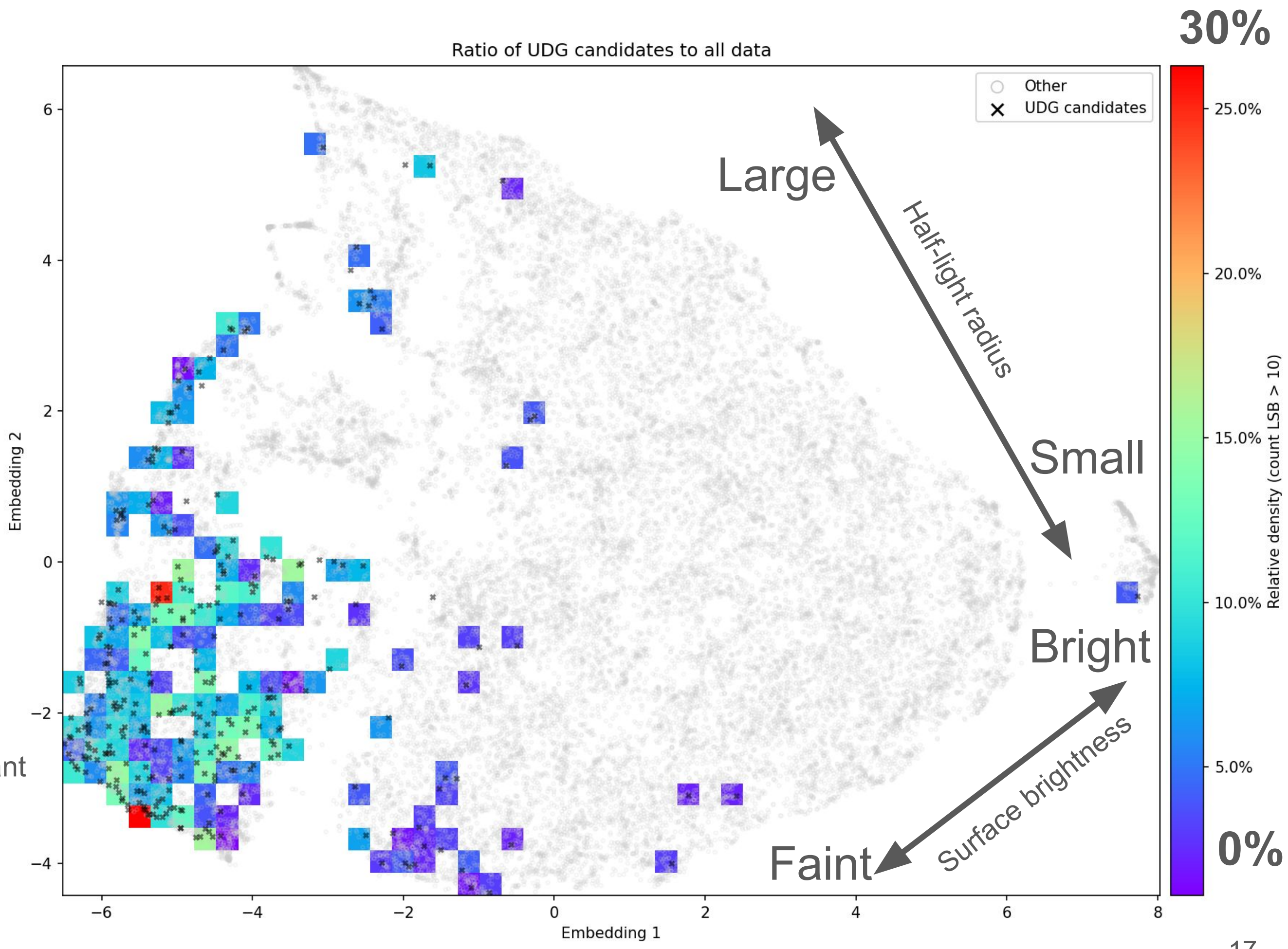


Additional galaxy properties: **Ultra diffuse galaxies**

Only available for galaxies associated with clusters (reliable redshifts are required)

UDG:

- **Low stellar masses**: (7.5 ≤log(M⋆/M☉)≤9)
- **Very diffuse** - up to a significant fraction of the Milky Way size
- Consist of even **99% of dark matter** or **less than 1%**

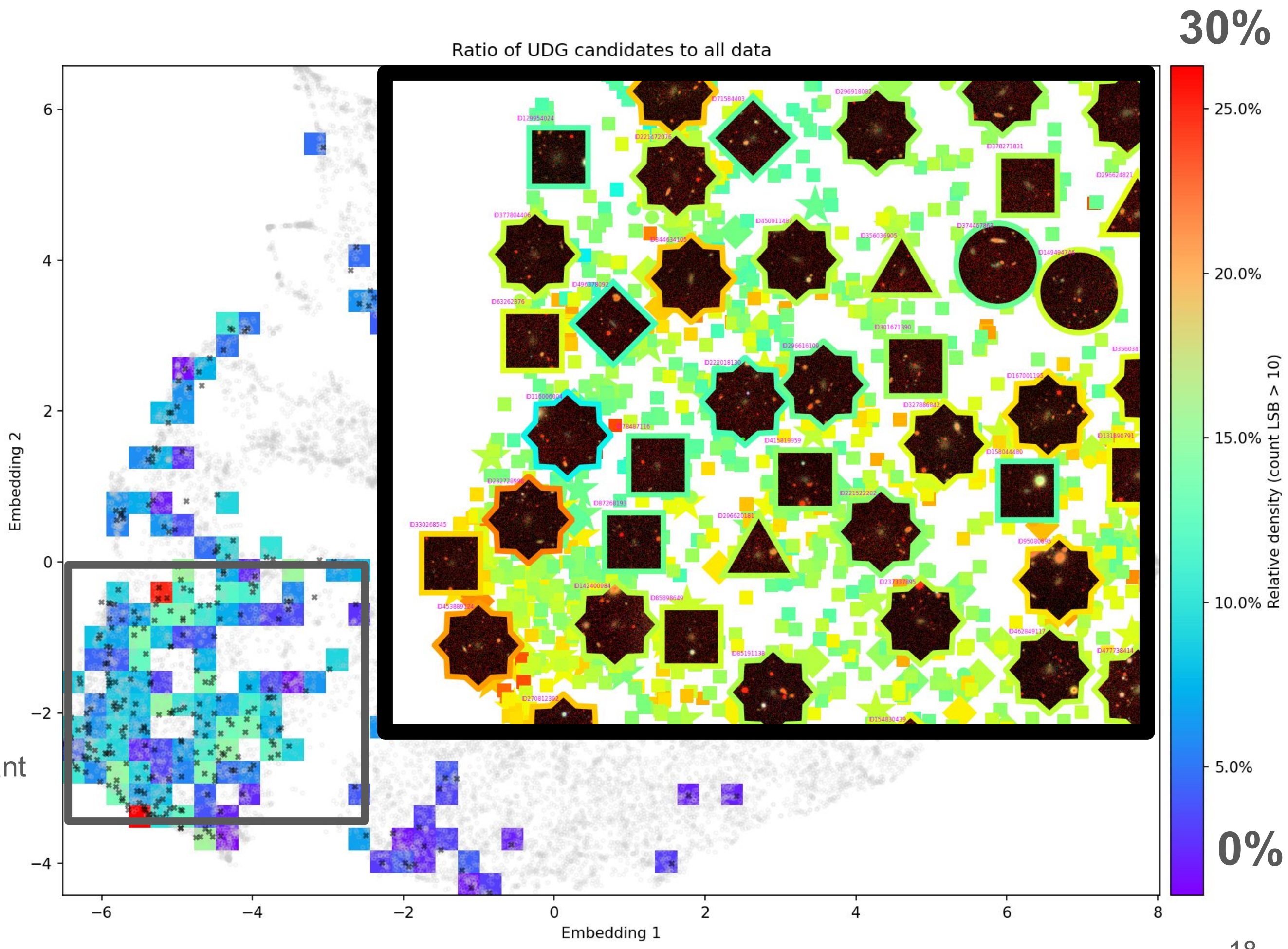


Additional galaxy properties: **Ultra diffuse galaxies**

Only available for galaxies associated with clusters (reliable redshifts are required)

UDG:

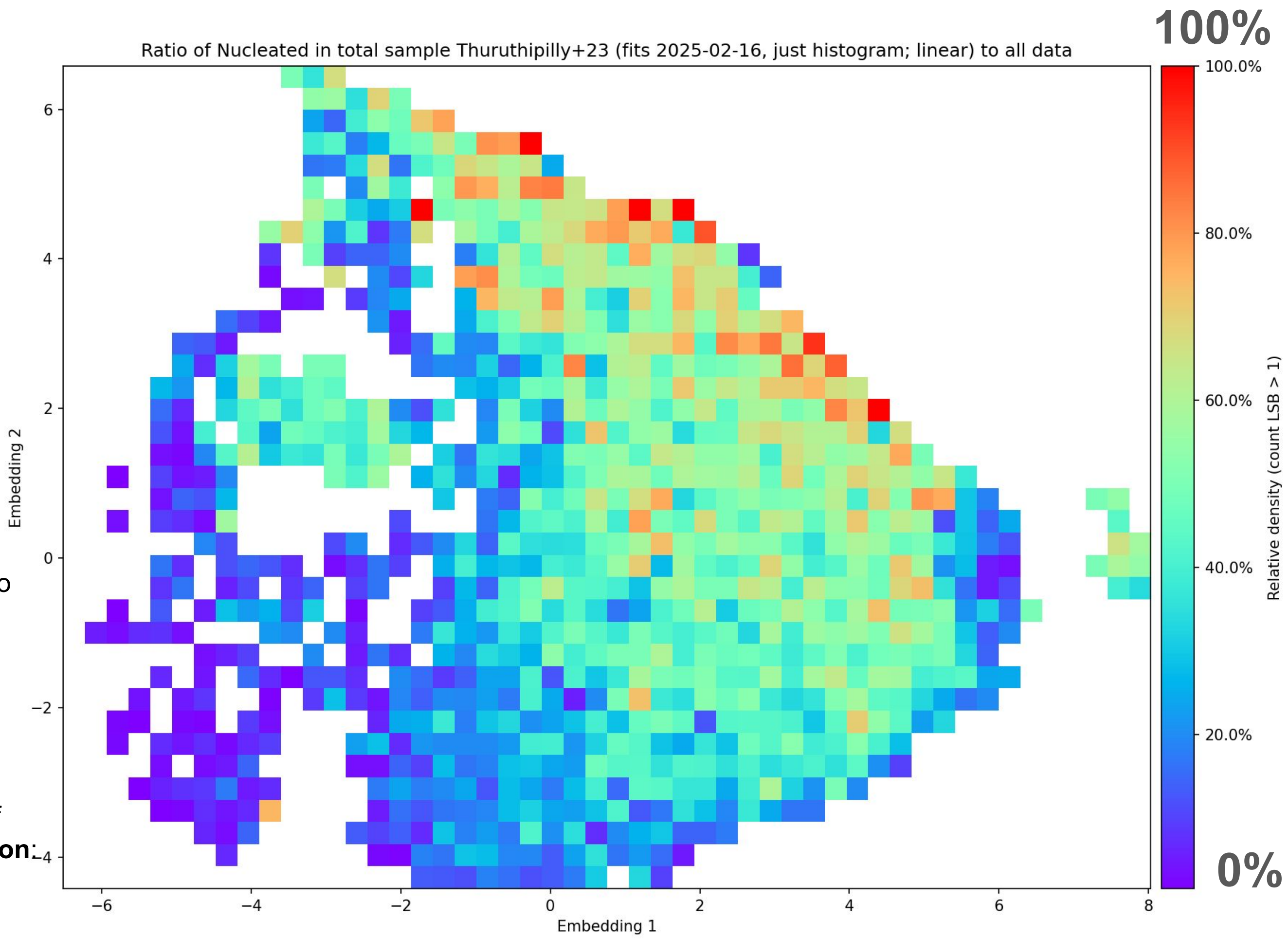
- **Low stellar masses**: ($7.5 \leq \log(M_\star/M_\odot) \leq 9$)
- **Very diffuse** - up to a significant fraction of the Milky Way size
- Consist of even **99% of dark matter** or **less than 1%**



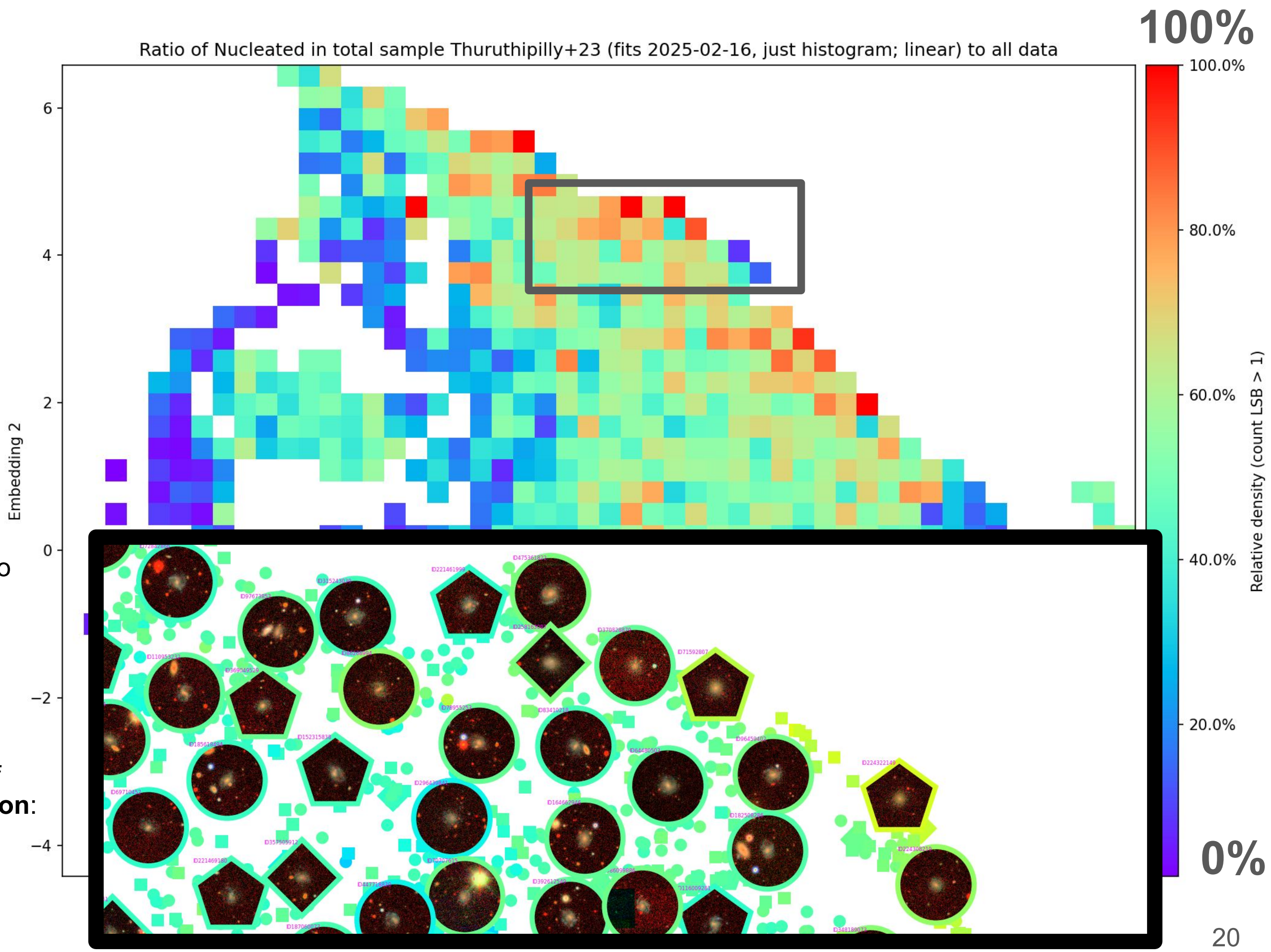
Additional galaxy properties: **Presence of nucleus**

Galaxy Evolution: Reveals how central nuclei shape galaxy growth.
Black Holes: Links nuclear star clusters to black hole formation.
Environment: Distinguishes external vs. internal influences on dwarfs.

Core/cusp problem: Observed galaxy rotation curves conflict with dark matter simulations. **Observation**: LSB and dwarf galaxies have flat-density cores. **Simulation**: Predicts steep, cuspy density profiles.



Additional galaxy properties: **Presence of nucleus**



Galaxy Evolution: Reveals how central nuclei shape galaxy growth.
Black Holes: Links nuclear star clusters to black hole formation.
Environment: Distinguishes external vs. internal influences on dwarfs.

Core/cusp problem: Observed galaxy rotation curves conflict with dark matter simulations. **Observation**: LSB and dwarf galaxies have flat-density cores. **Simulation**: Predicts steep, cuspy density profiles. (de Blok, 2010)

Re-analysis of all candidates found by Thuruthipilly’s methods

- Evaluate nucleation by **comparing quality of morphology fits**
  - *(Sersic profile fitting using Galfit, based on Yagi 2016 method)*
  - Comparison of BIC metric between **single and double-component** morphological fits
- **Filter** by:
  - fit quality in g-band (Reduced χ^2, passing, range, relative error),
  - LSB definition (g-band)
  - fit quality (other bands)

Fit success ☑️, Range ☑️, Fit quality ☑️, Low error ☑️, LSB fit ☑️, i-band acceptable ☑️

Dataset	Paper	Nucleus?	Count
New candidates		No	**11,567**
		Yes	**9,705**
Thuruthipilly ∪ Tanoglitis	Thur. 2024	No	1,834
		Yes	1,965
	Tanoglitis 2021	No	11,511
		Yes	7,539

21,272 LSB cantidates

Fit success ☑️, Range ☑️, Fit quality ☑️, Low error ☑️, LSB fit ☑️, i-band acceptable ☑️, LSB SourceExtractor ☑️

Dataset	Paper	Nucleus?	Count
New candidates		No	**865**
		Yes	**824**
Thuruthipilly ∪ Tanoglitis	Thur. 2024	No	1,834
		Yes	1,965
	Tanogitis 2021	No	11,032
		Yes	6,559

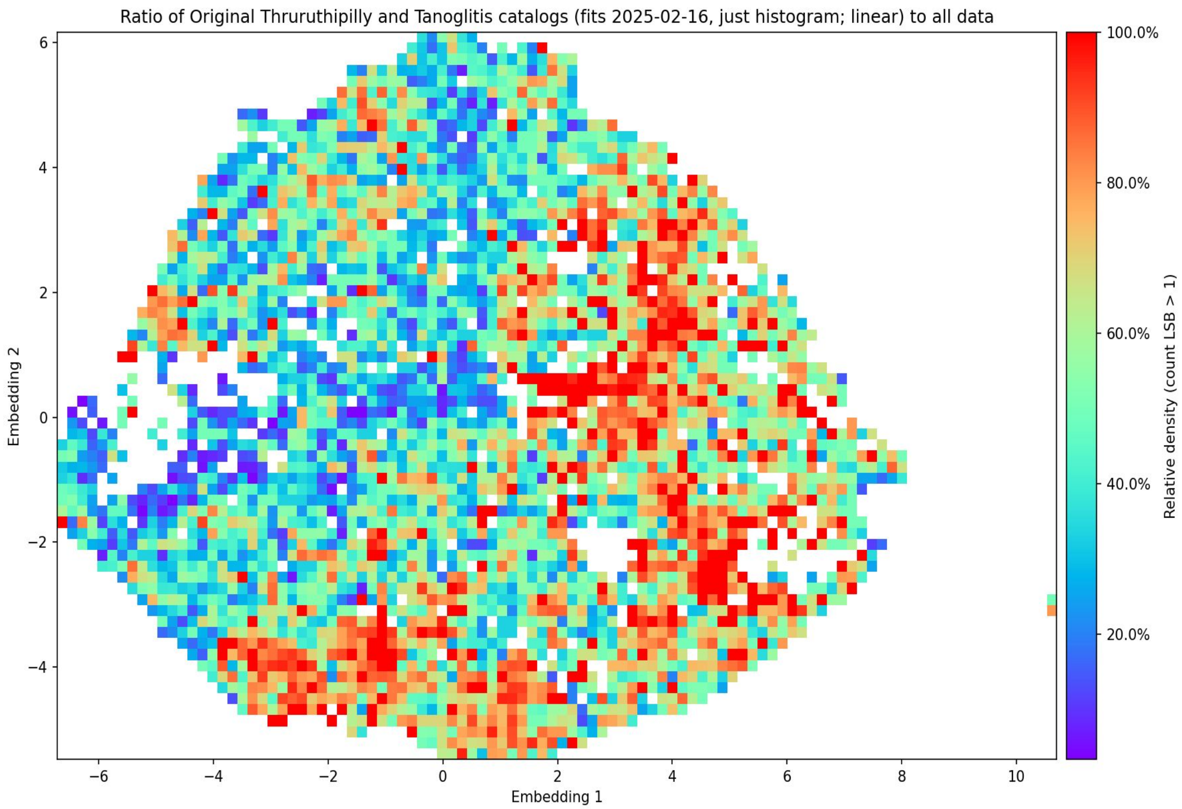
1,689 LSB candidates

Yagi, Masafumi, et al. “Catalog of Ultra-Diffuse Galaxies in the Coma Cluster from Subaru Imaging Data.” *The Astrophysical Journal Supplement Series*, vol. 225, no. 1, 2016, p. 11, https://doi.org/10.3847/0067-0049/225/1/11

Re-analysis of all candidates found by Thuruthipilly's methods

$$\frac{N_{\text{Thuruthipilly}} + N_{\text{Tanoglitis}}}{N_{\text{Thuruthipilly}} + N_{\text{Tanoglitis}} + N_{\text{new}}}$$

Low ratio ⇒
More galaxies from the new dataset



Conclusion

- Understanding of LSBs is critical for a **complete picture of galaxy formation and evolution**
  - 30-60% of the galaxy population
- The work of our group significantly **increased the size of known LSB catalog**
  - + possible additional size increase due to refined fitting (potentially 1000s of new LSBs)
- The LSB catalog was extended with **nucleus presence information**
  - Useful for future analyses
  - Nucleating is **localized** in low dimensional embeddings of the data
- The methods developed using DES catalogue will be transferred to **LSST and Euclid**
- Embedding methods such as UMAP seem to be useful tool for visualizing diversity of LSB galaxies

Thank you

michal.vrabel@ncbj.gov.pl

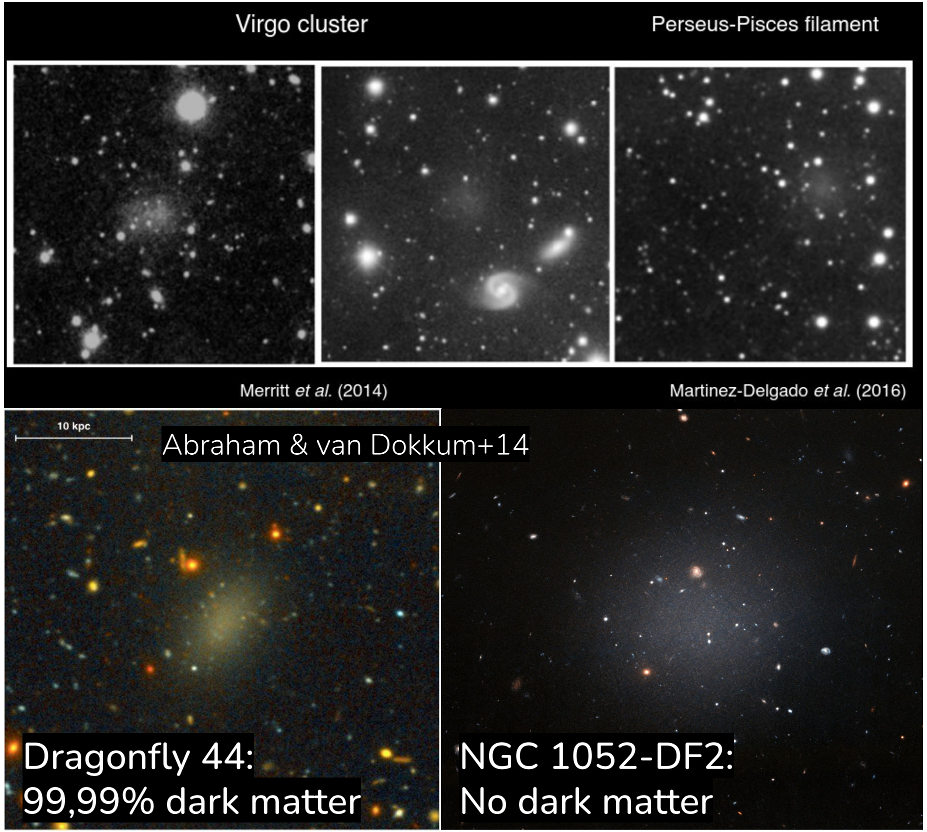
Ludwika Pasteura 7, 02-093 Warszawa

NGC 1052-DF2
Abraham & van Dokkum+14

INTRODUCTION What are subtypes of low surface brightness galaxies?

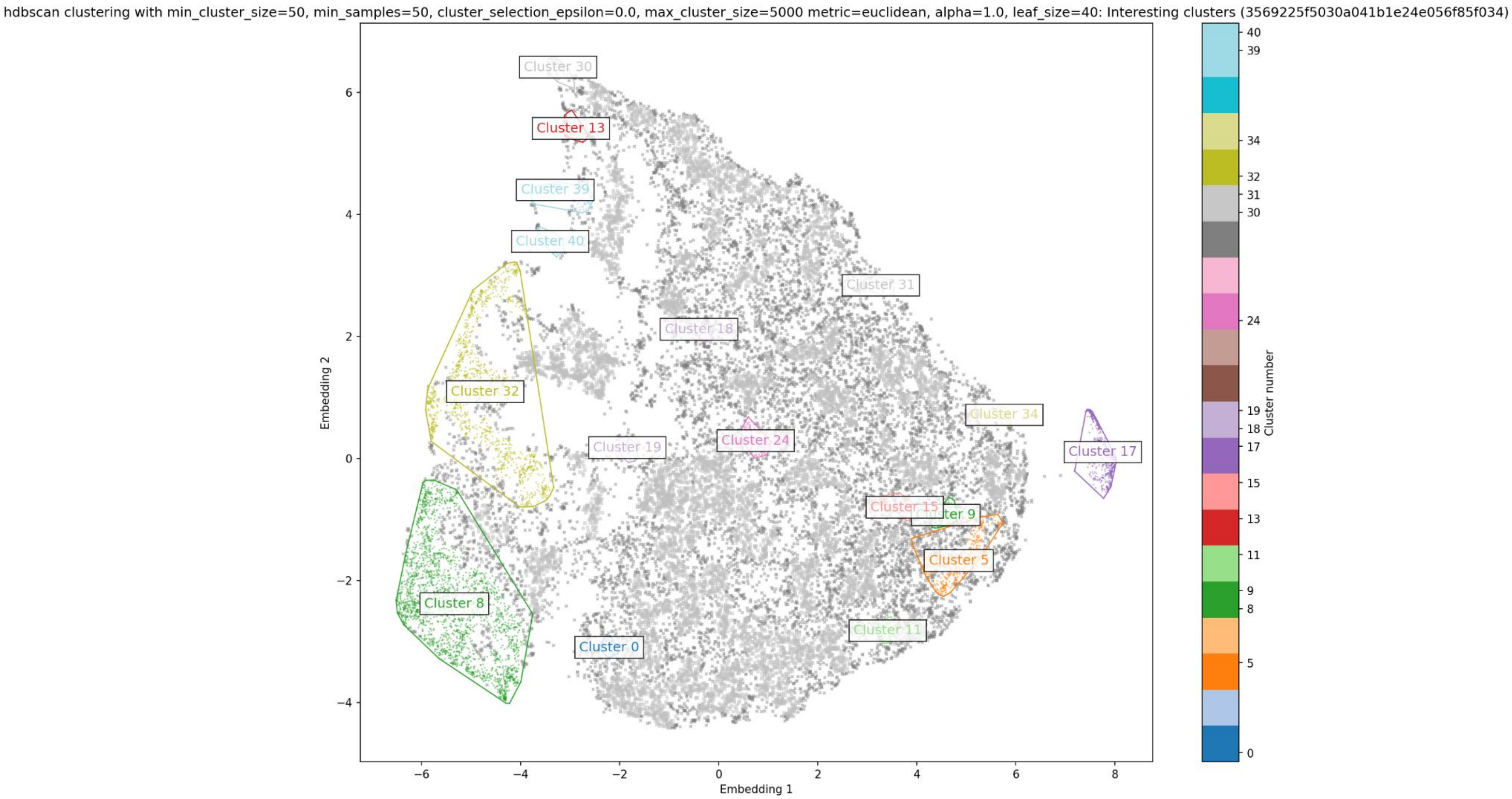
Ultra Diffuse Galaxies (UDG)

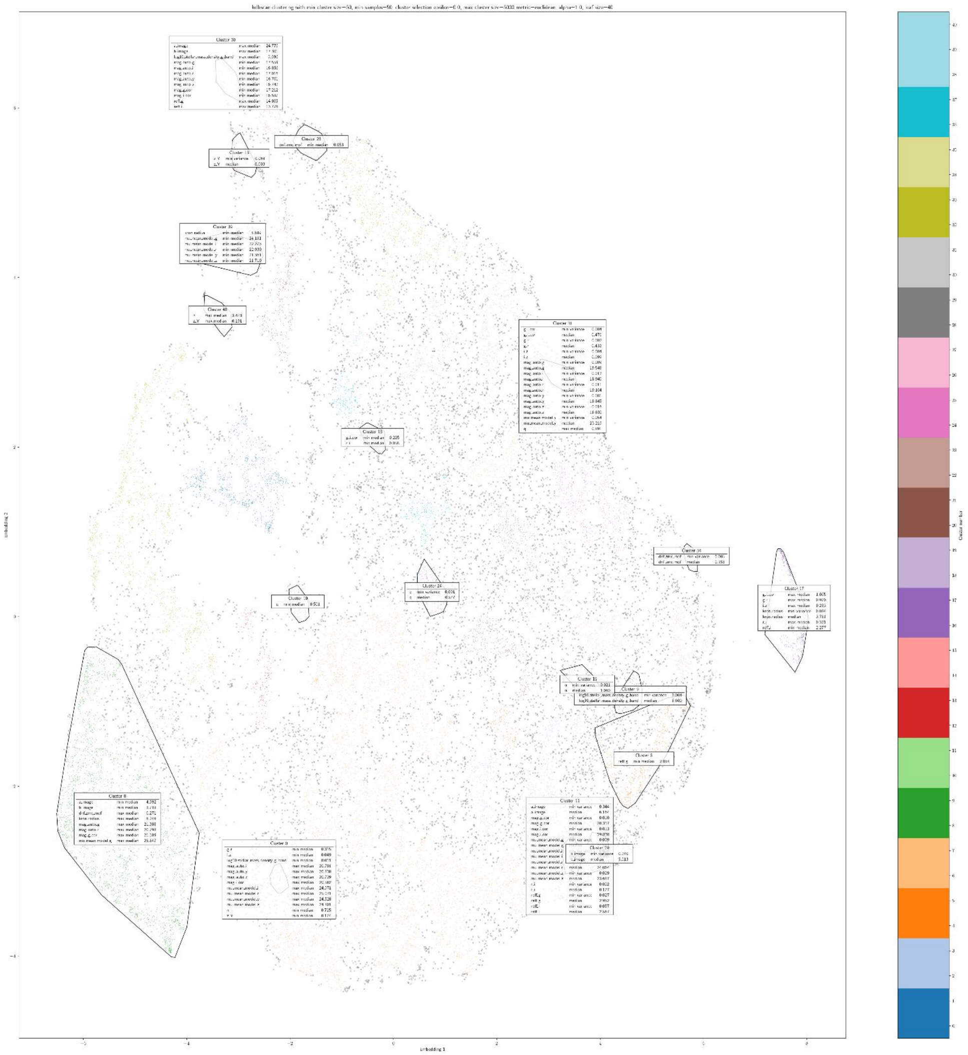
- **Low stellar masses** ($7.5 \le \log(M_\star/M_\odot) \le 9$),
- **Very diffuse** - up to a significant fraction of the Milky Way size
- Mostly found in clusters → but not only (debate still ongoing if those found elsewhere can be classified as UDGs)
- They can consist of even
  - **99% of dark matter** or
  - **Less than 1%**

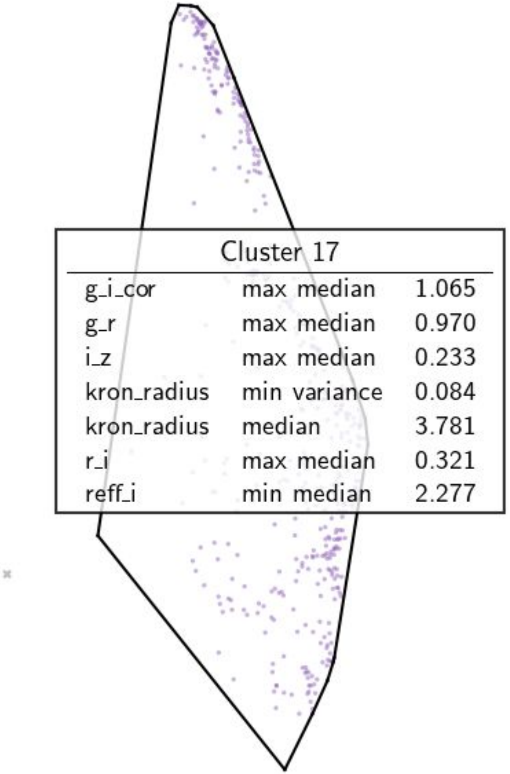


Clustering of LSBs

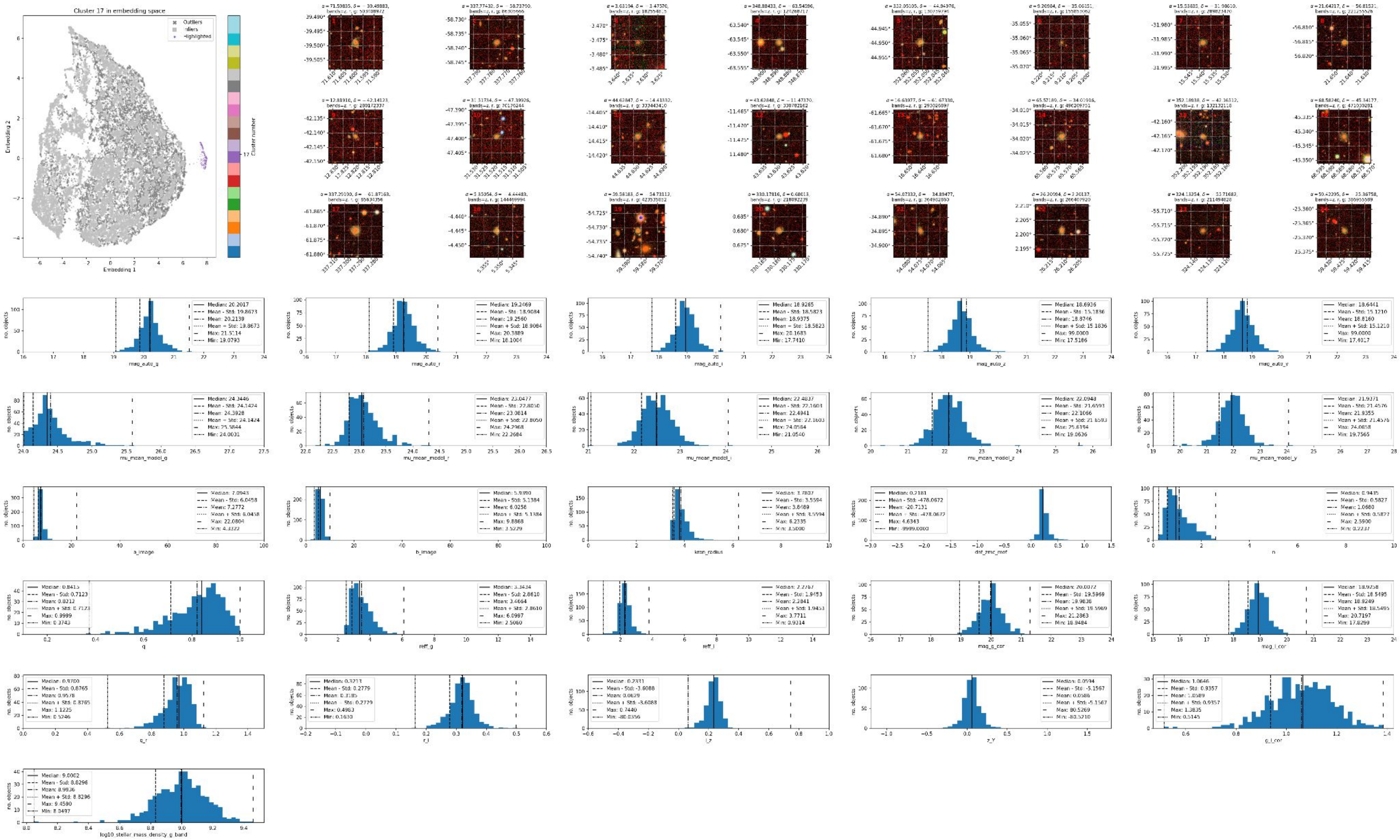
- The method is still under development
- Multiple routes
  - Global / local features
  - Embedding before clustering / Embedding and clustering optimized together
  - Fixed number of clusters / Density-based


hdbscan clustering with min_cluster_size=50, min_samples=50, cluster_selection_epsilon=0.0, max_cluster_size=5000 metric=euclidean, alpha=1.0, leaf_size=40: Interesting clusters (3569225f5030a041b1e24e056f85f034)
Cluster 30
Cluster 13
Cluster 39
Cluster 40
Cluster 31
Cluster 18
Cluster 32
Cluster 34
Cluster 17
Cluster 19
Cluster 24
Cluster 15
Cluster 9
Cluster 5
Cluster 8
Cluster 11
Cluster 0
Embedding 1
Embedding 2
Cluster number

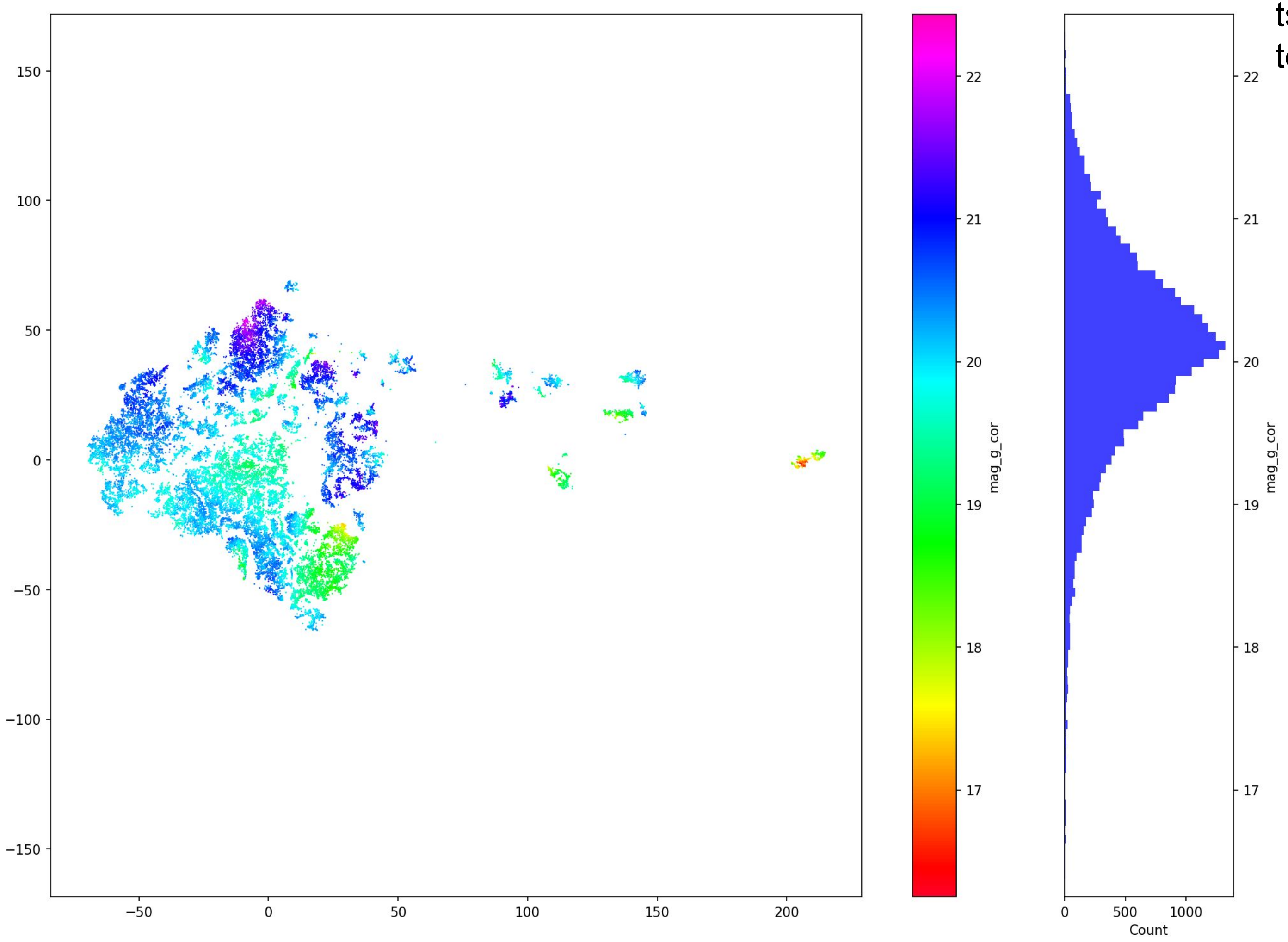


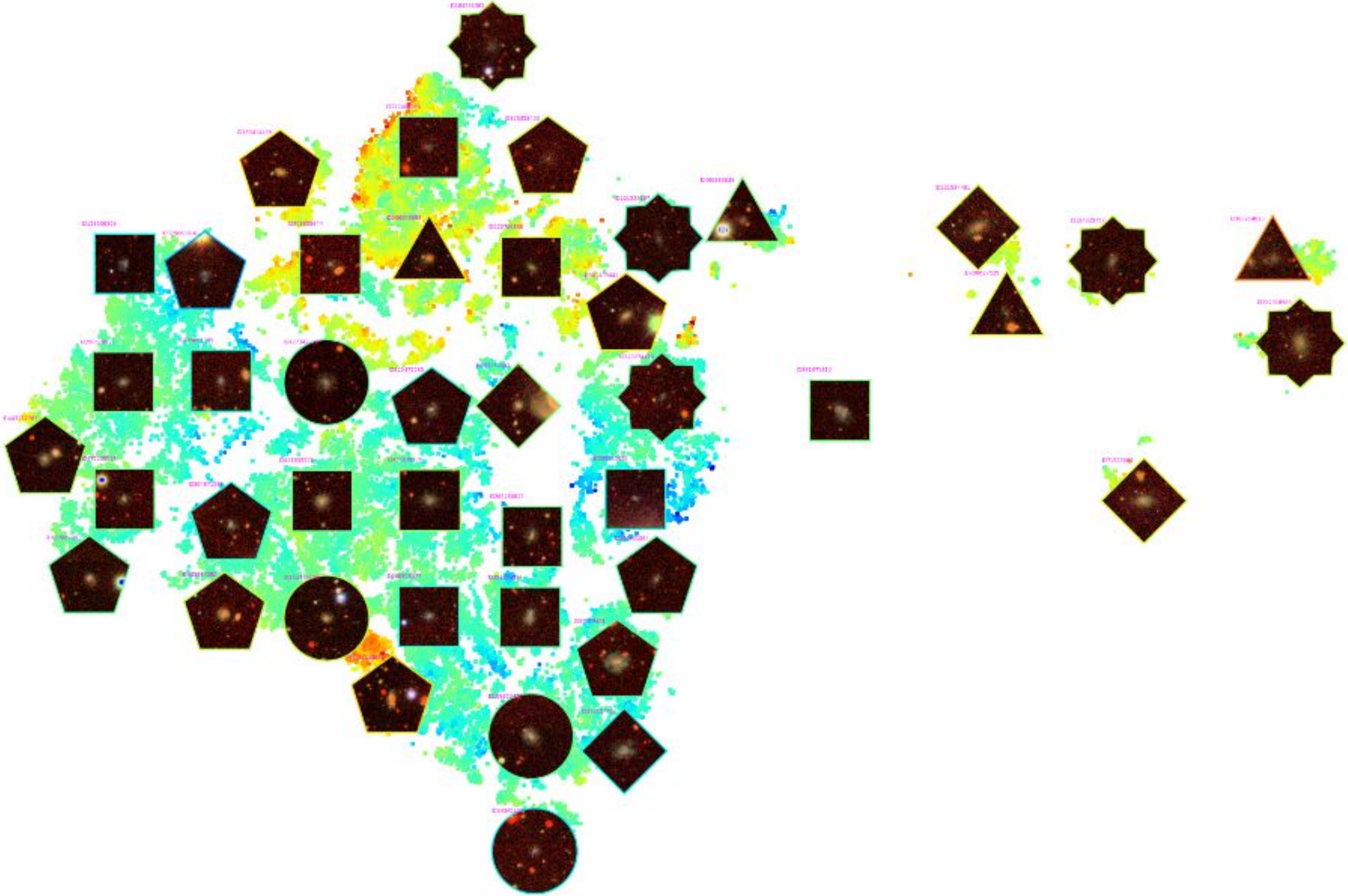


Cluster 17
g_i_cor max median 1.065
g_r max median 0.970
i_z max median 0.233
kron_radius min variance 0.084
kron_radius median 3.781
r_i max median 0.321
reff_i min median 2.277

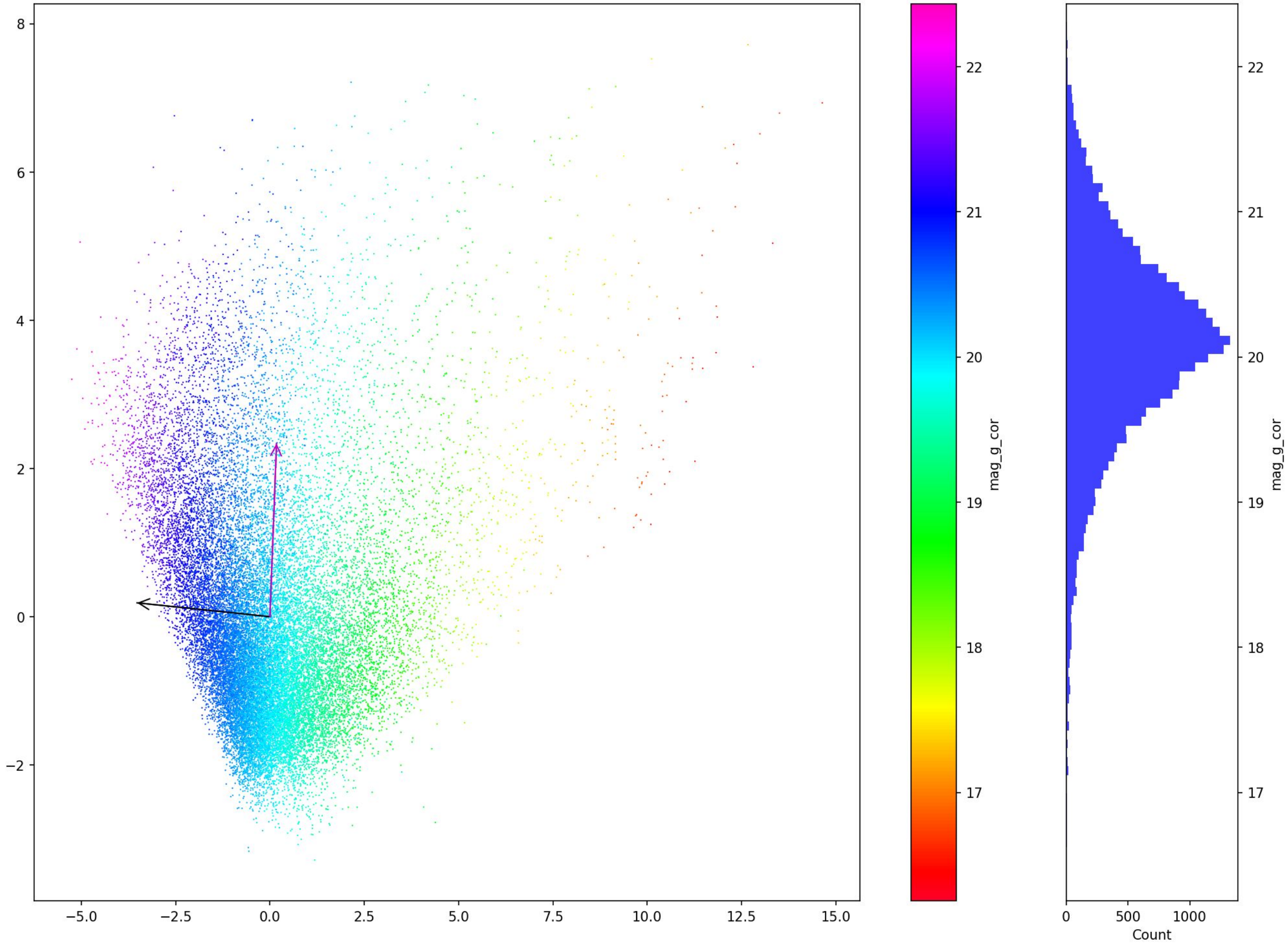

Cluster 17 in embedding space
Outliers
Inliers
Highlighted
Embedding 1
Embedding 2
Cluster number
mag_auto_g
mag_auto_r
mag_auto_i
mag_auto_z
mag_auto_y
mu_mean_model_g
mu_mean_model_r
mu_mean_model_i
mu_mean_model_z
mu_mean_model_y
a_image
b_image
kron_radius
n
q
reff_g
reff_i
mag_g_cor
mag_i_cor
g_r
r_i
i_z
z_Y
g_i_cor
log10_stellar_mass_density_g_band

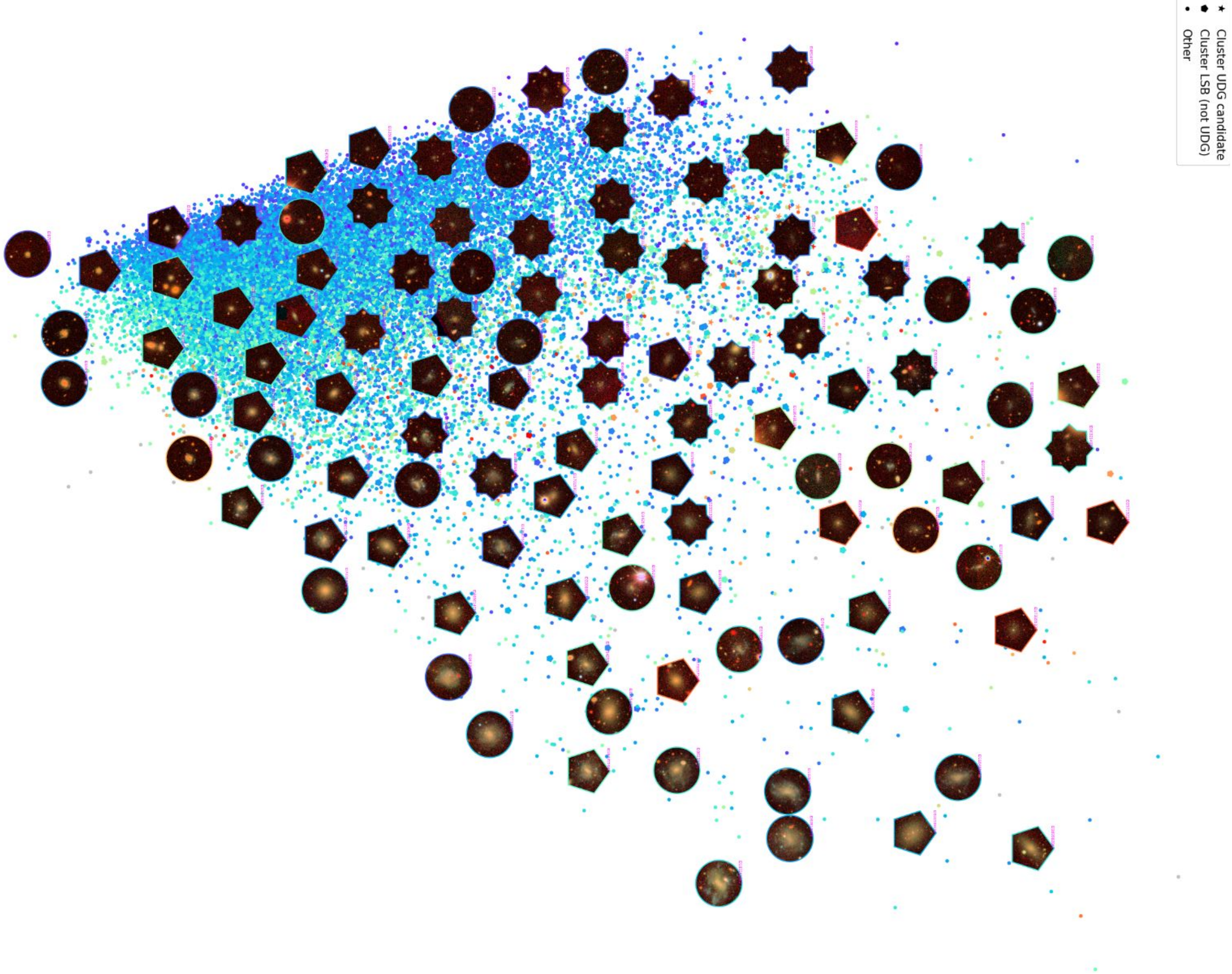
tsnecuda_perp_75_lr_10_n_iter_10000_66d7f33b



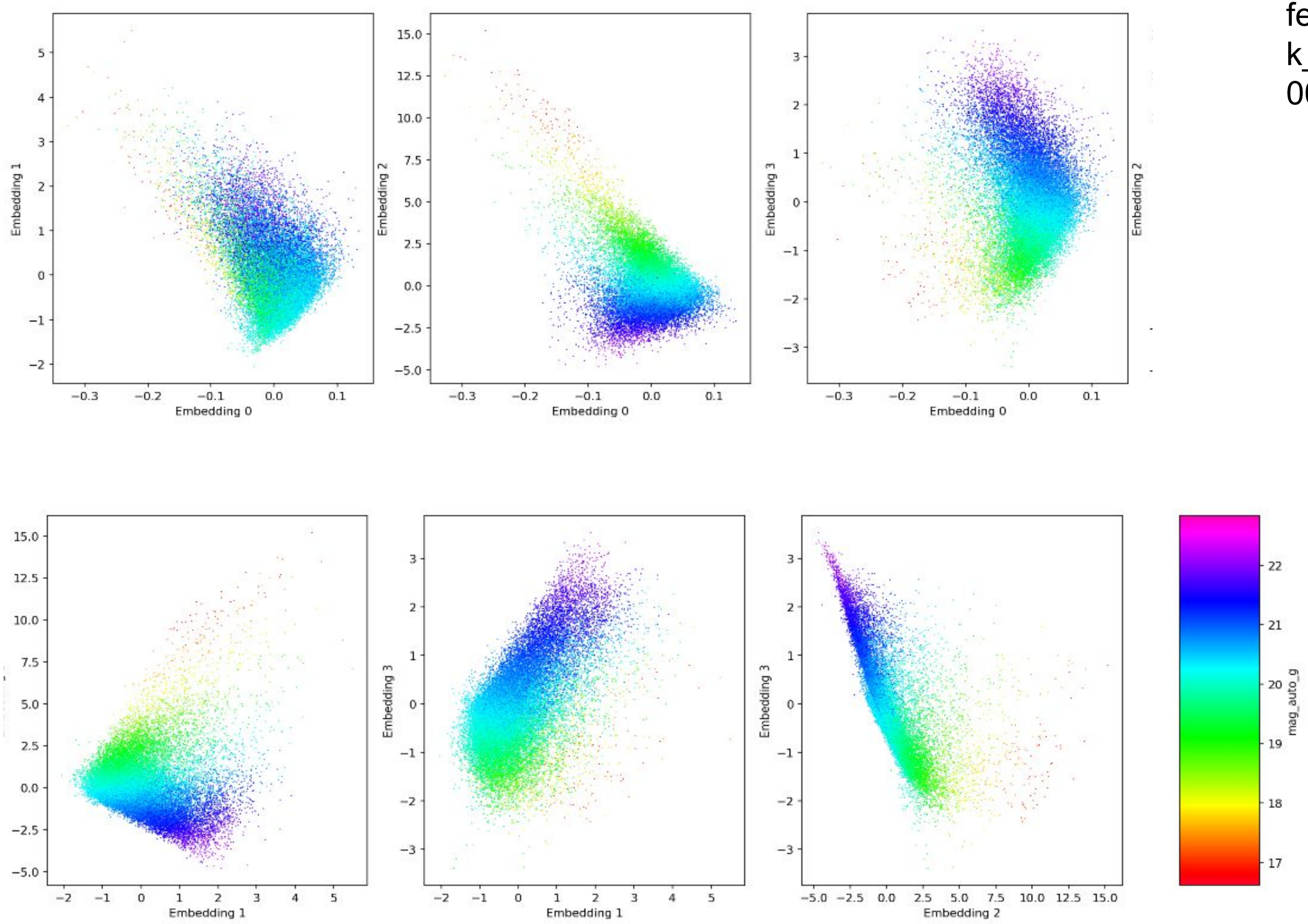


pca_n_2_51e95b15

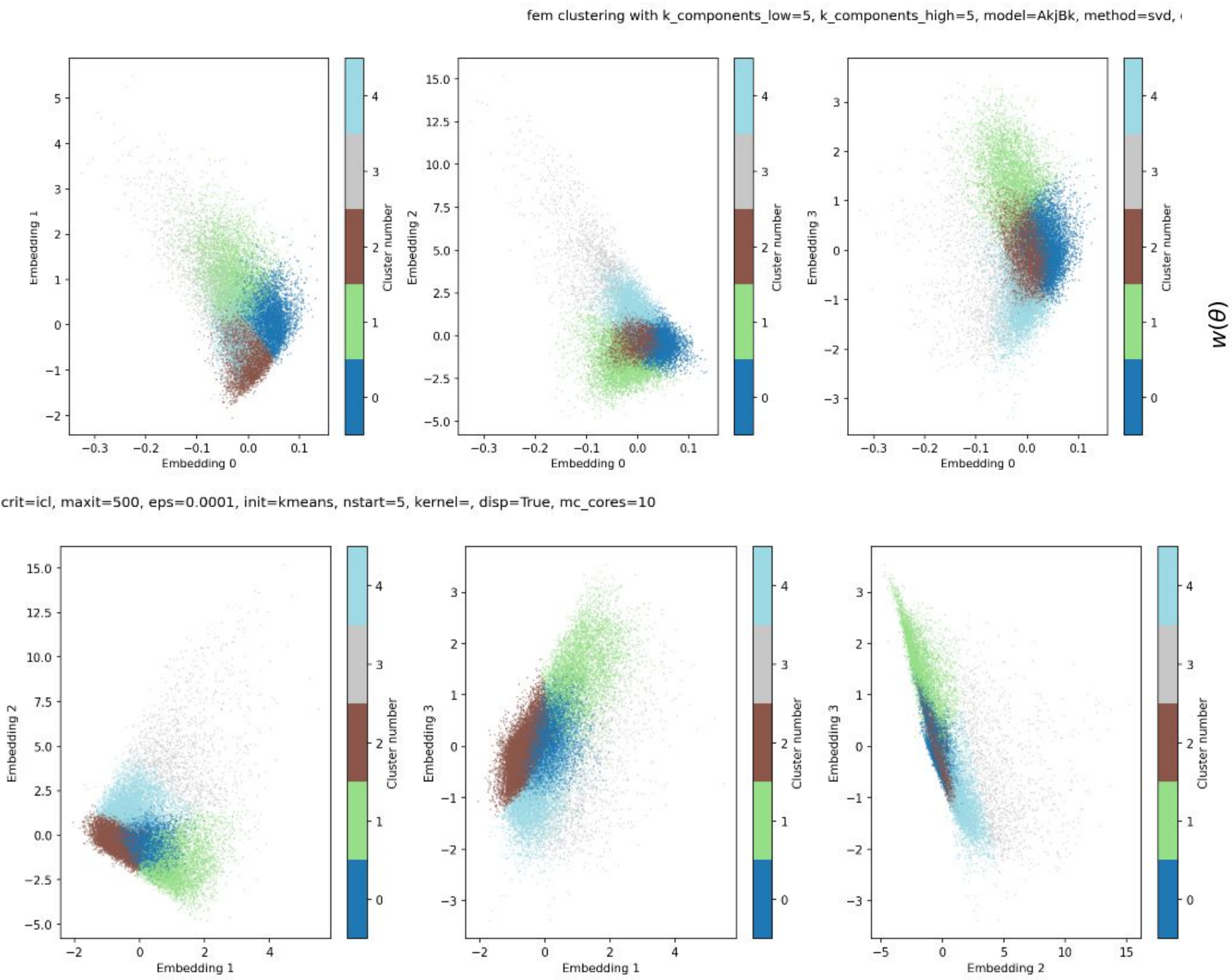


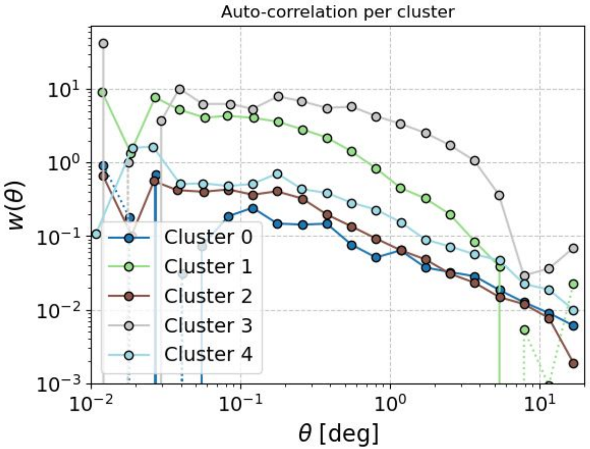

Cluster UDG candidate
Cluster LSB (not UDG)
Other

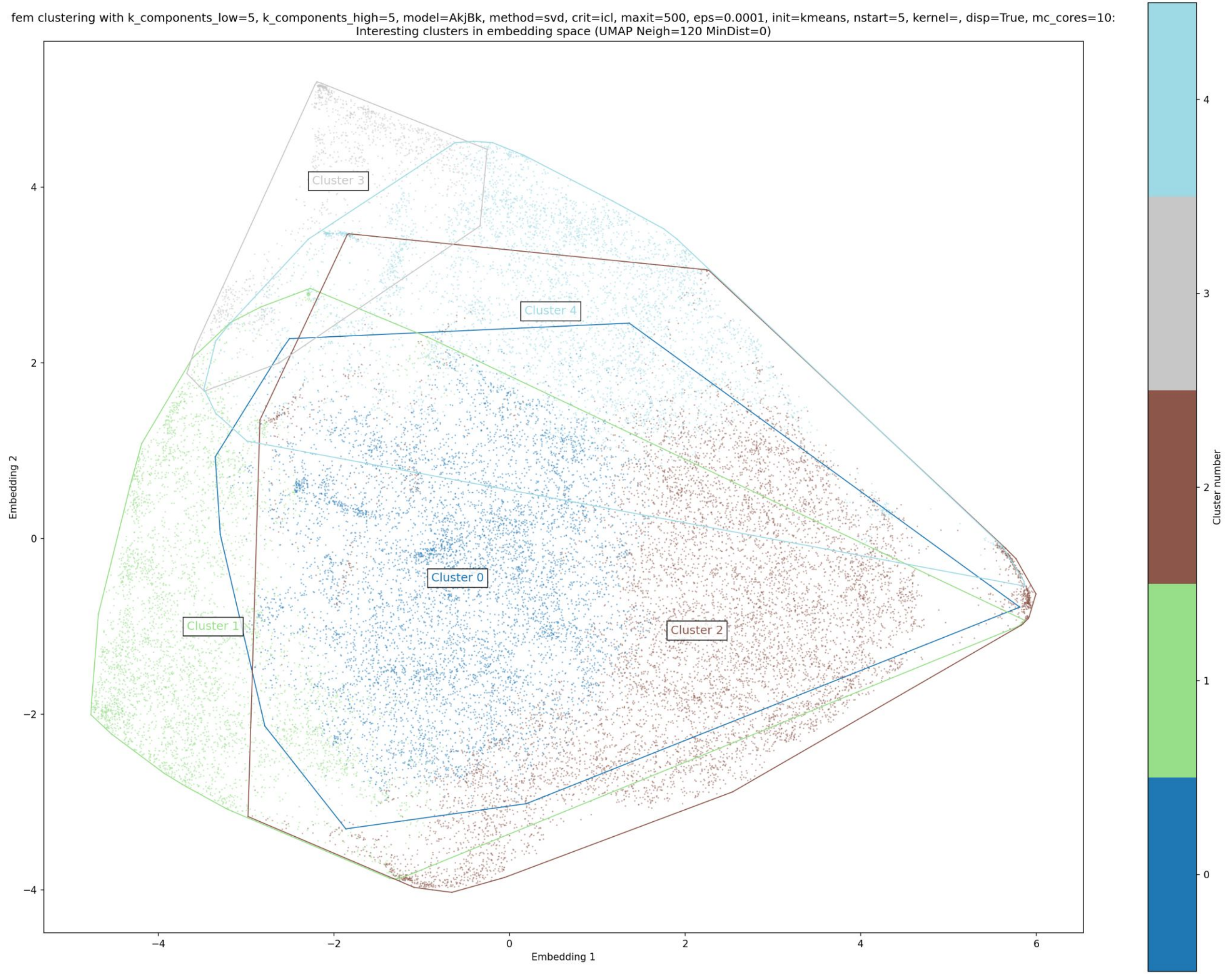
fem_k_5_5_model_AkjB
k_method_svd_maxit_5
00_d72662c1

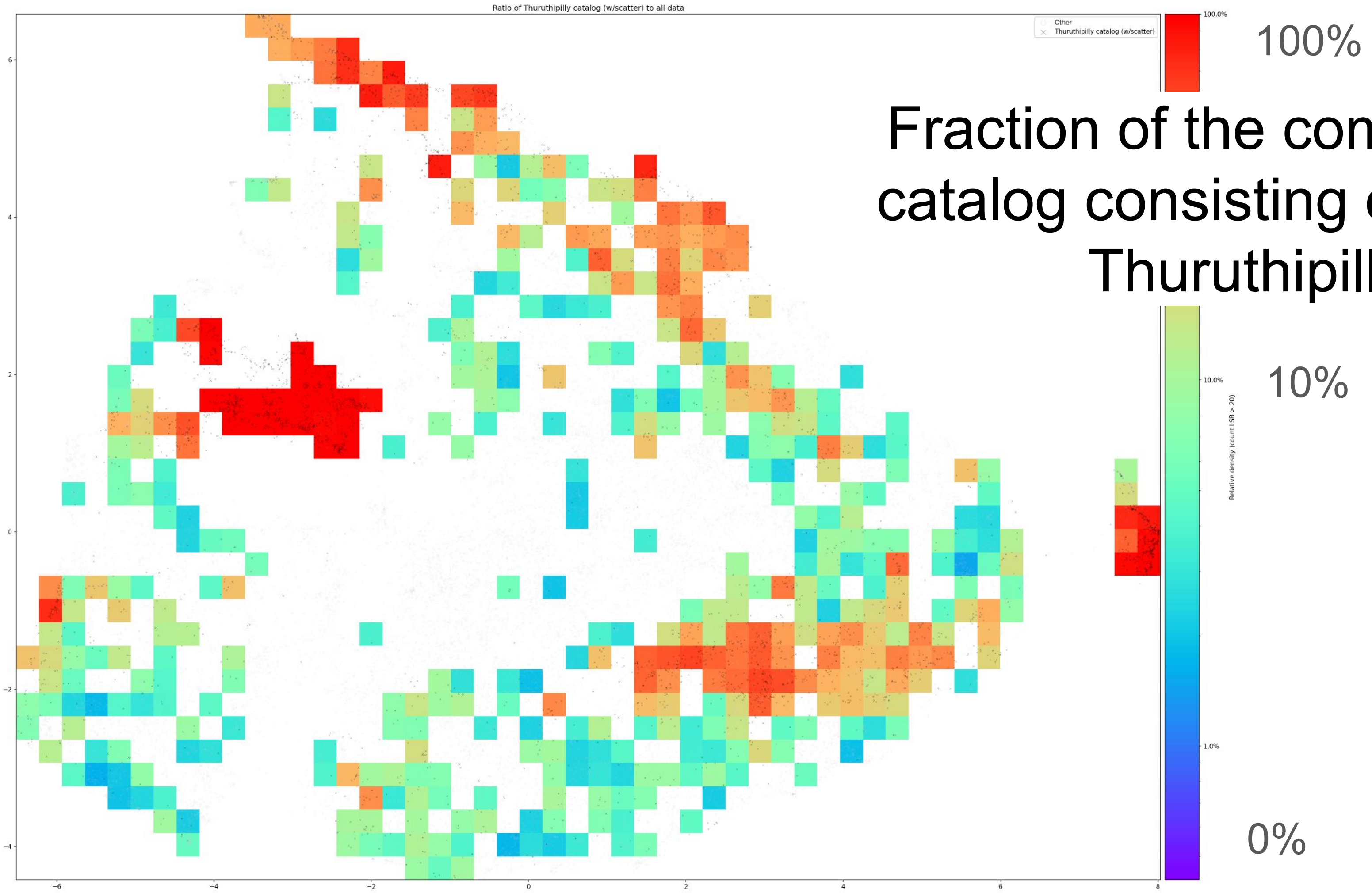


fem_k_5_5_model_AkjBk_met
hod_svd_maxit_500_d72662c1






fem clustering with k_components_low=5, k_components_high=5, model=AkjBk, method=svd, crit=icl, maxit=500, eps=0.0001, init=kmeans, nstart=5, kernel=, disp=True, mc_cores=10:
Interesting clusters in embedding space (UMAP Neigh=120 MinDist=0)
Cluster 3
Cluster 4
Cluster 0
Cluster 1
Cluster 2
Embedding 1
Embedding 2
Cluster number
4
3
2
1
0
−4
−2
0
2
4
6



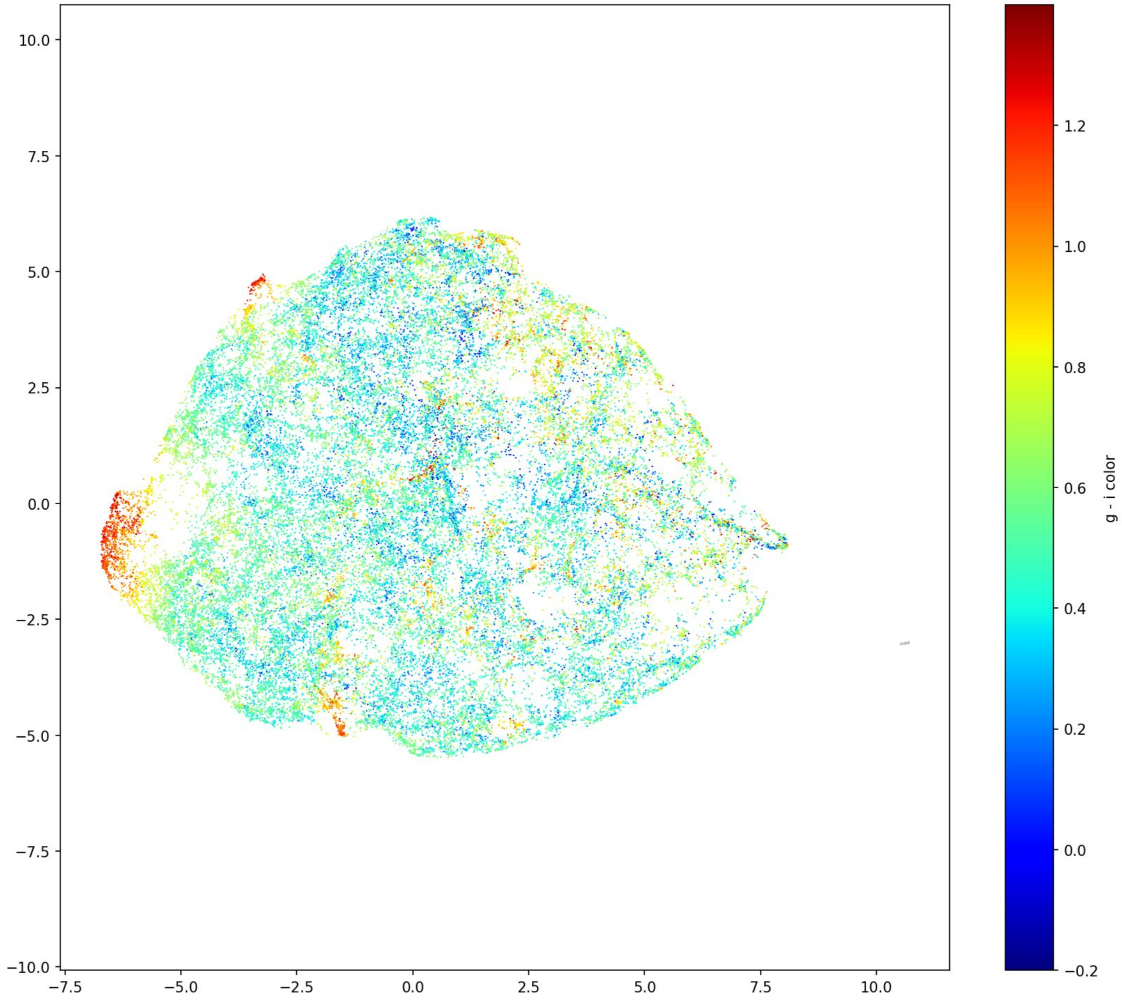
Fraction of the combined catalog consisting of new Thuruthipilly data

More about the extended dataset

Re-analysis of all candidates found by Thuruthipilly's methods

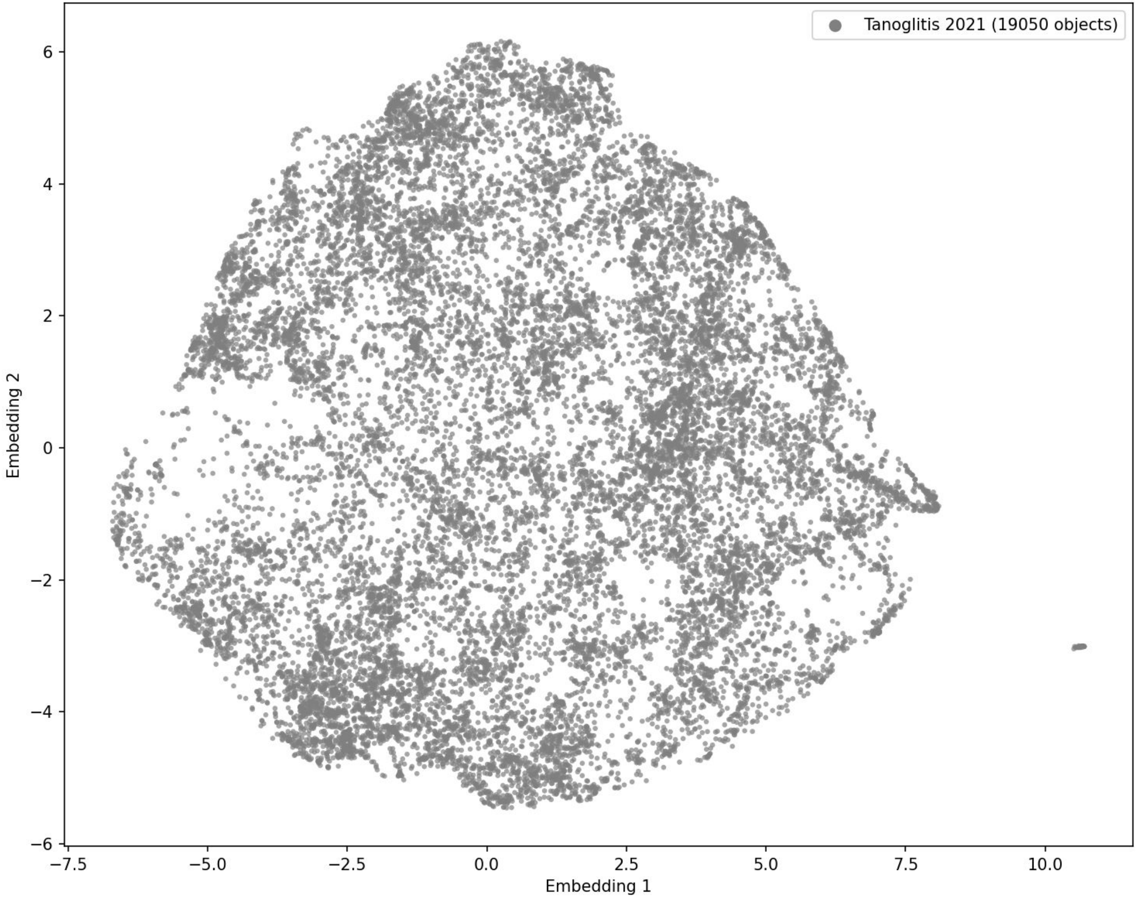
UMAP embedding

g - i color



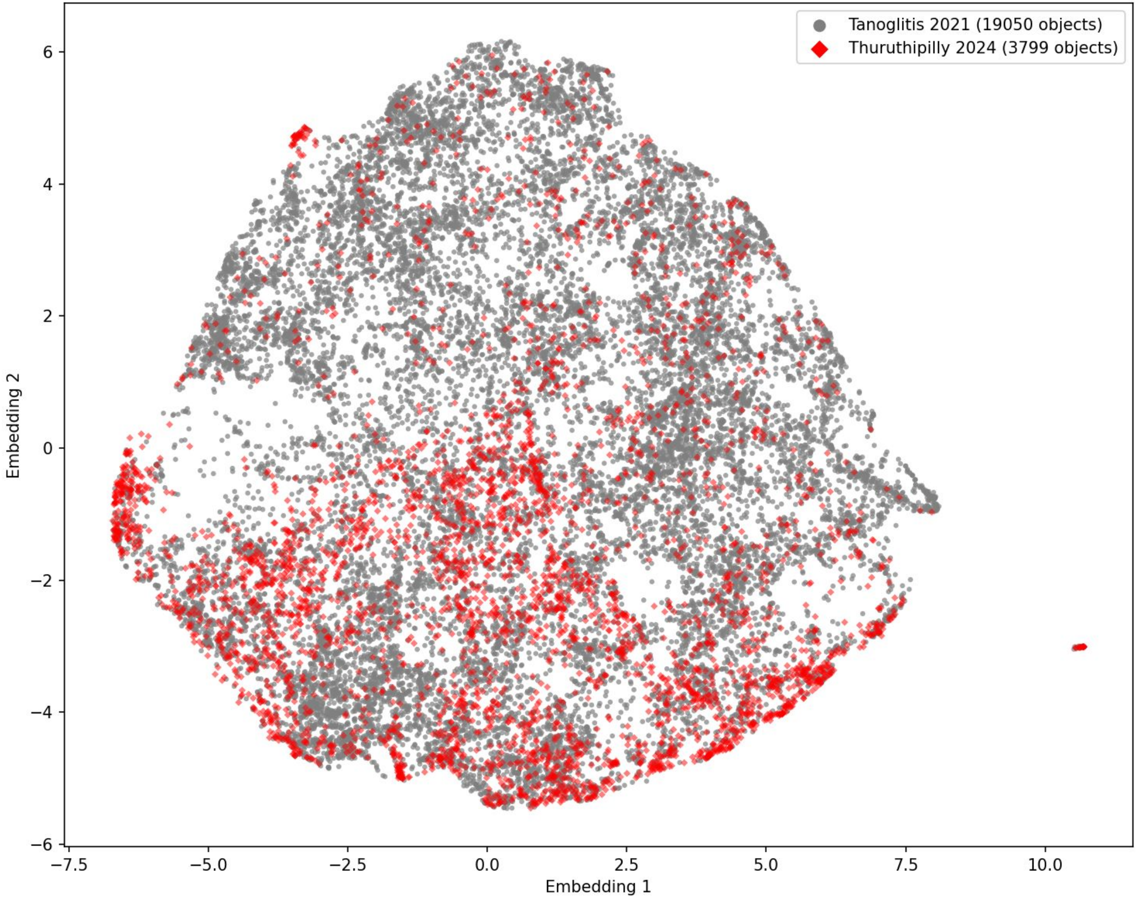
Re-analysis of all candidates found by Thuruthipilly's methods

Tanoglitis 2021



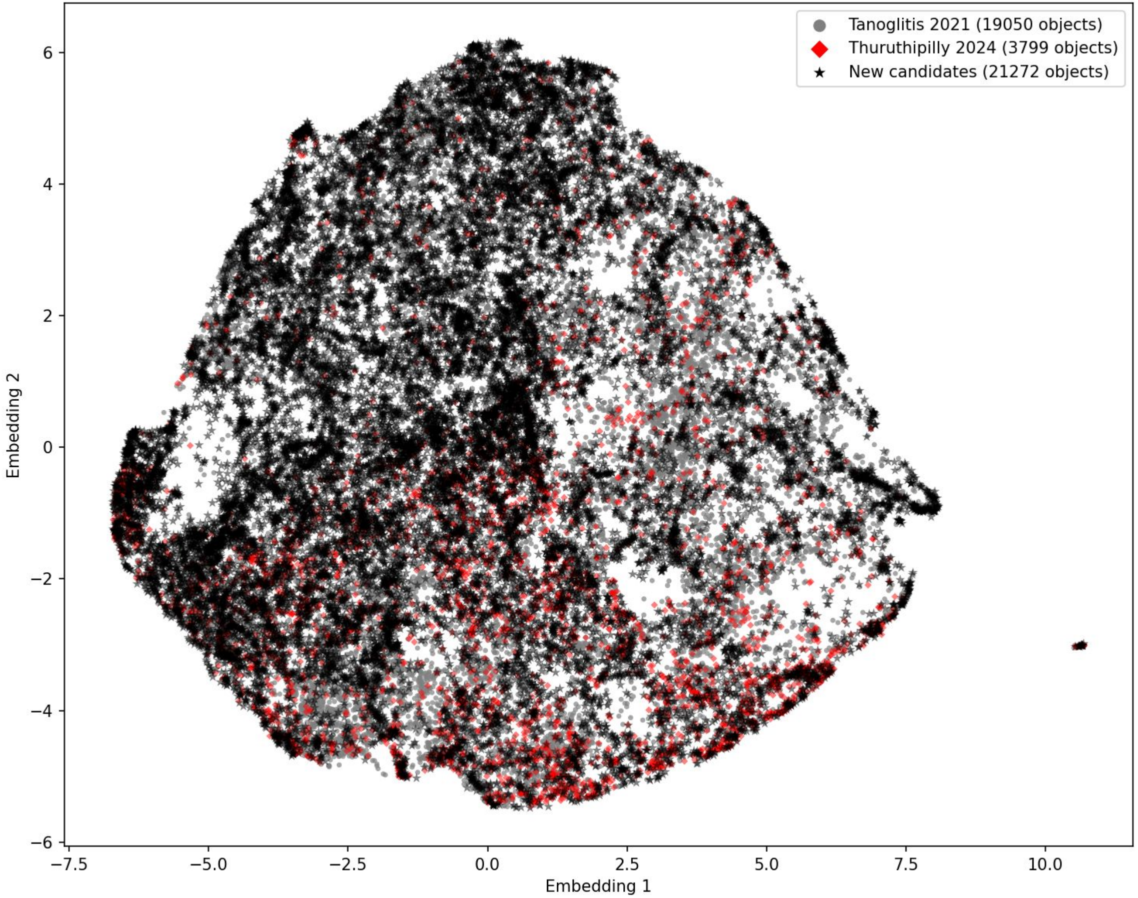
Re-analysis of all candidates found by Thuruthipilly's methods

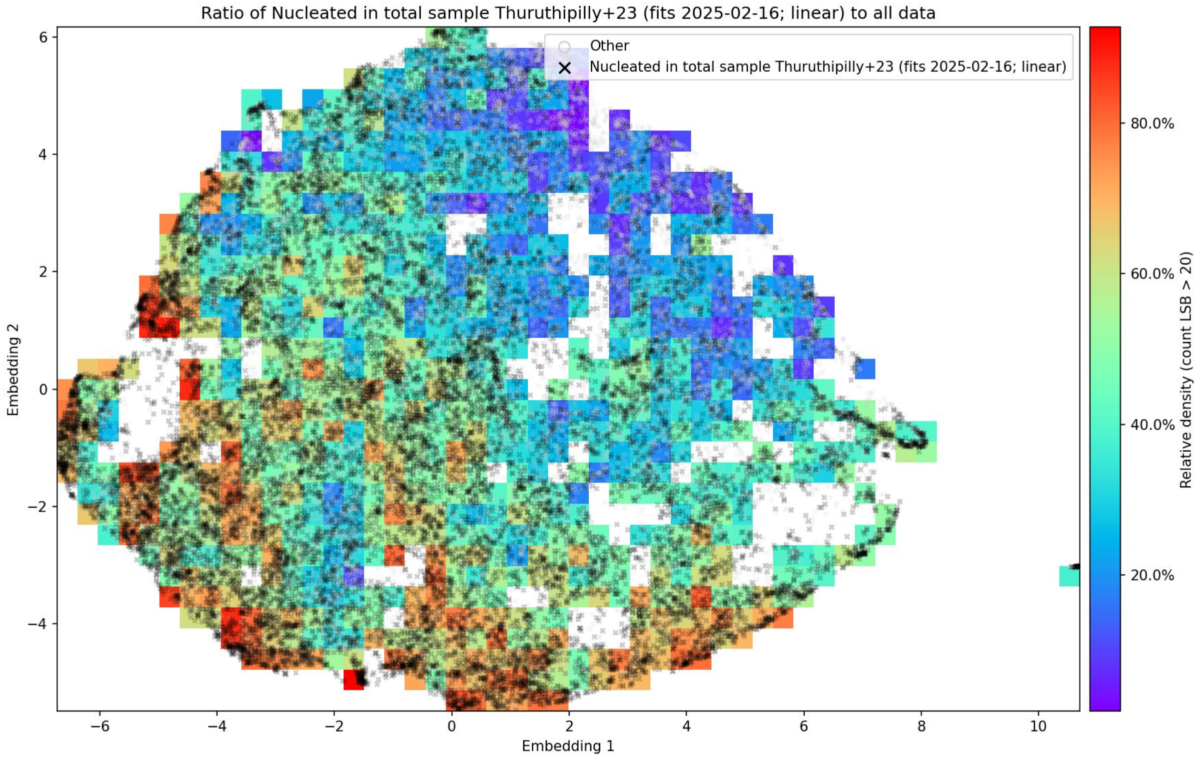
Thuruthpilly 2021

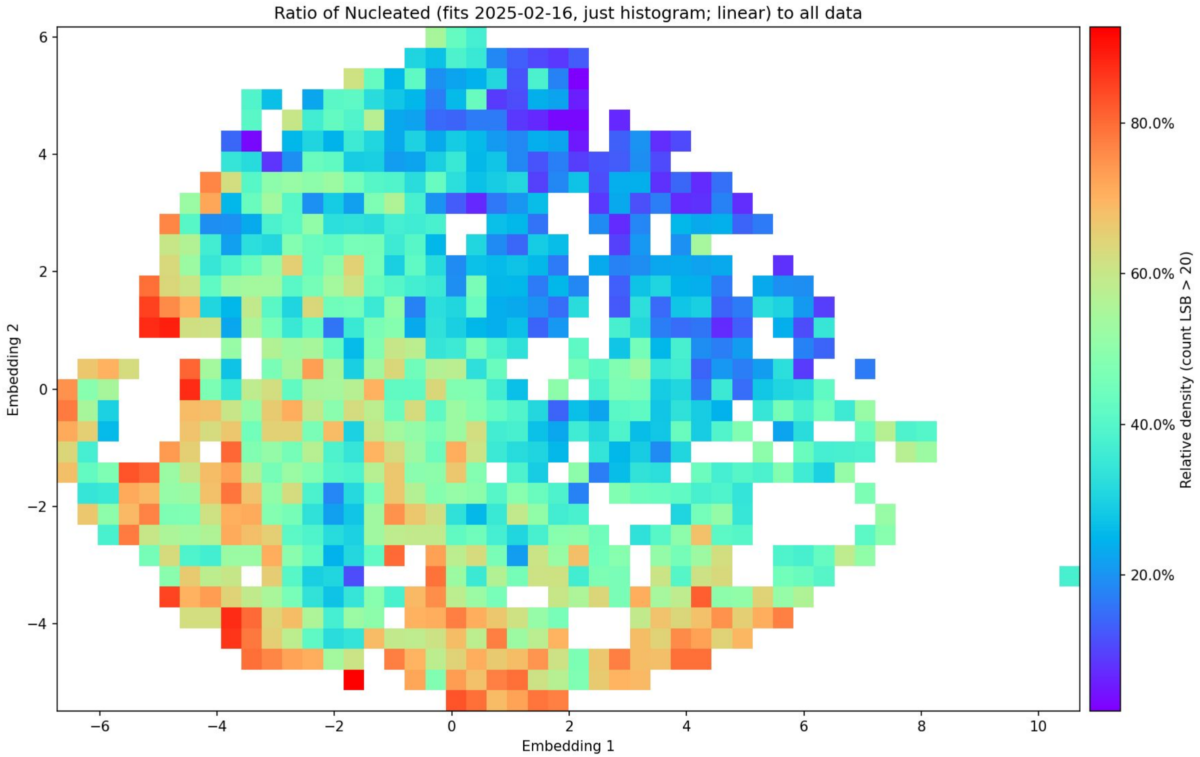


Re-analysis of all candidates found by Thuruthipilly's methods

New candidates



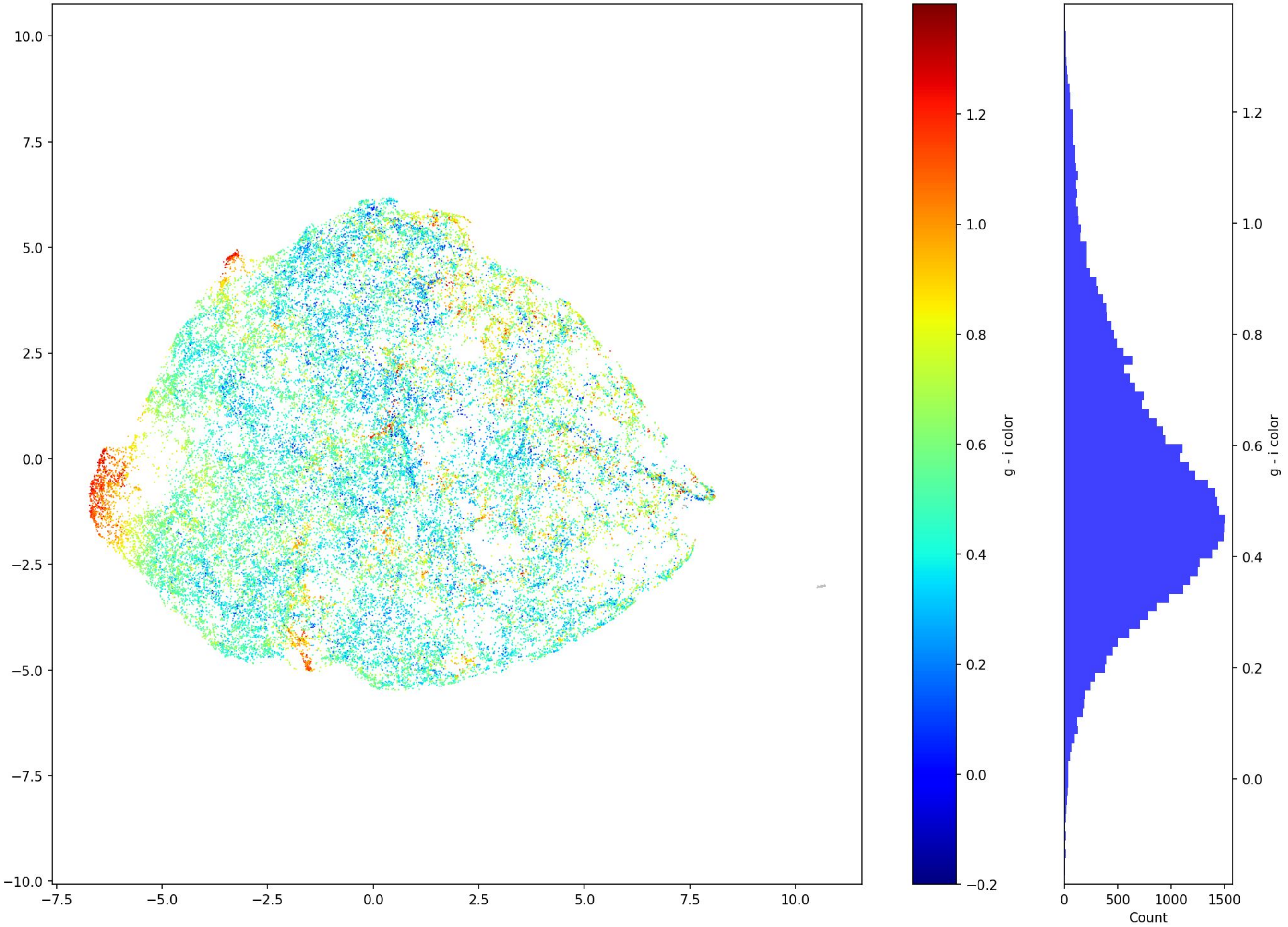

Ratio of Nucleated in total sample Thuruthipilly+23 (fits 2025-02-16; linear) to all data
Other
Nucleated in total sample Thuruthipilly+23 (fits 2025-02-16; linear)
Embedding 1
Embedding 2
Relative density (count LSB > 20)
80.0%
60.0%
40.0%
20.0%
−6
−4
−2
0
2
4
6
8
10



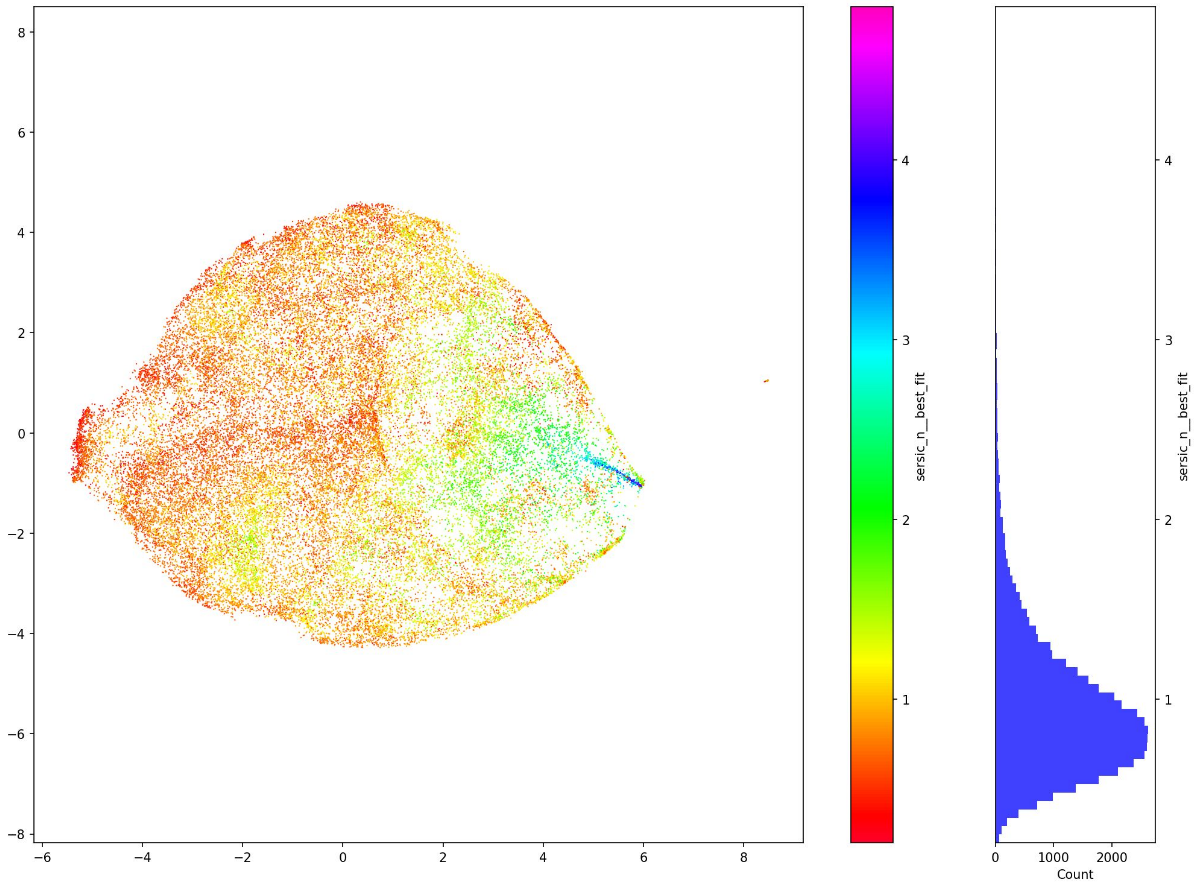
Ratio of Nucleated (fits 2025-02-16, just histogram; linear) to all data
Embedding 2
Embedding 1
−6
−4
−2
0
2
4
6
8
10
Relative density (count LSB > 20)
80.0%
60.0%
40.0%
20.0%

Embedding of the “new” catalog

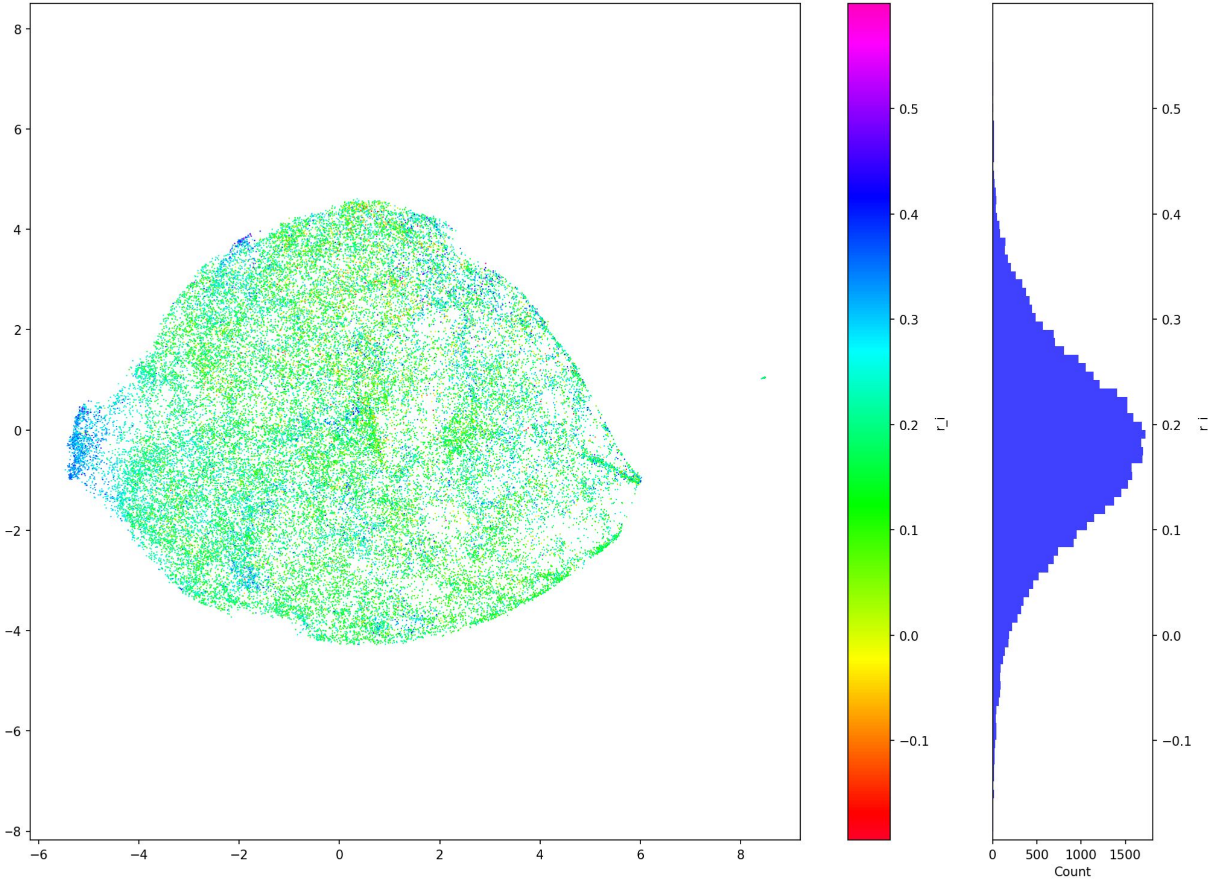
- Similarly region of “red” galaxies



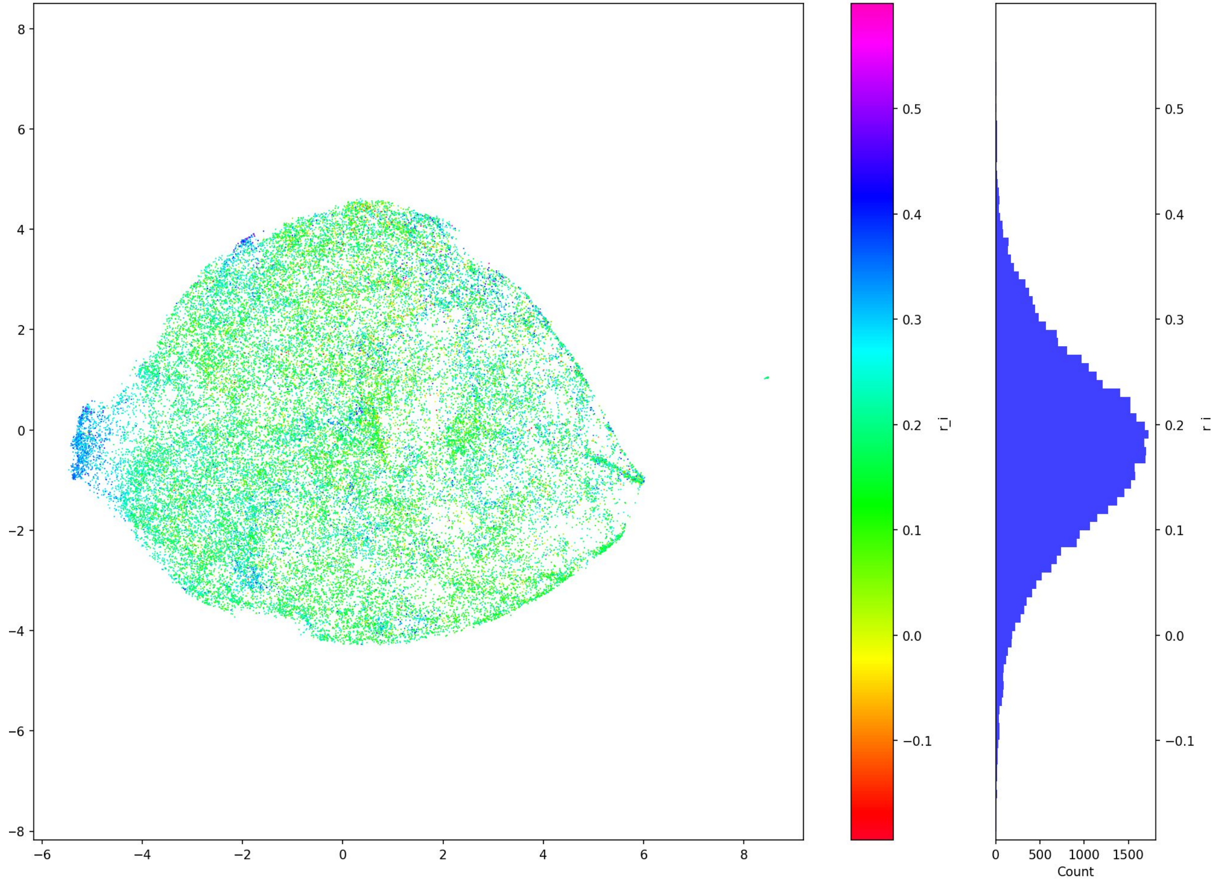
Embedding of the "new" catalog



Embedding of the “new” catalog



Embedding of the “new” catalog



Redshifts of UDGs in Thuruthipilly & Tanoglitis datasets