

Finding colliding galaxies in big data

Aidan Cotter

W. J. Pearson, S. Dey, B. Margalef-Bentabol,
A. Guzmán-Ortega, V. Rodríguez-Gómez

Graduate school seminar 5th March 2026



ESA/Hubble & NASA, J. Dalcanton, Dark Energy
Survey/DOE/FNAL/DECam/CTIO/NOIRLab/NSF/AURA

Outline



Introduction



Introduction

Classification techniques



Introduction

Classification techniques

Can morphologies beat AI?



Introduction

Classification techniques

Can morphologies beat AI?

Future developments



Introduction

Classification techniques

Can morphologies beat AI?

Future developments

Summary



Introduction



Classification techniques

Can morphologies beat AI?

Future developments

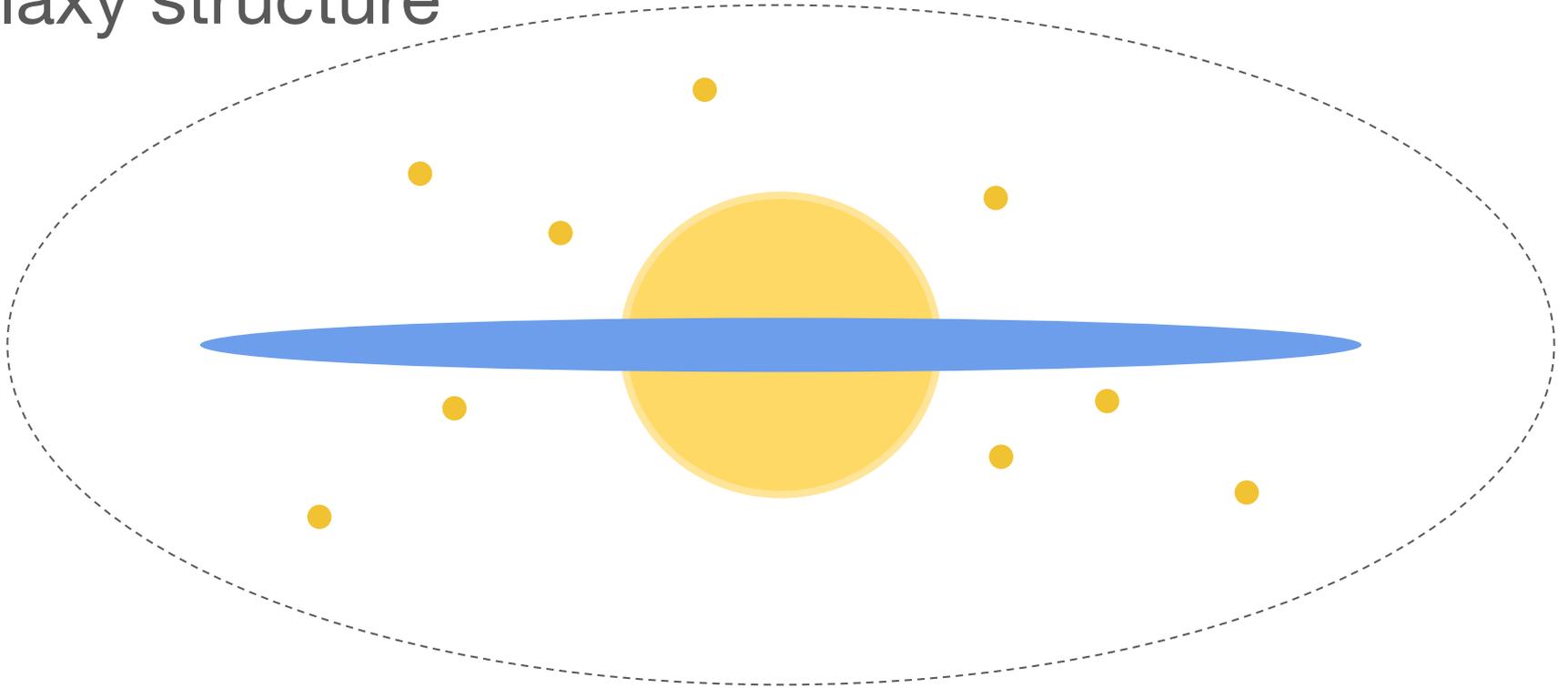
Summary



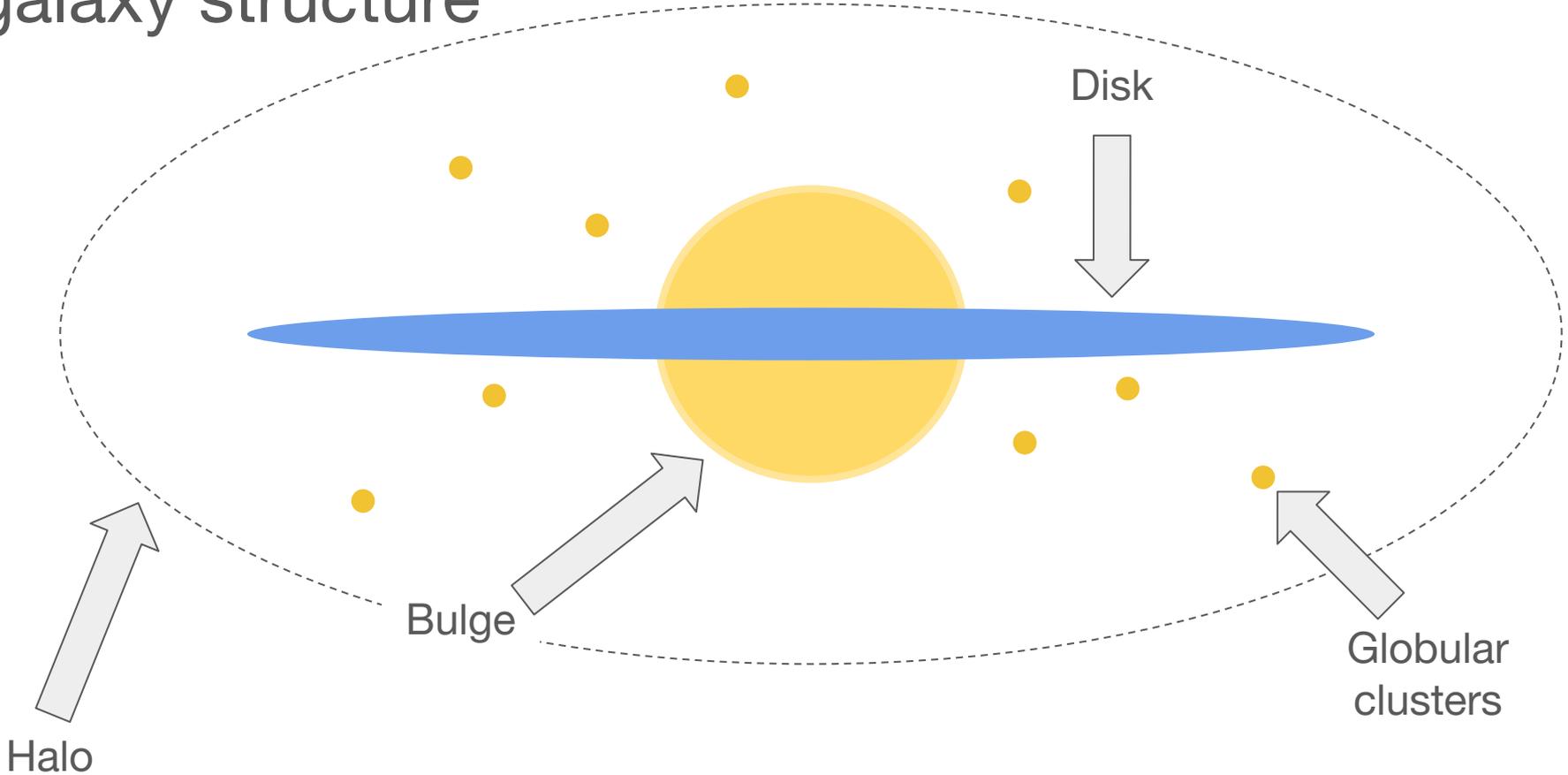
An overview of galaxy structure



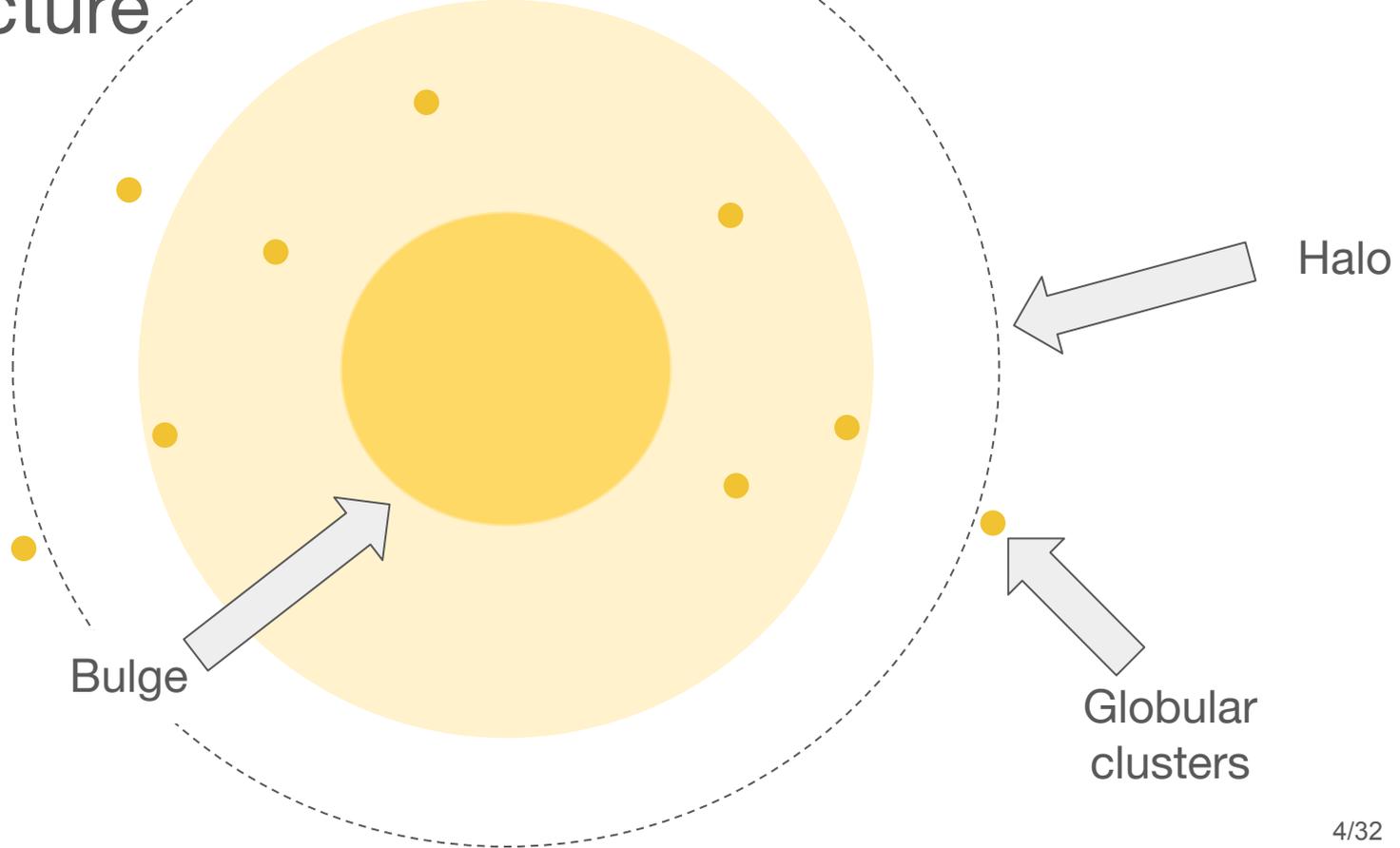
An overview of galaxy structure

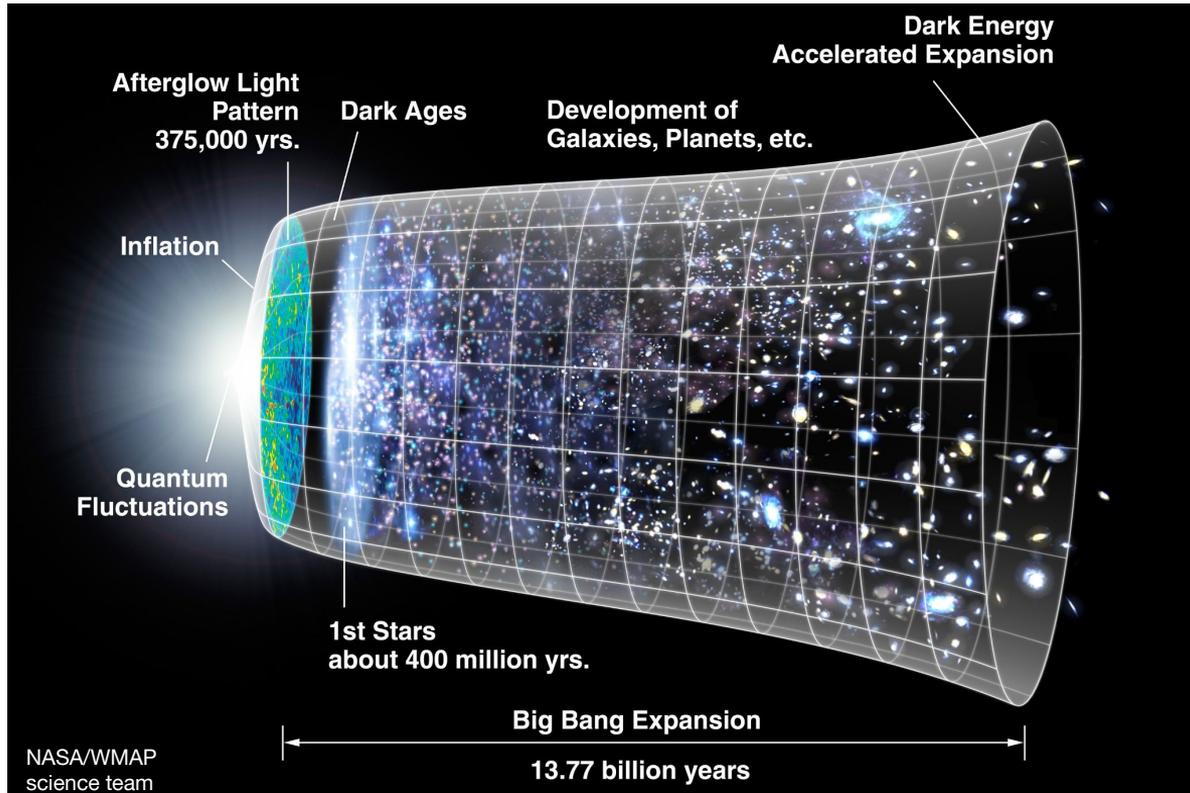


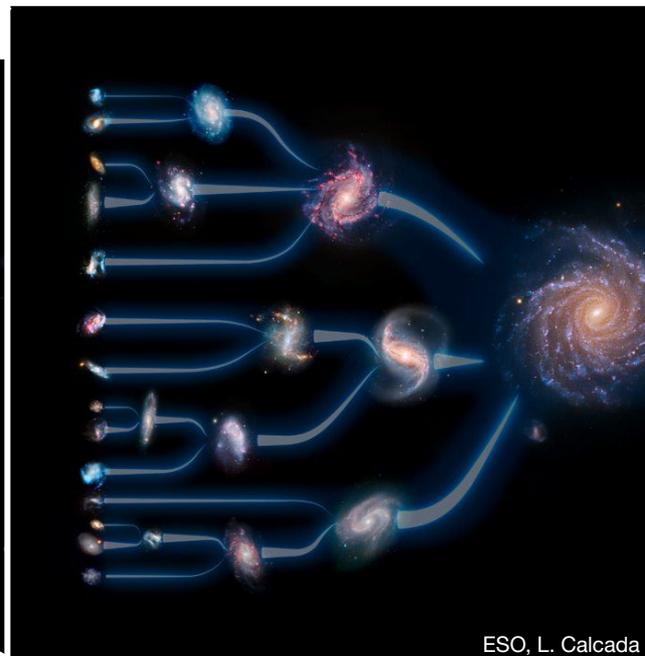
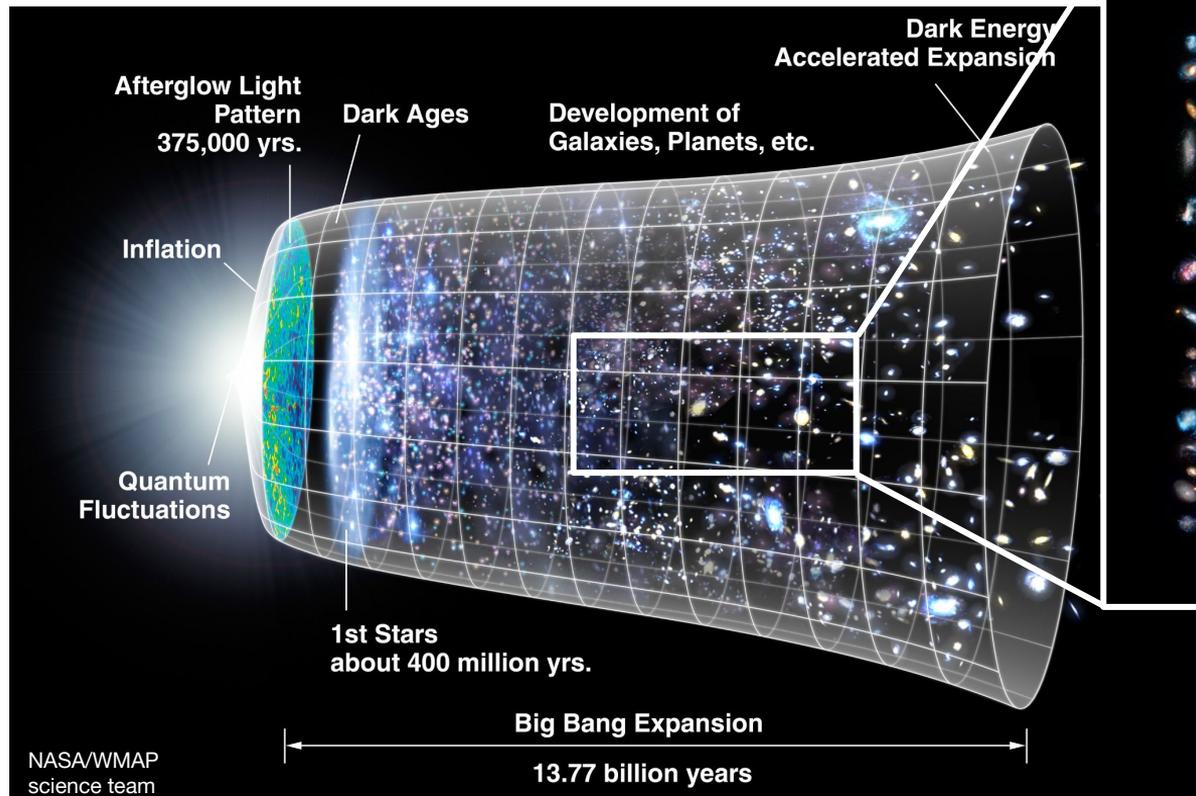
An overview of galaxy structure



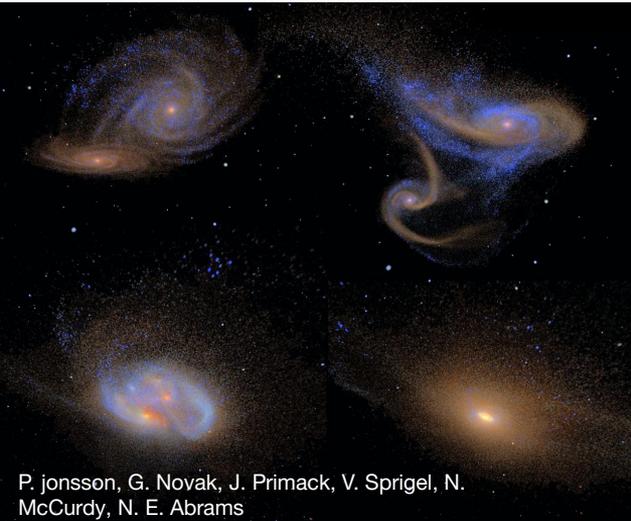
An overview of galaxy structure

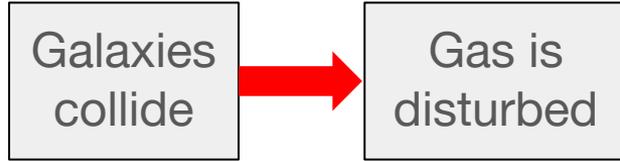




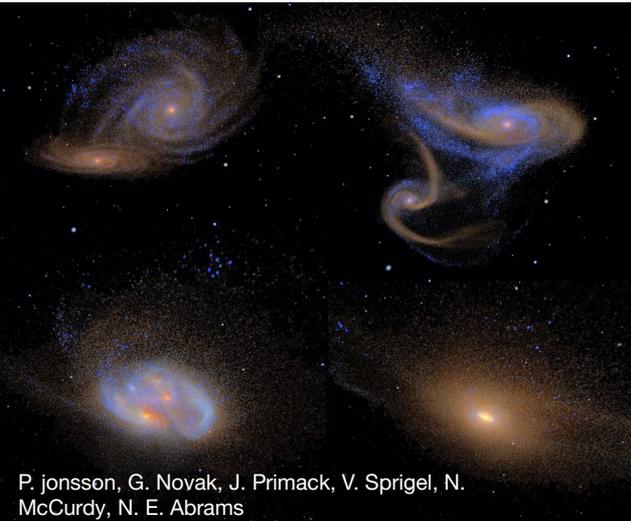


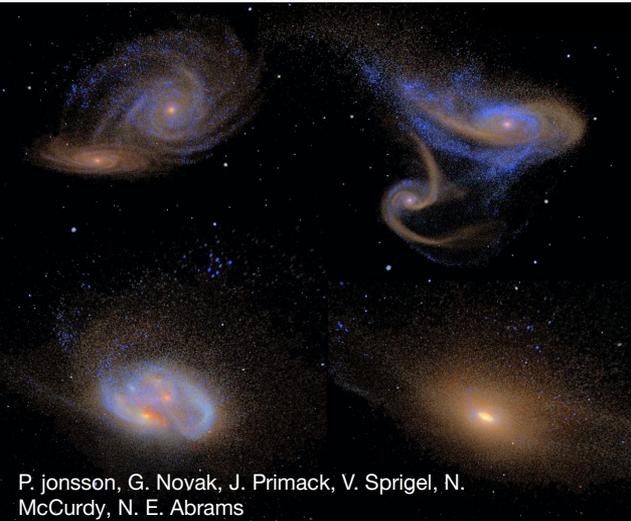
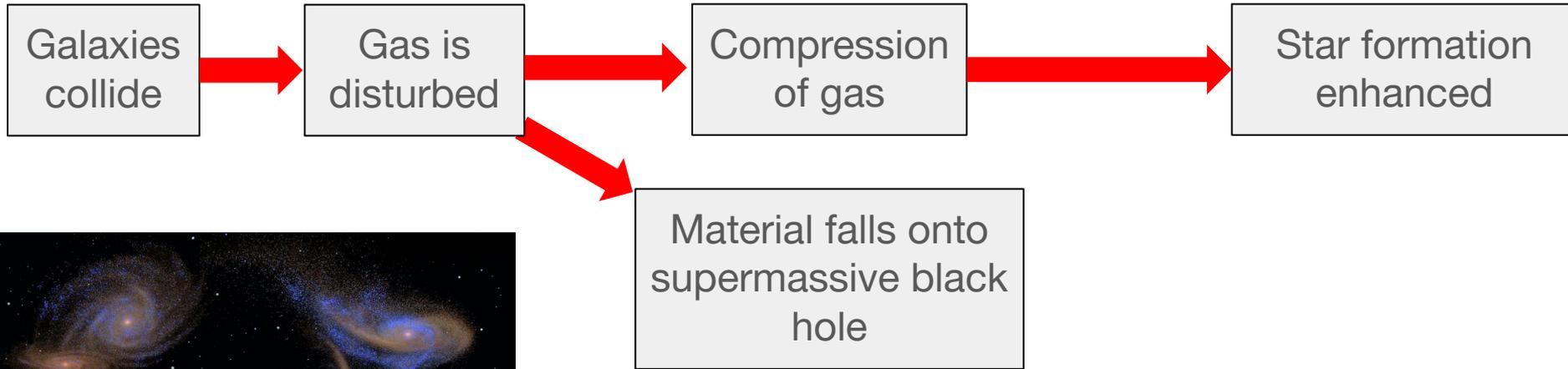
Galaxies
collide

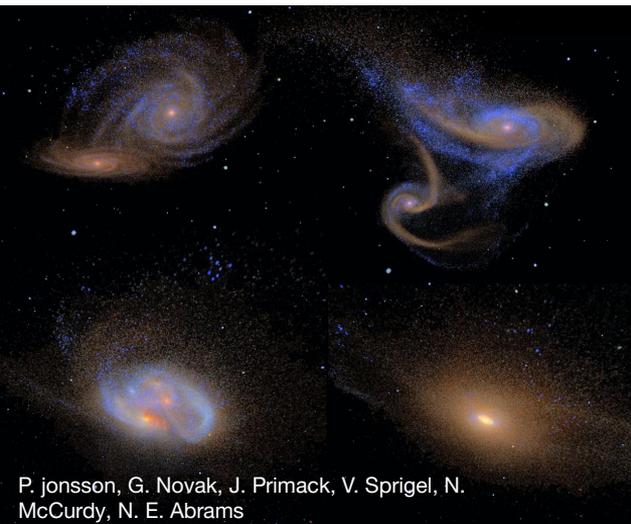
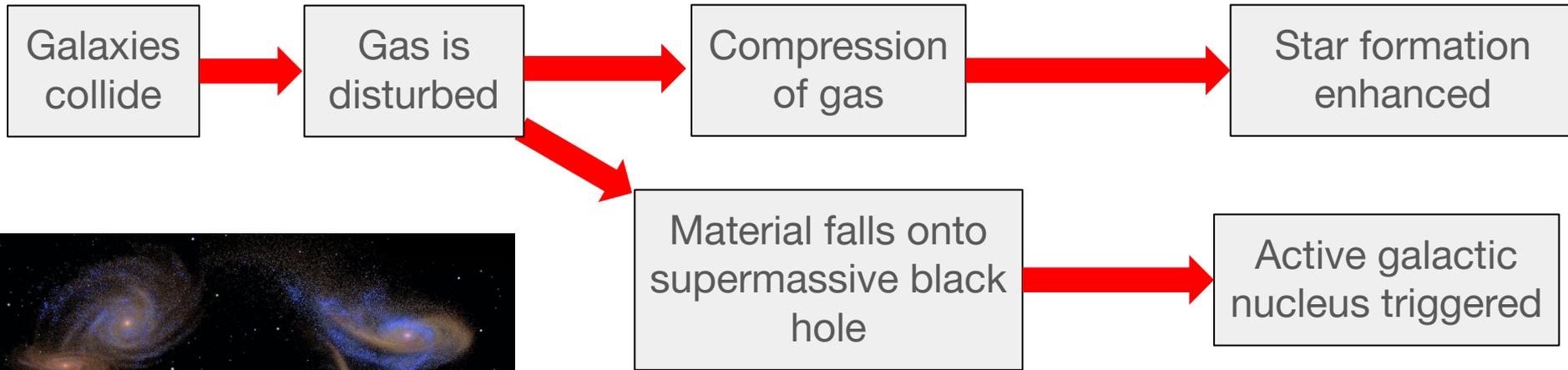


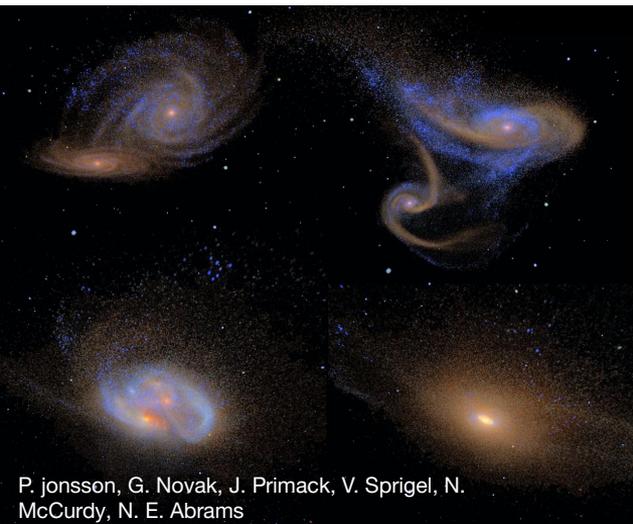
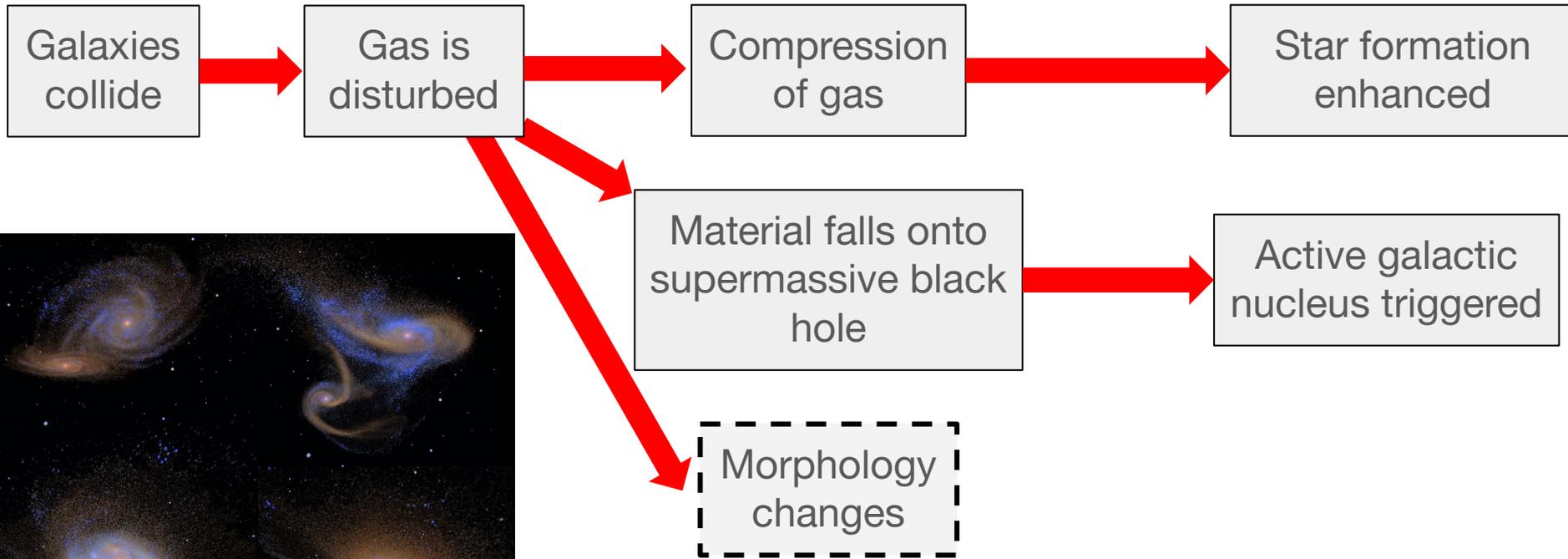












Challenges in classifying mergers

Irregular morphology and complicated dynamics

- Methods based on modelling are impractical

Challenges in classifying mergers

Irregular morphology and complicated dynamics

- Methods based on modelling are impractical

Projection problems

- Cannot always see pairs

Challenges in classifying mergers

Irregular morphology and complicated dynamics

- Methods based on modelling are impractical

Projection problems

- Cannot always see pairs

Difference between surveys

- Inconsistent classifications

Introduction

Classification techniques



Can morphologies beat AI?

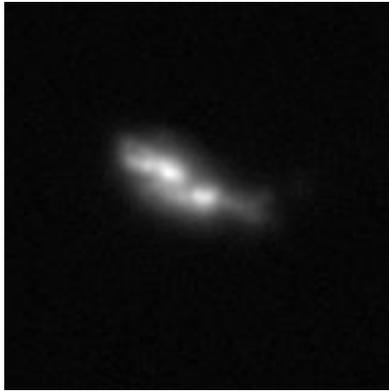
Future developments

Summary



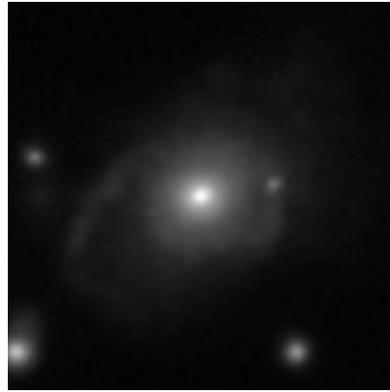
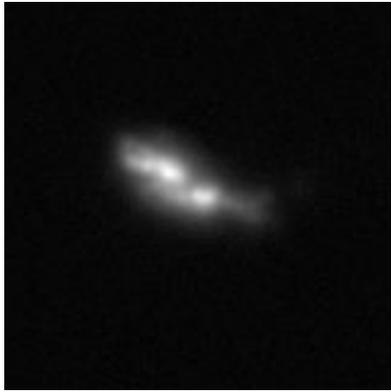
Visual classification

Classification techniques



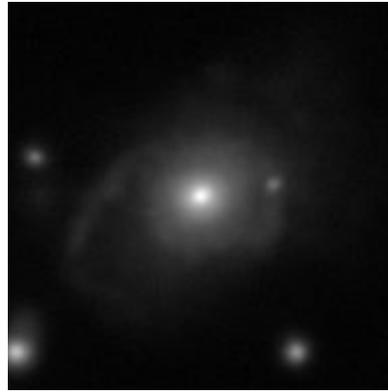
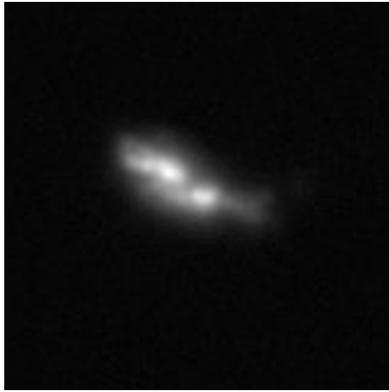
Visual classification

Classification techniques



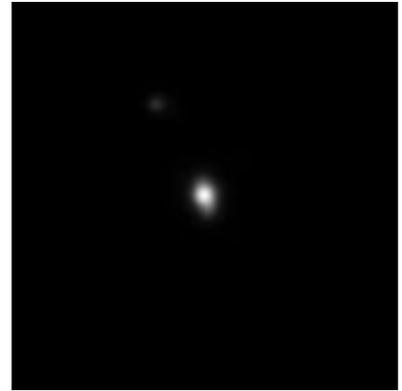
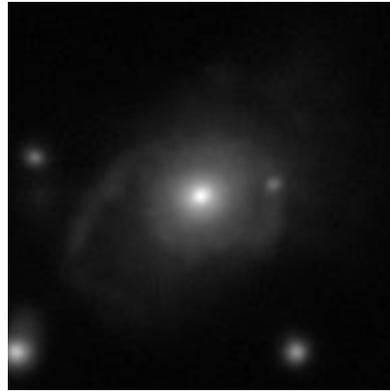
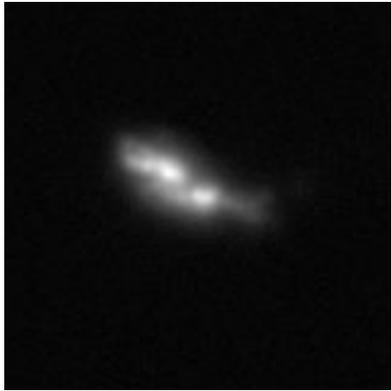
Visual classification

Classification techniques



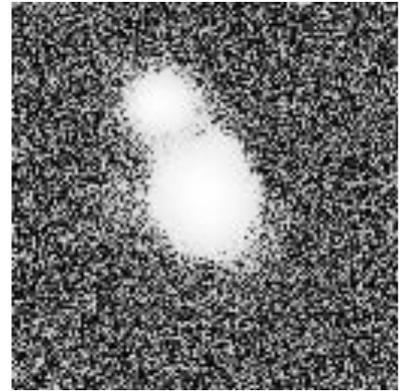
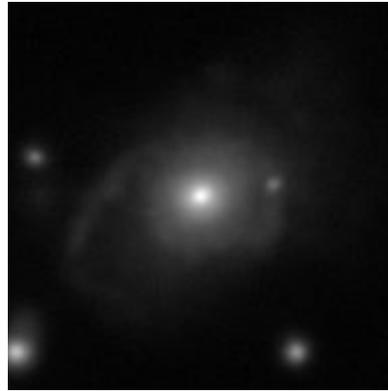
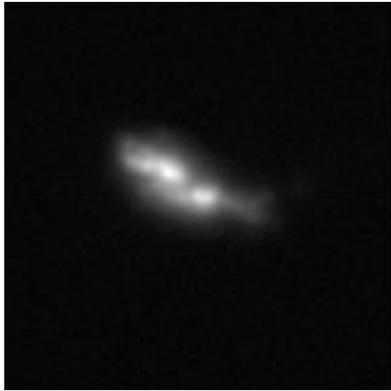
Visual classification

Classification techniques



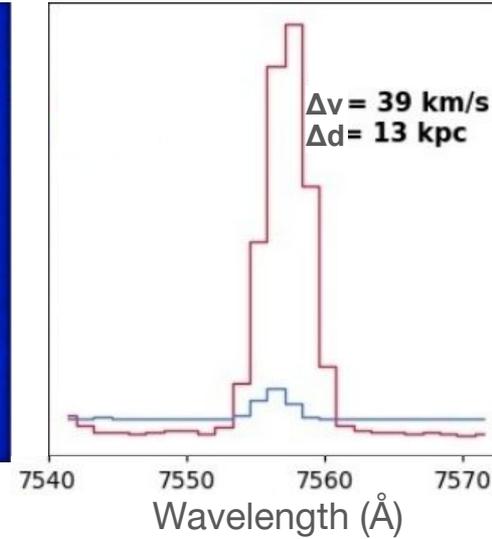
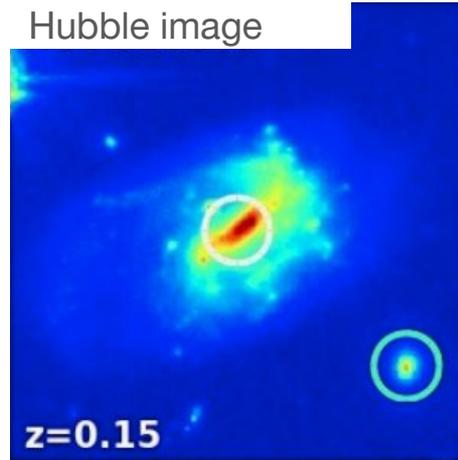
Visual classification

Classification techniques

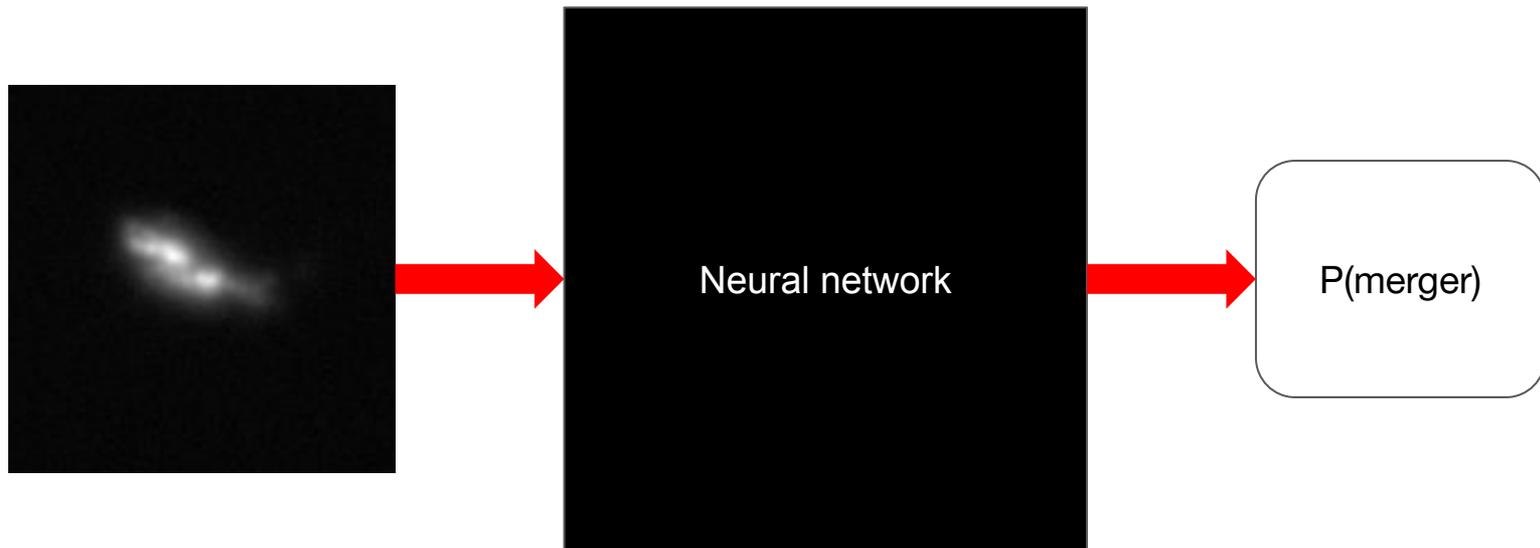


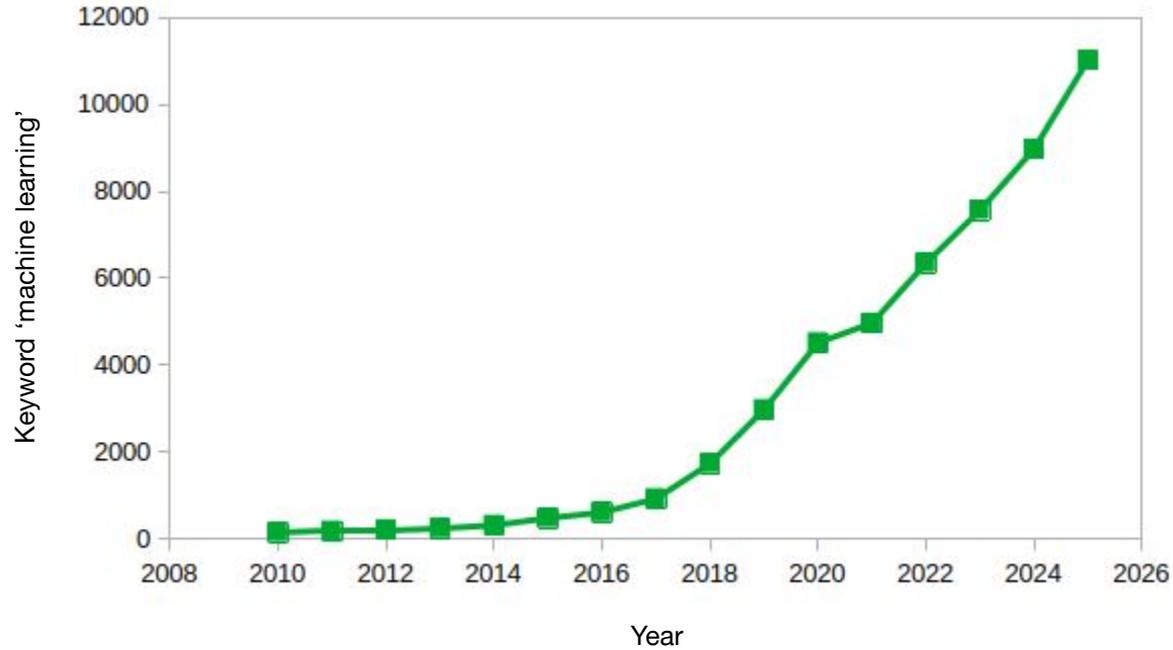
Close pairs method

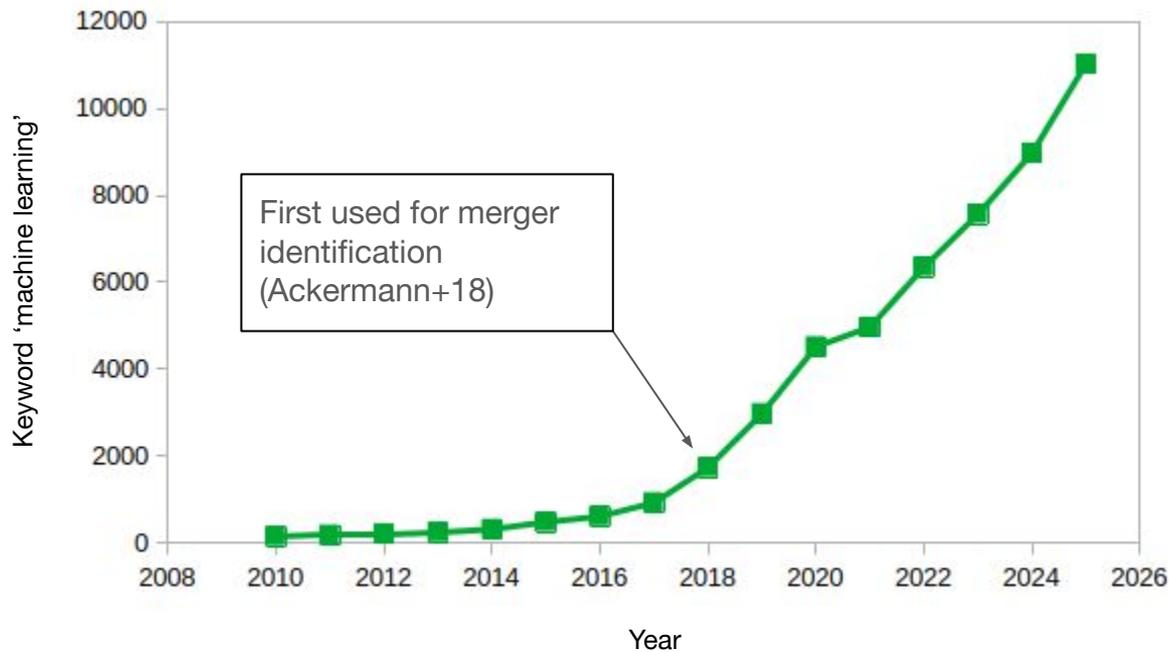
Classification techniques



Zoomed in spectrum

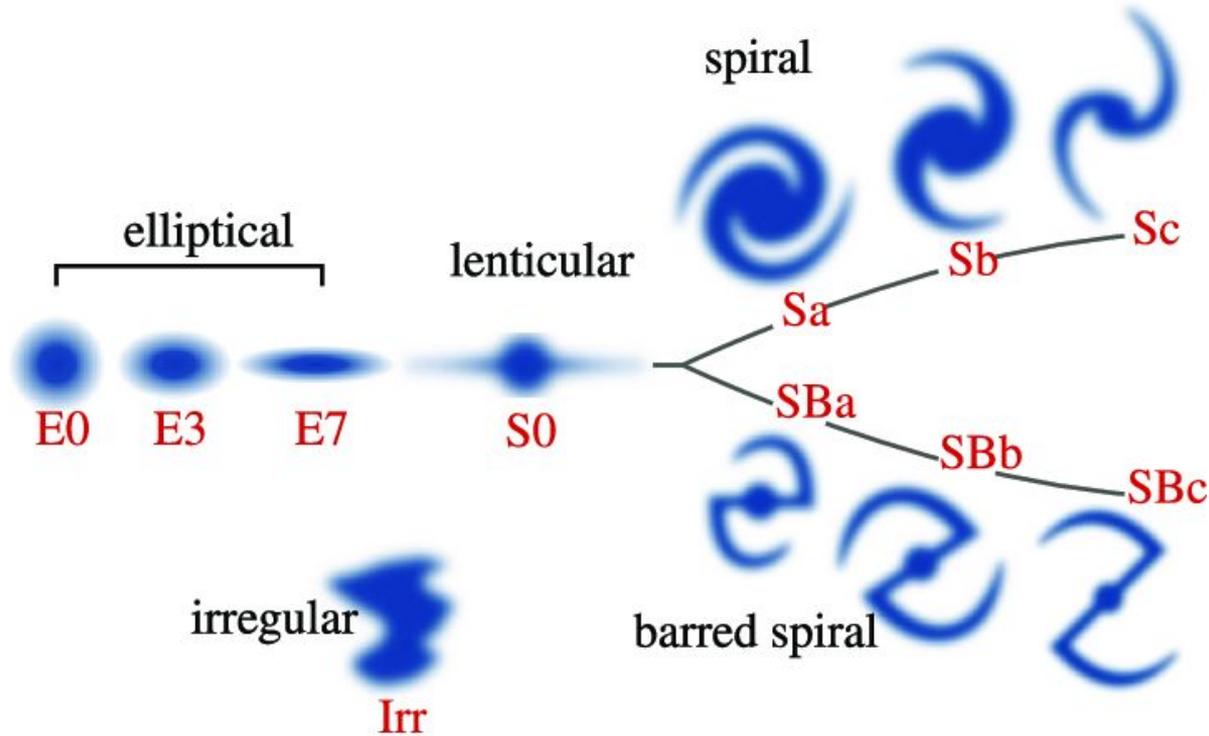






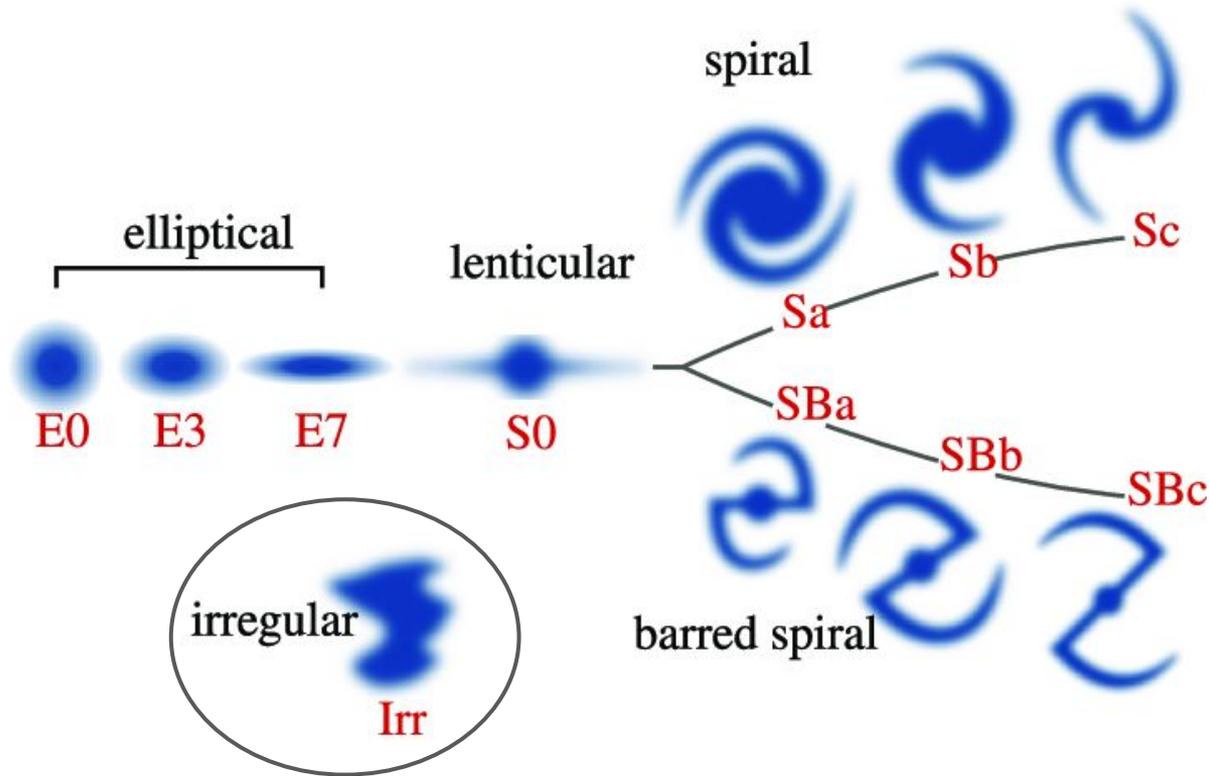
Morphological classifiers

Classification techniques



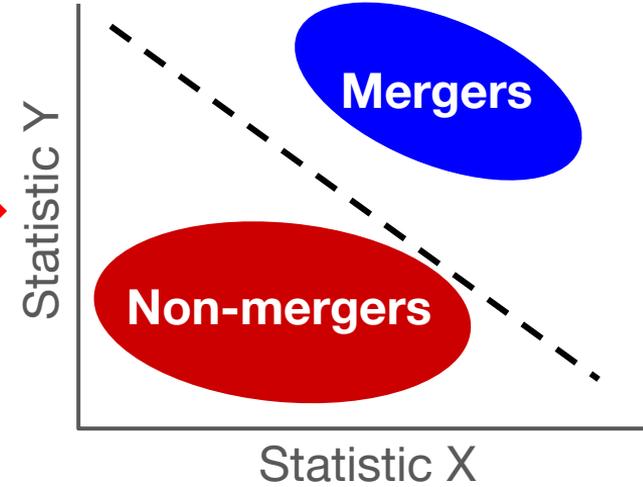
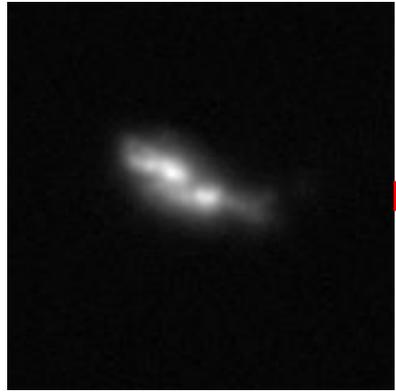
Morphological classifiers

Classification techniques



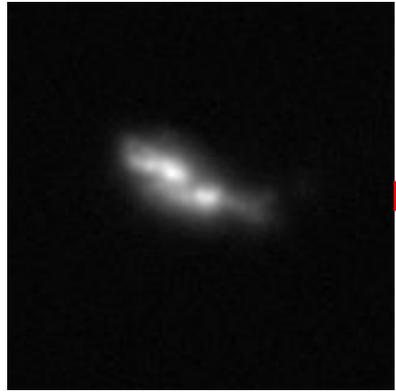
Morphological classifiers

Classification techniques

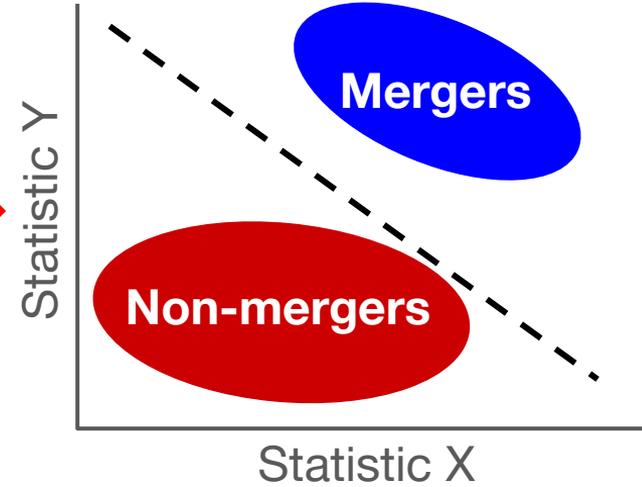


Morphological classifiers

Classification techniques



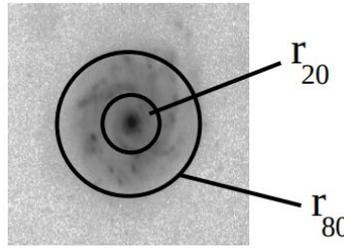
C, A, S, G, M_{20}



Morphological classifiers

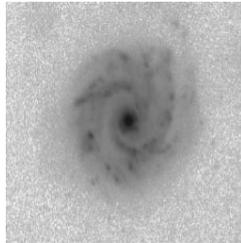
Classification techniques

Concentration

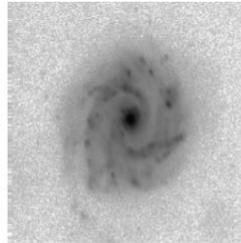


$$C = 5 \log\left(\frac{r_{80}}{r_{20}}\right)$$

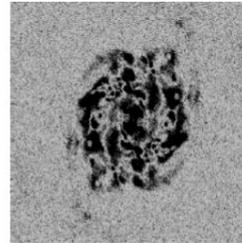
Asymmetry



I



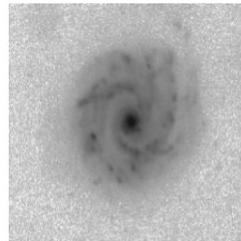
R



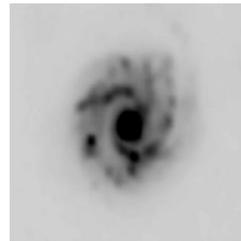
abs(I-R)

$$A = \frac{\text{abs}(I-R)}{I}$$

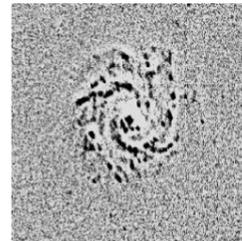
Smoothness



I



B



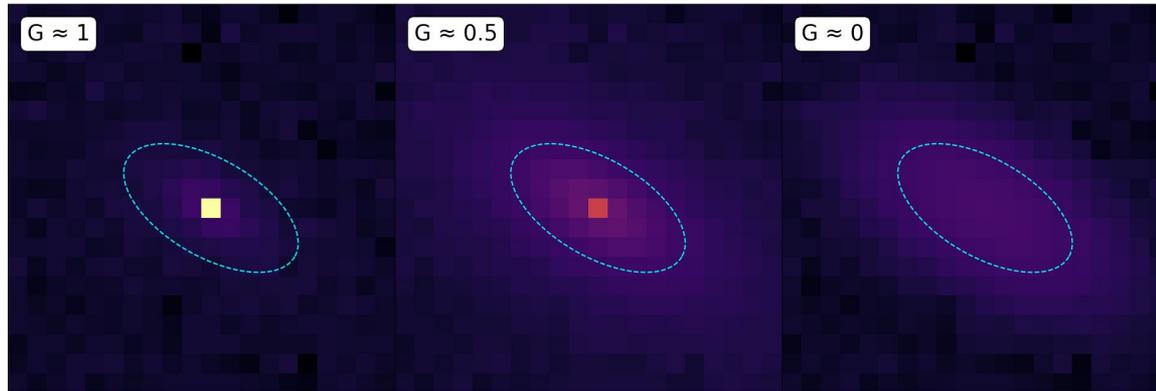
I-B

$$S = \frac{I-B}{I}$$

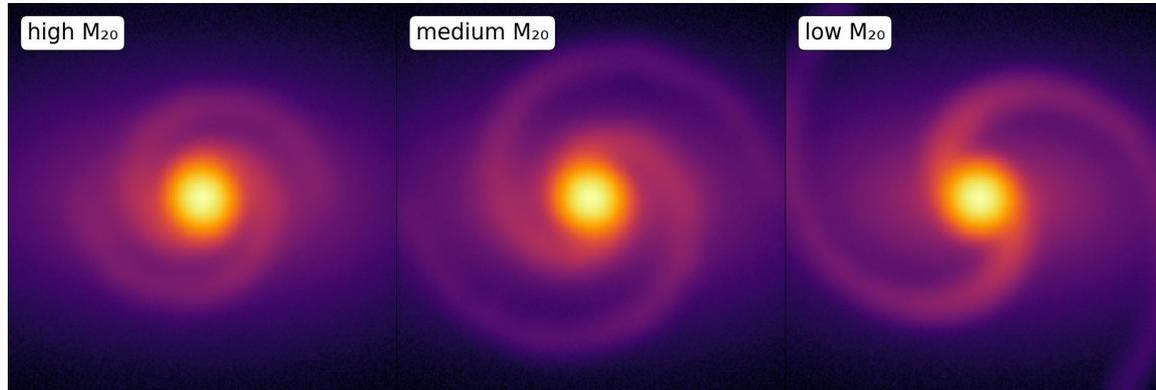
Morphological classifiers

Classification techniques

Gini index



Second order moment of light (brightest 20%)



Morphological classifiers

Classification techniques

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

Number of pixels

Pixel values

$G \approx 1$

$G \approx 0.5$

$G \approx 0$

high M_{20}

medium M_{20}

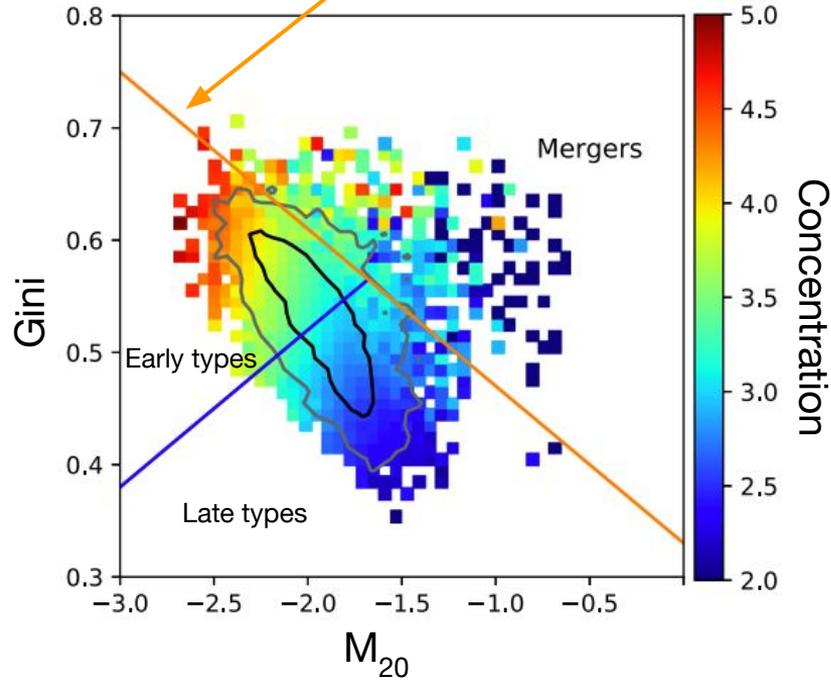
low M_{20}

$$M_{20} = \log_{10} \left(\frac{\sum_i \mu_i}{\mu_{total}} \right), \text{ while } \sum_i X_i < 0.2 X_{total}$$
$$\mu_i = X_i \left((x_i - x_c)^2 + (y_i - y_c)^2 \right)$$

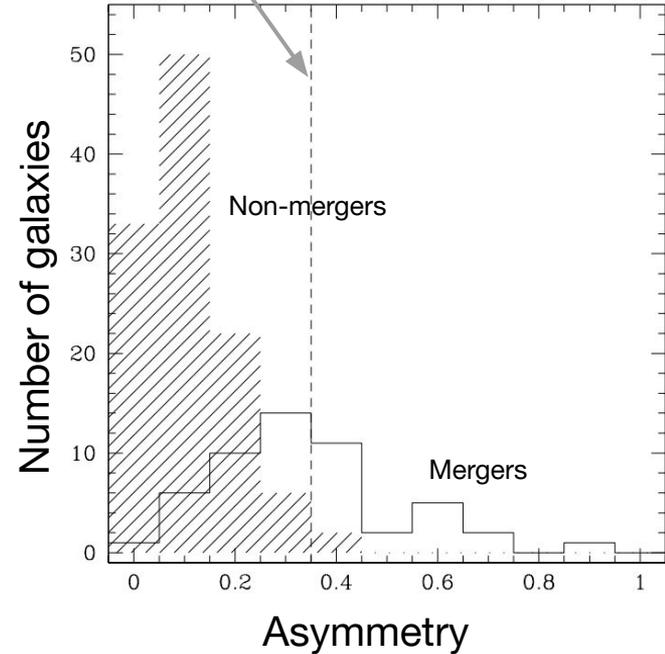
Morphological classifiers

Classification techniques

$$G > -0.14M_{20} + 0.33$$



$$A > 0.35$$



	Scalable	Accurate	Resource friendly
Visual inspection			
Close pairs method			
Machine learning			
Morphological classifiers			

	Scalable	Accurate	Resource friendly
Visual inspection			
Close pairs method			
Machine learning			
Morphological classifiers			

Introduction

Classification techniques

Can morphologies beat AI?



Future developments

Summary



A&A, 687, A24 (2024)
<https://doi.org/10.1051/0004-6361/202348239>
© The Authors 2024

**Astronomy
&
Astrophysics**

Galaxy merger challenge: A comparison study between machine learning-based detection methods

B. Margalef-Bentabol¹ , L. Wang^{1,2}, A. La Marca^{1,2} , C. Blanco-Prieto³, D. Chudy⁴, H. Domínguez-Sánchez⁵,
A. D. Goulding⁶, A. Guzmán-Ortega⁷ , M. Huertas-Company⁸, G. Martín^{9,10} , W. J. Pearson¹¹ ,
V. Rodríguez-Gomez⁷, M. Walmsley¹², R. W. Bickley¹³, C. Bottrell¹⁴ , C. Conselice¹², and D. O’Ryan¹⁵

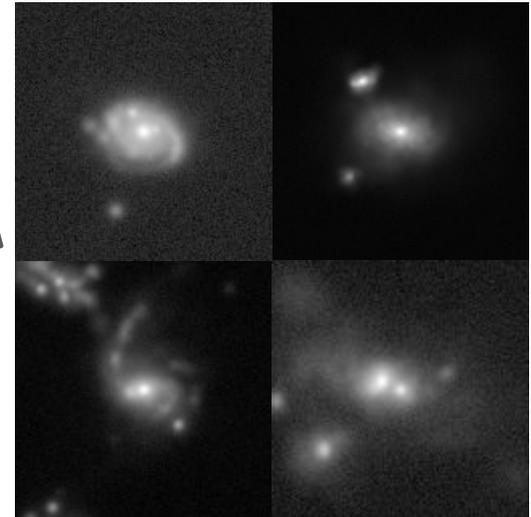
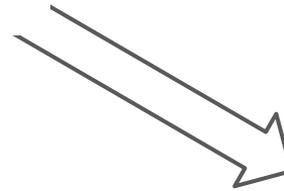
A&A, 687, A24 (2024)
<https://doi.org/10.1051/0004-6361/202348239>
© The Authors 2024

**Astronomy
&
Astrophysics**

Galaxy merger challenge: A comparison study between machine learning-based detection methods

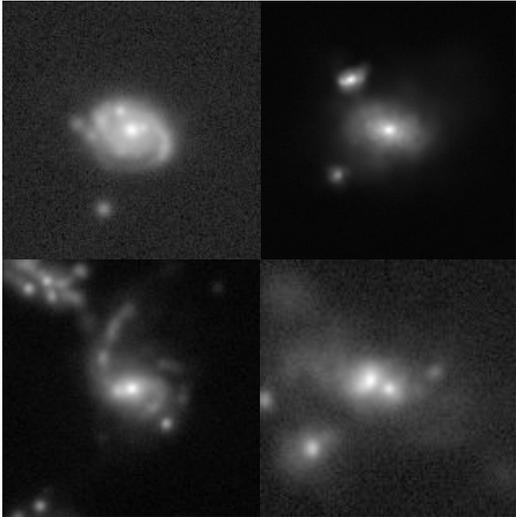
B. Margalef-Bentabol¹, L. Wang^{1,2}, A. La Marca^{1,2}, C. Blanco-Prieto³, D. Chudy⁴, H. Domínguez-Sánchez⁵,
A. D. Goulding⁶, A. Guzmán-Ortega⁷, M. Huertas-Company⁸, G. Martín^{9,10}, W. J. Pearson¹¹,
V. Rodríguez-Gomez⁷, M. Walmsley¹², R. W. Bickley¹³, C. Bottrell¹⁴, C. Conselice¹², and D. O’Ryan¹⁵

1. Mass ratio $> 1:4$
2. Merging in the next 800 Myrs or in the past 300 Myrs
3. Stellar mass $> 10^9 M_{\odot}$



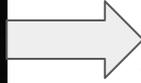
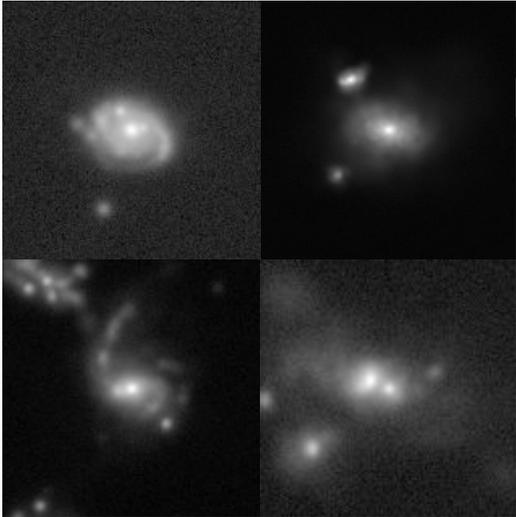
Updated merger morphologies

Can morphologies beat AI?



Updated merger morphologies

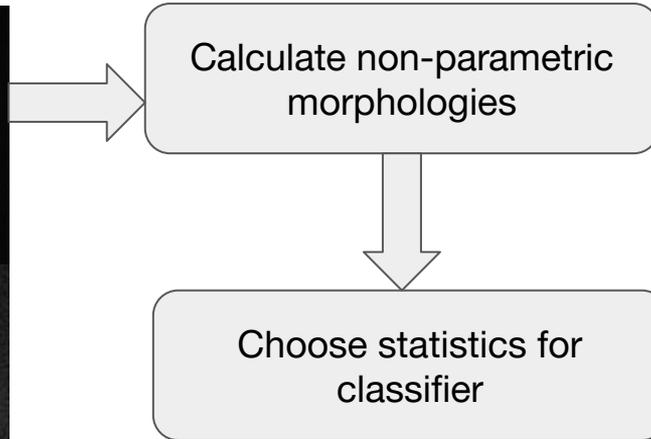
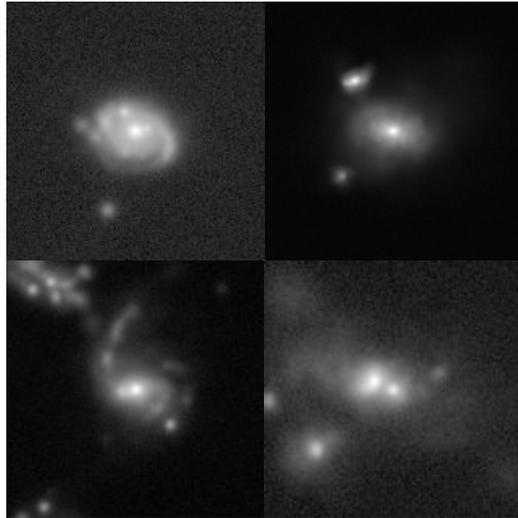
Can morphologies beat AI?



Calculate non-parametric morphologies

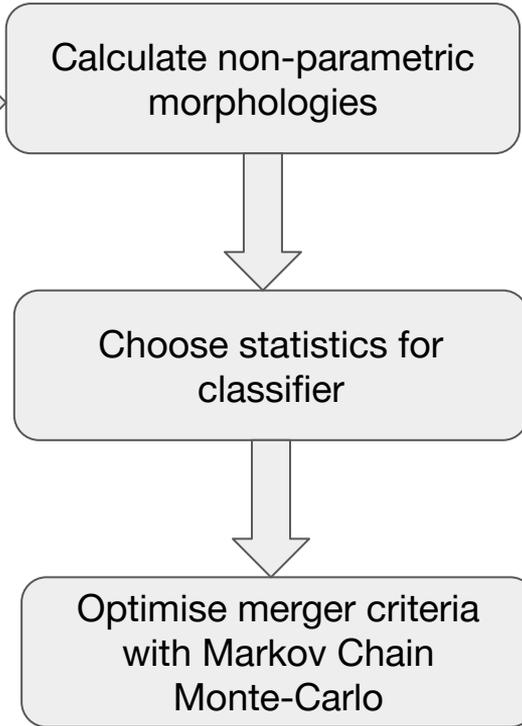
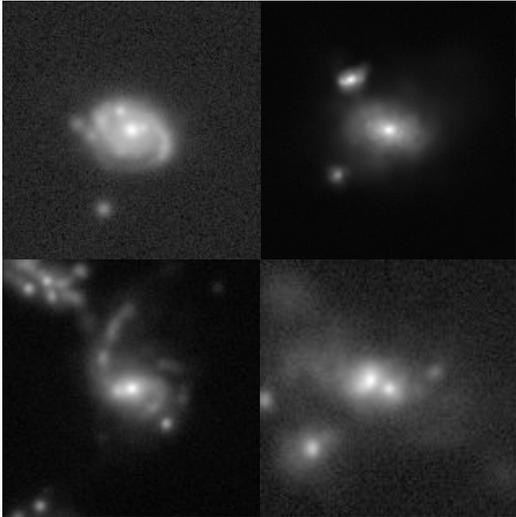
Updated merger morphologies

Can morphologies beat AI?



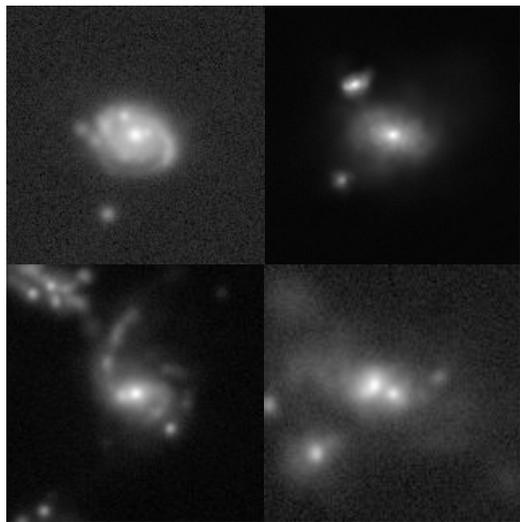
Updated merger morphologies

Can morphologies beat AI?



Updated merger morphologies

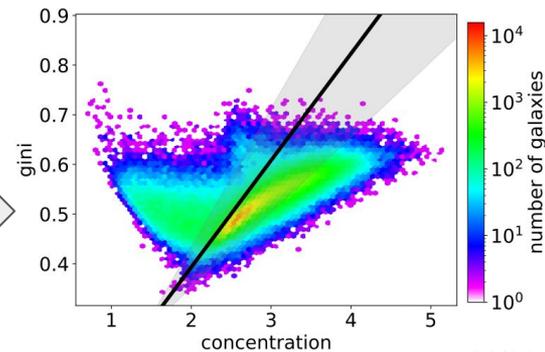
Can morphologies beat AI?



Calculate non-parametric morphologies

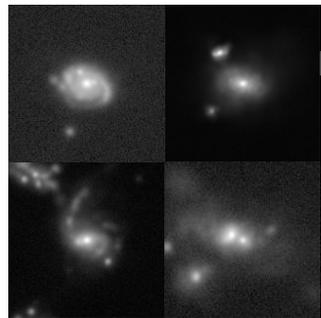
Choose statistics for classifier

Optimise merger criteria with Markov Chain Monte-Carlo



Updated merger morphologies

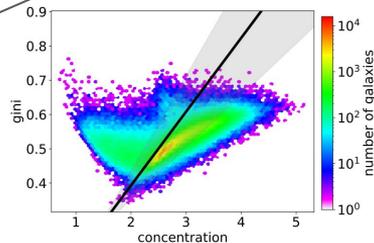
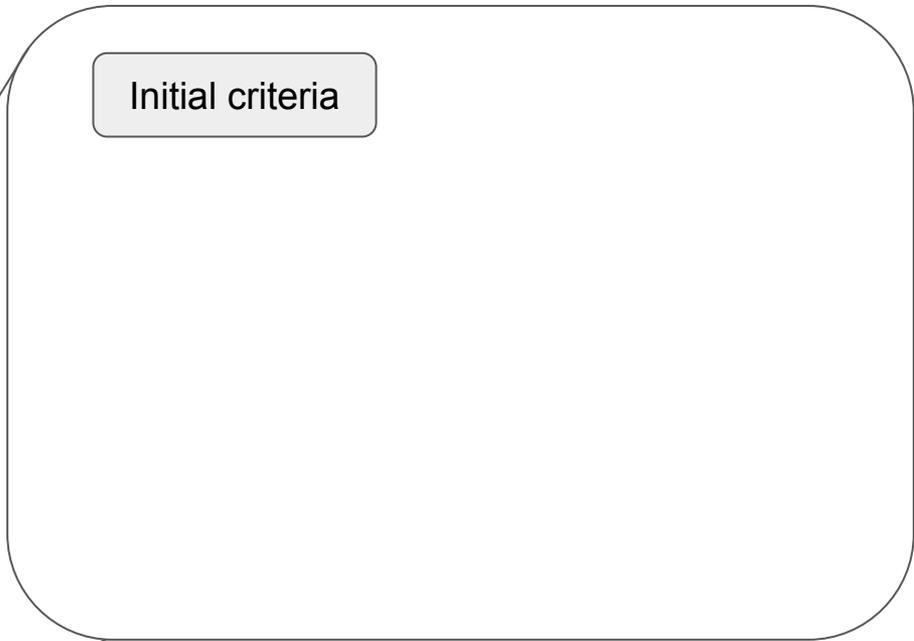
Can morphologies beat AI?



Calculate non-parametric morphologies

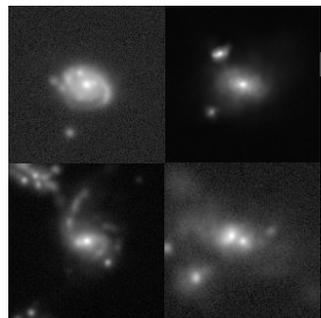
Choose statistics for classifier

Optimise merger criteria with Markov Chain Monte-Carlo



Updated merger morphologies

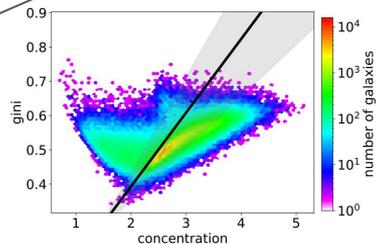
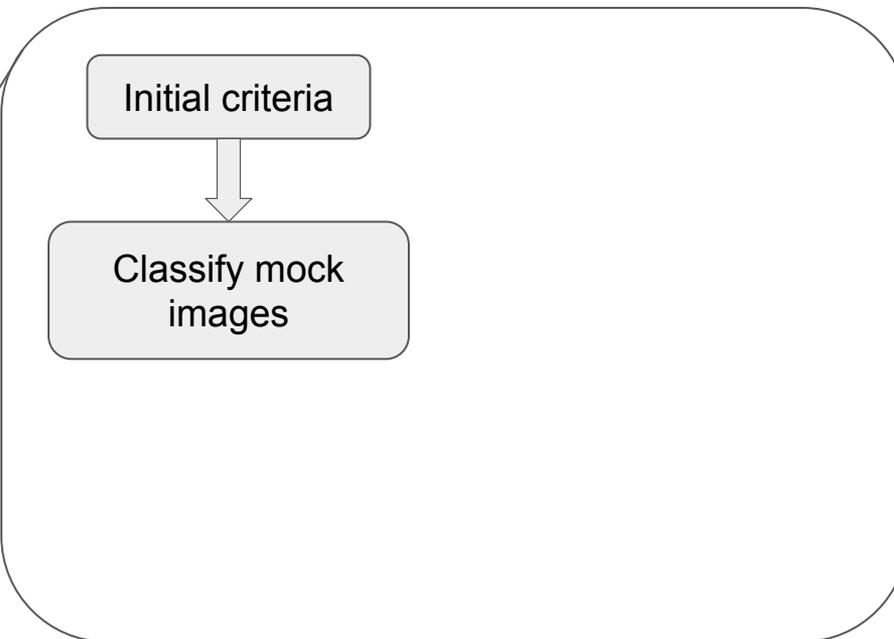
Can morphologies beat AI?



Calculate non-parametric morphologies

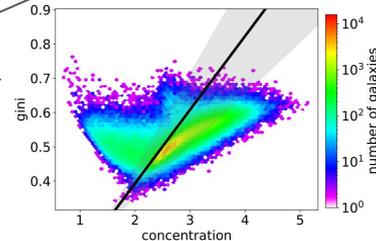
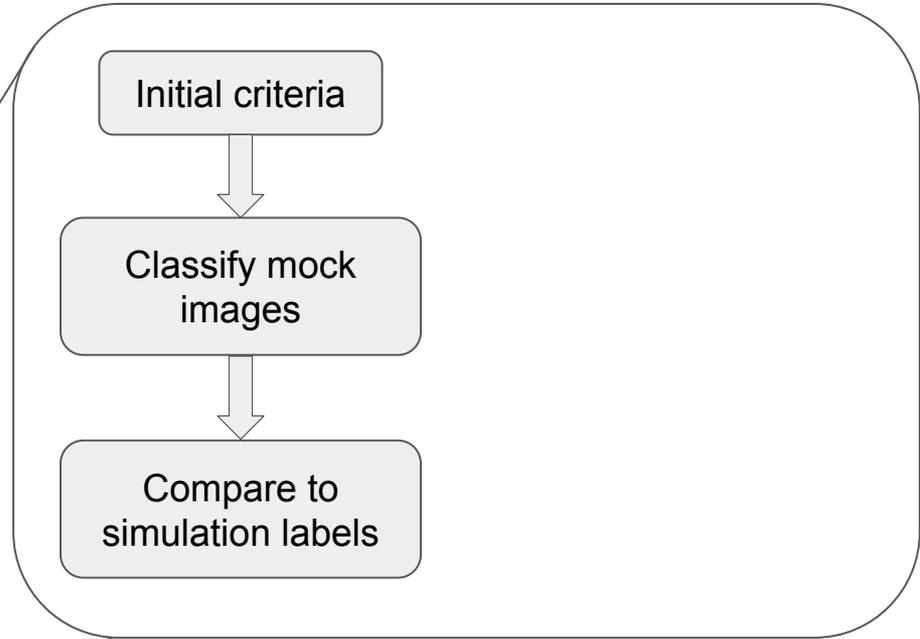
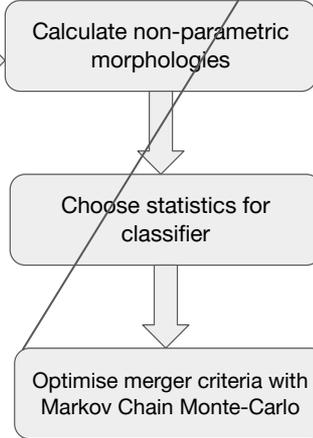
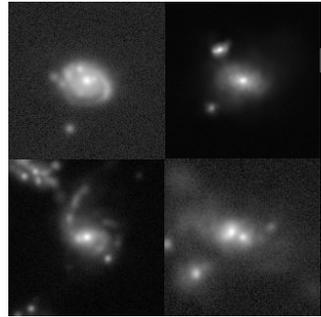
Choose statistics for classifier

Optimise merger criteria with Markov Chain Monte-Carlo



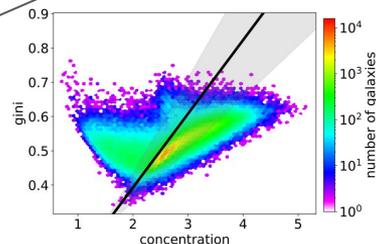
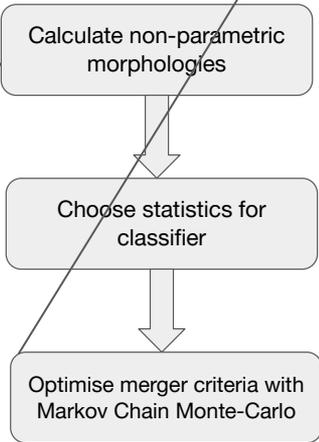
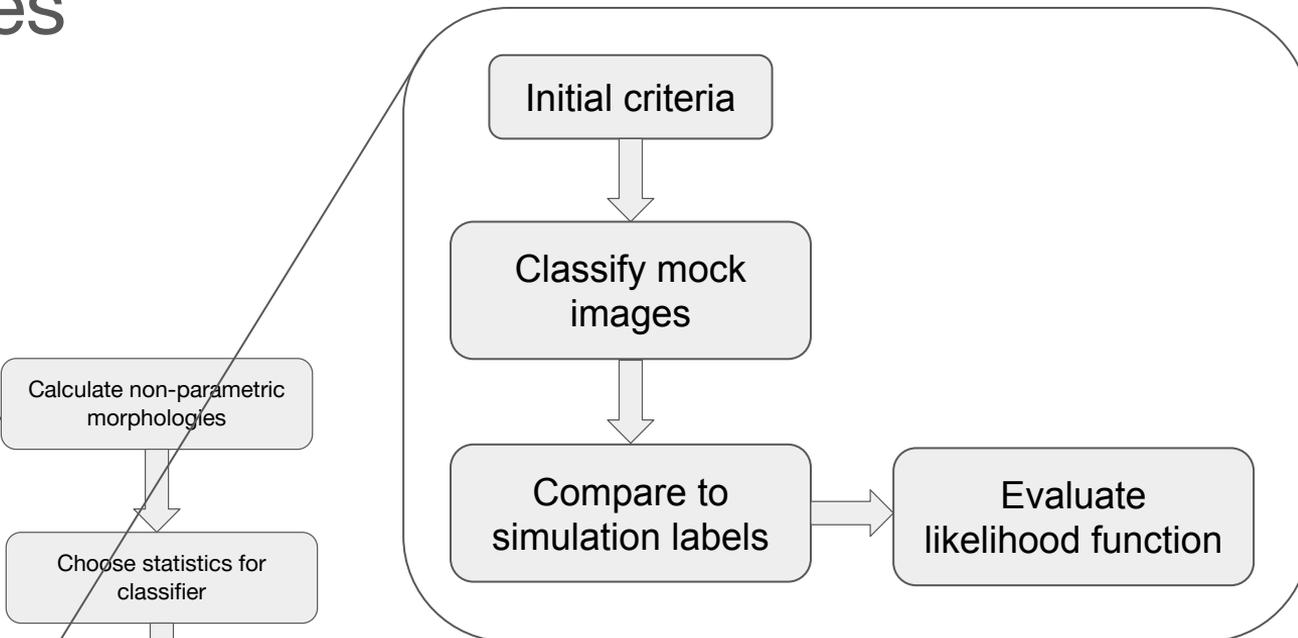
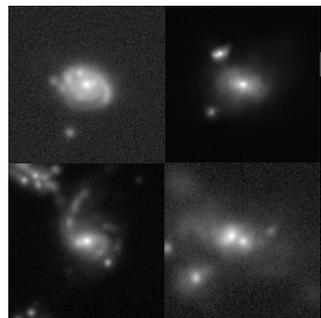
Updated merger morphologies

Can morphologies beat AI?



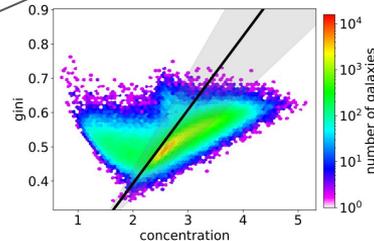
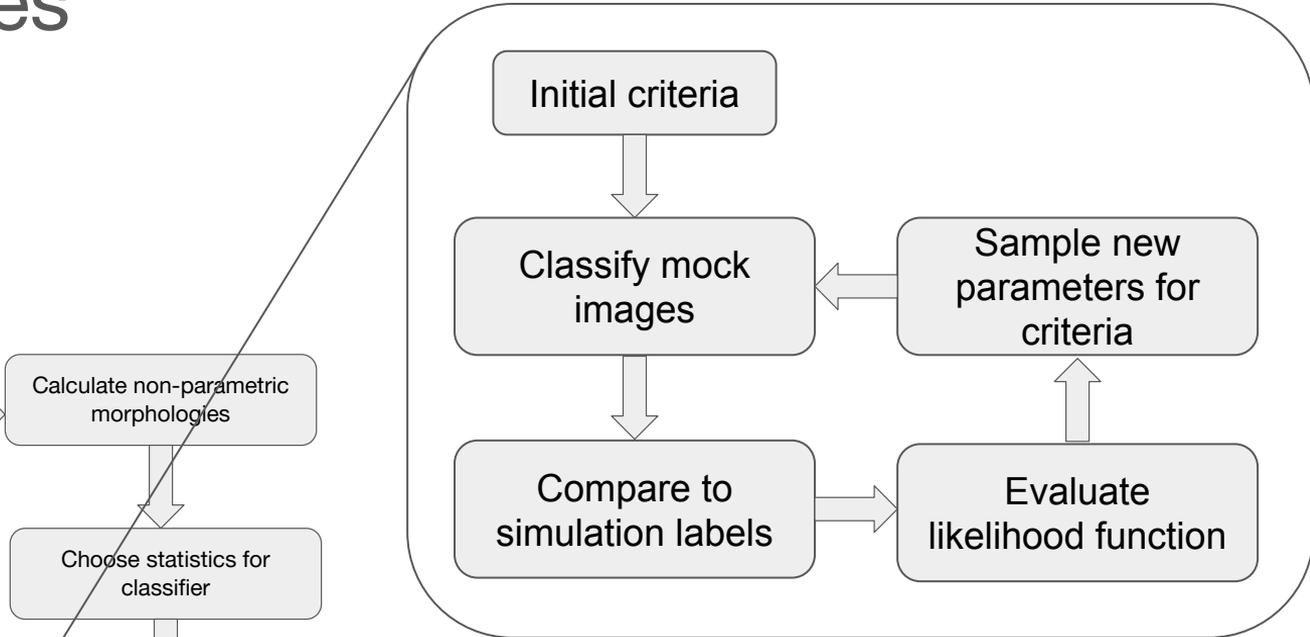
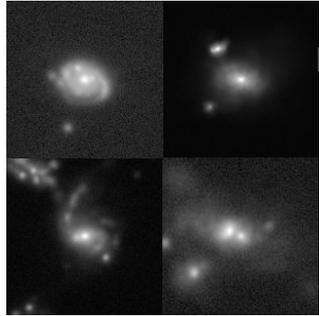
Updated merger morphologies

Can morphologies beat AI?



Updated merger morphologies

Can morphologies beat AI?



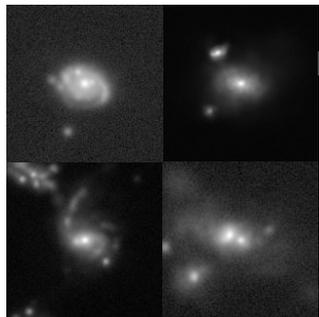
Updated merger morphologies

Can morphologies beat AI?

$$P(TP, FP) = e - e^{-\frac{FP+TP}{TP}}$$

True Positive

False Positive



Calculate non-parametric morphologies

Choose statistics for classifier

Optimise merger criteria with Markov Chain Monte-Carlo

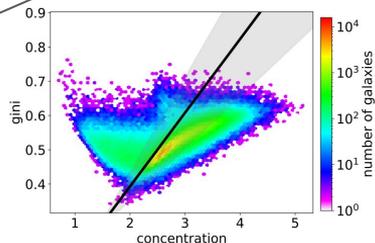
Initial criteria

Classify mock images

Sample new parameters for criteria

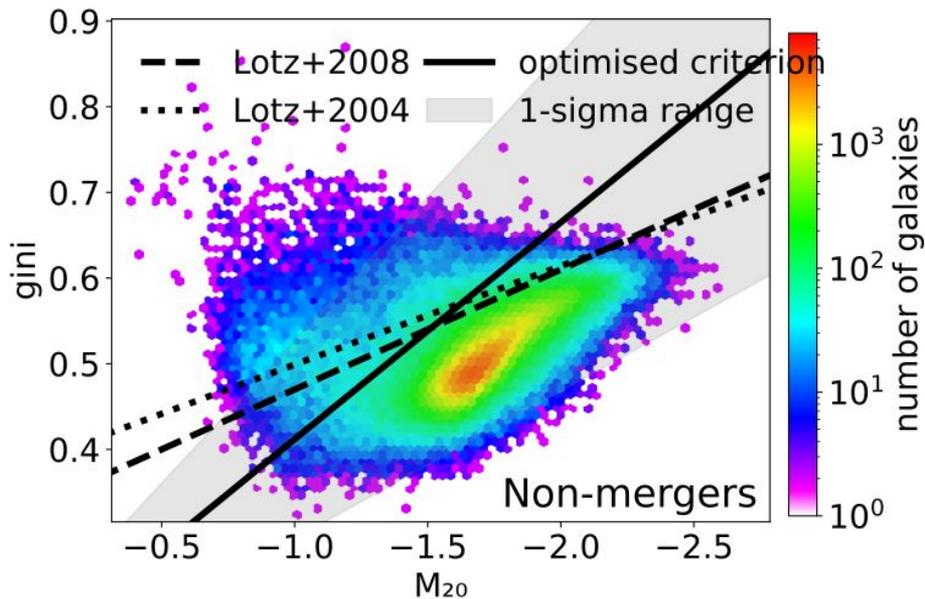
Compare to simulation labels

Evaluate likelihood function



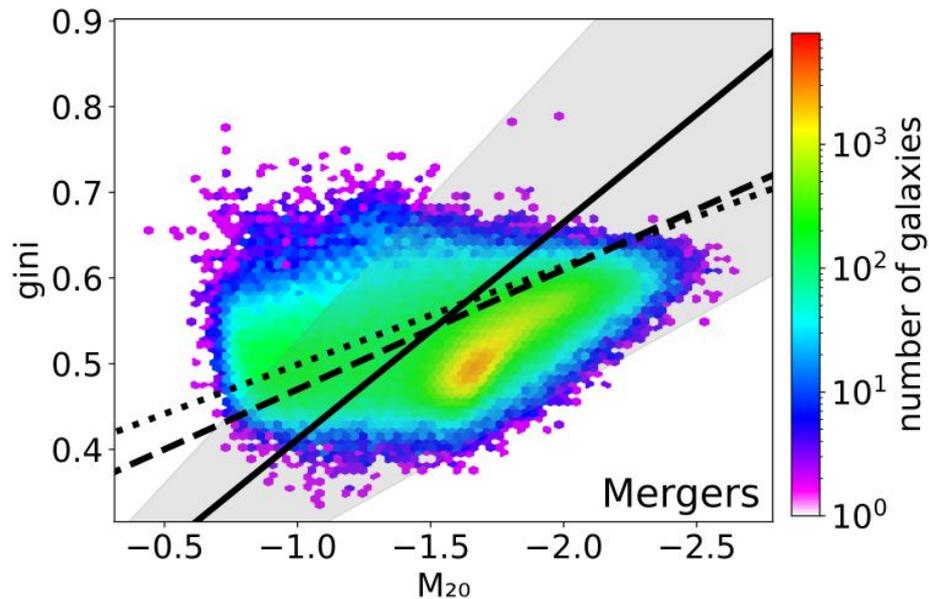
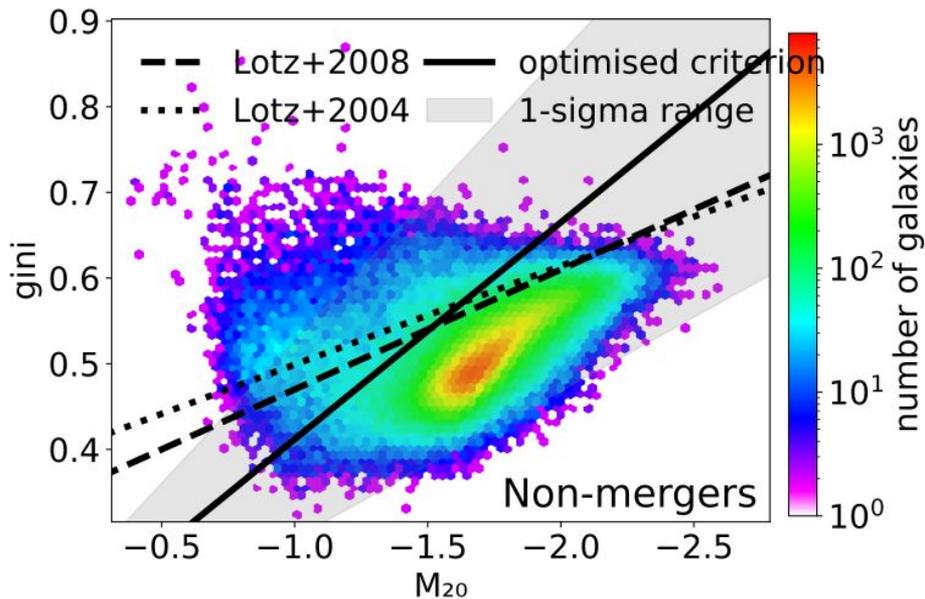
Updated merger morphologies

Can morphologies beat AI?



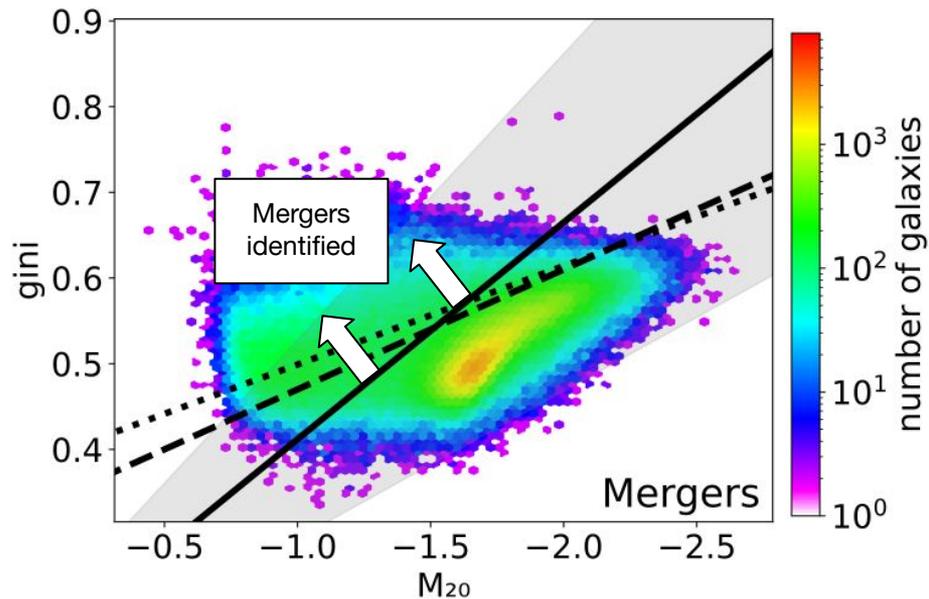
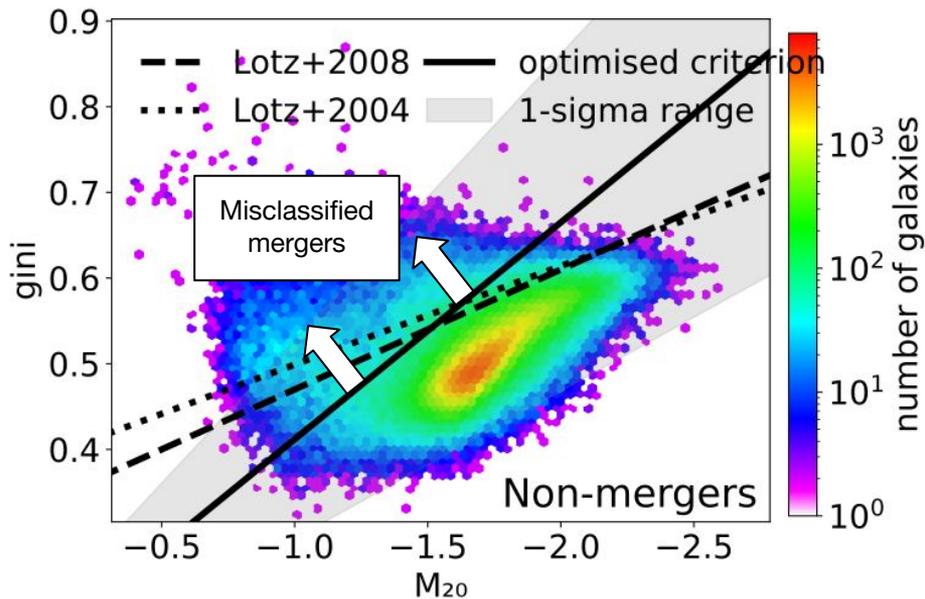
Updated merger morphologies

Can morphologies beat AI?



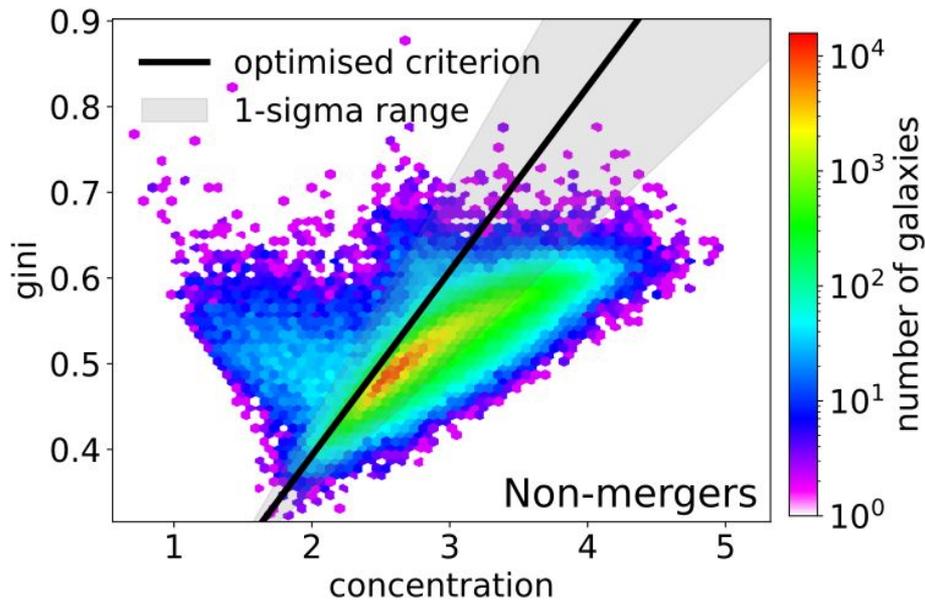
Updated merger morphologies

Can morphologies beat AI?



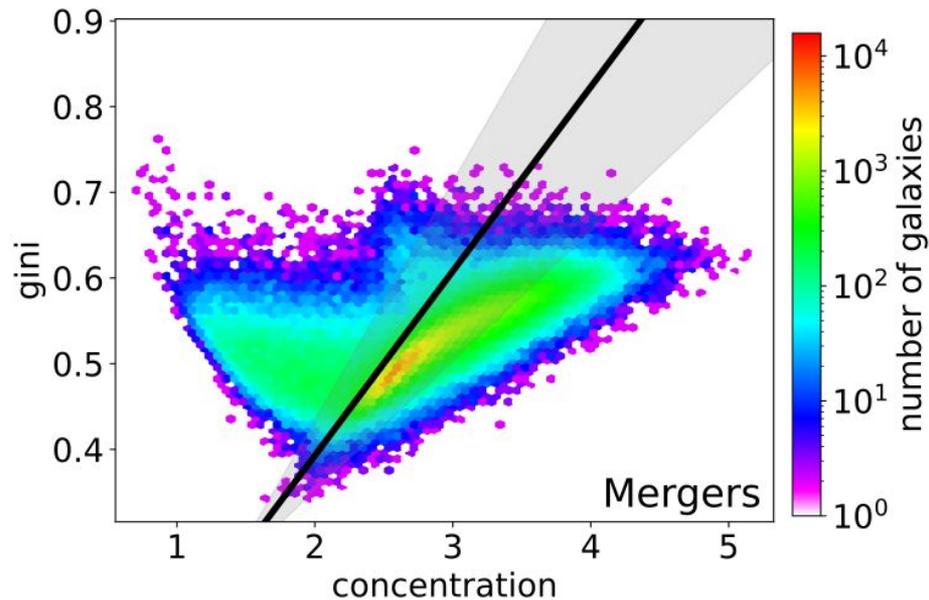
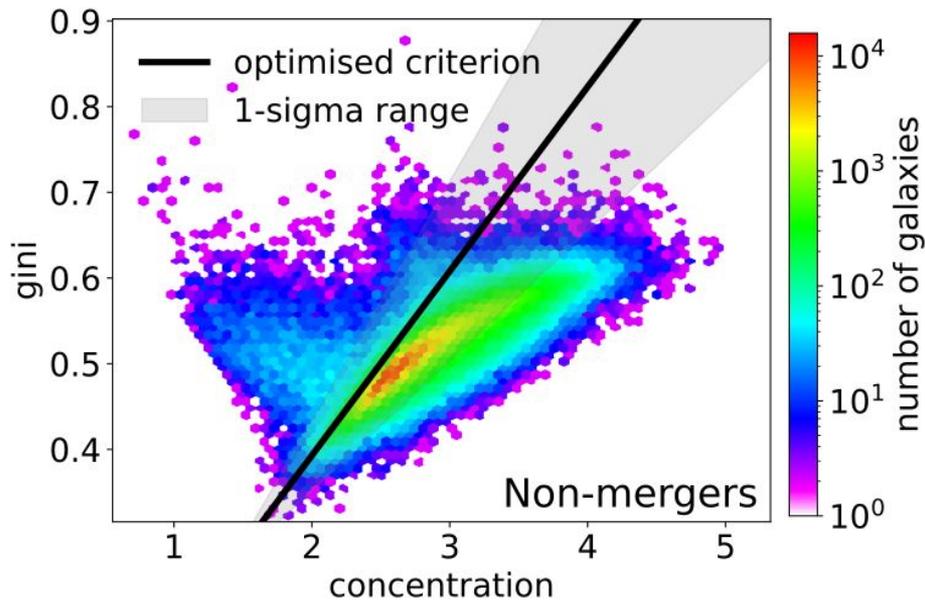
Updated merger morphologies

Can morphologies beat AI?



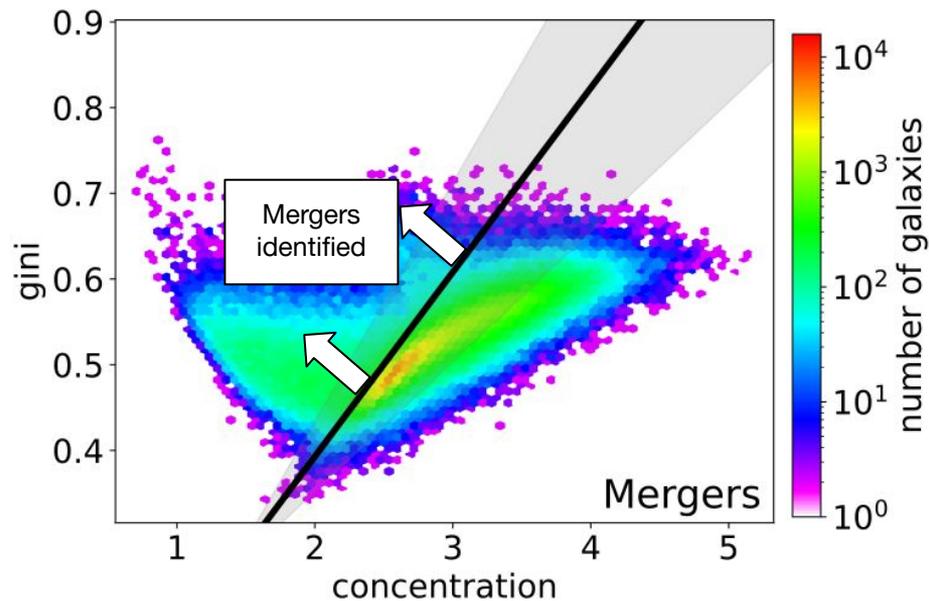
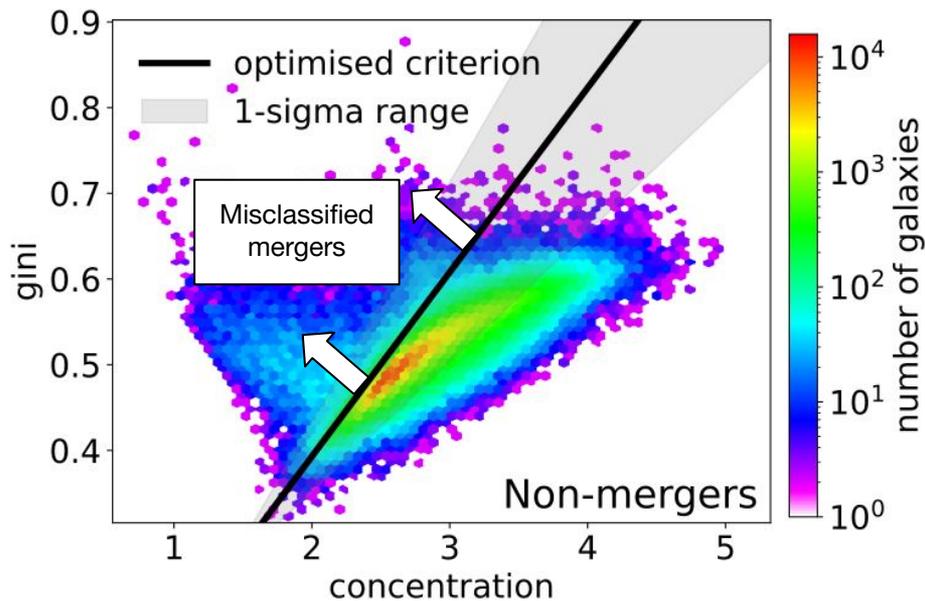
Updated merger morphologies

Can morphologies beat AI?



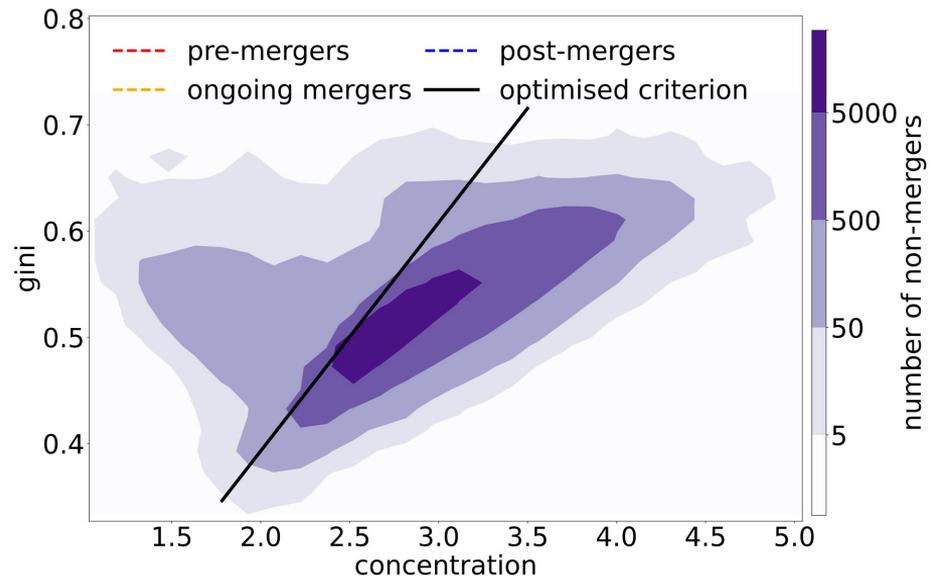
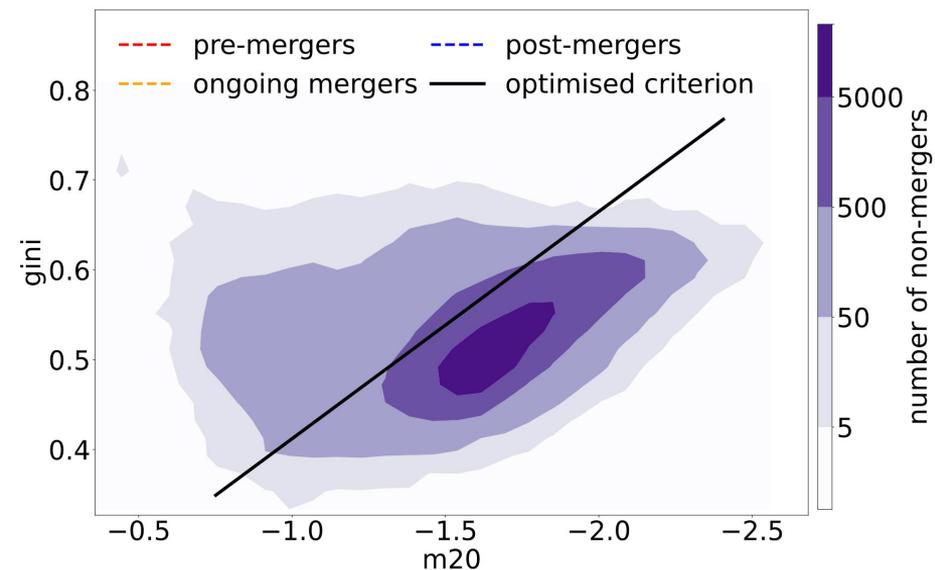
Updated merger morphologies

Can morphologies beat AI?



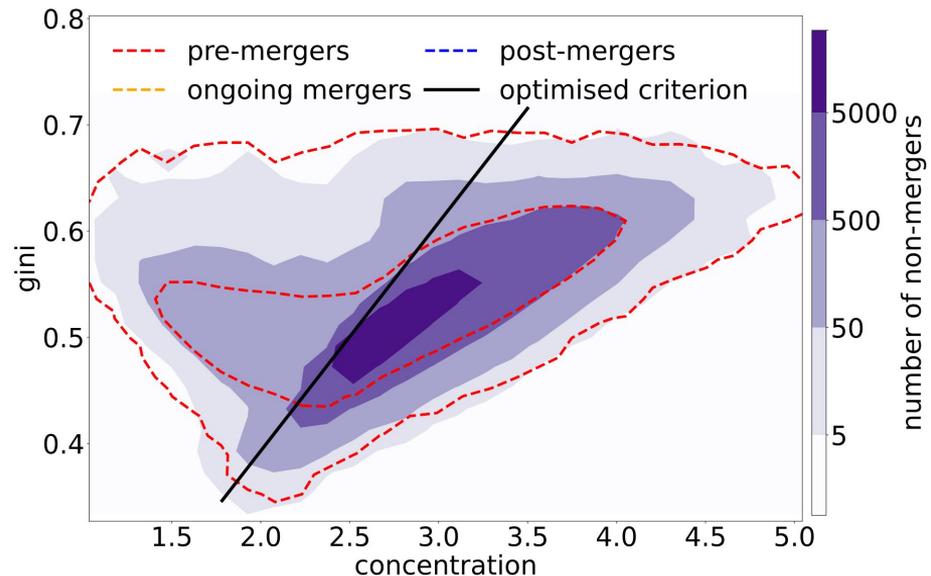
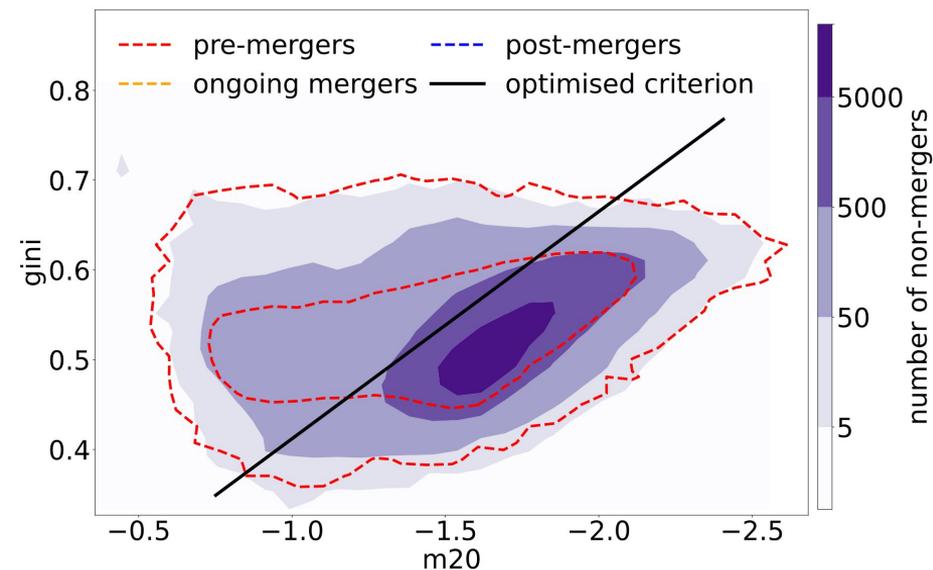
Updated merger morphologies

Can morphologies beat AI?



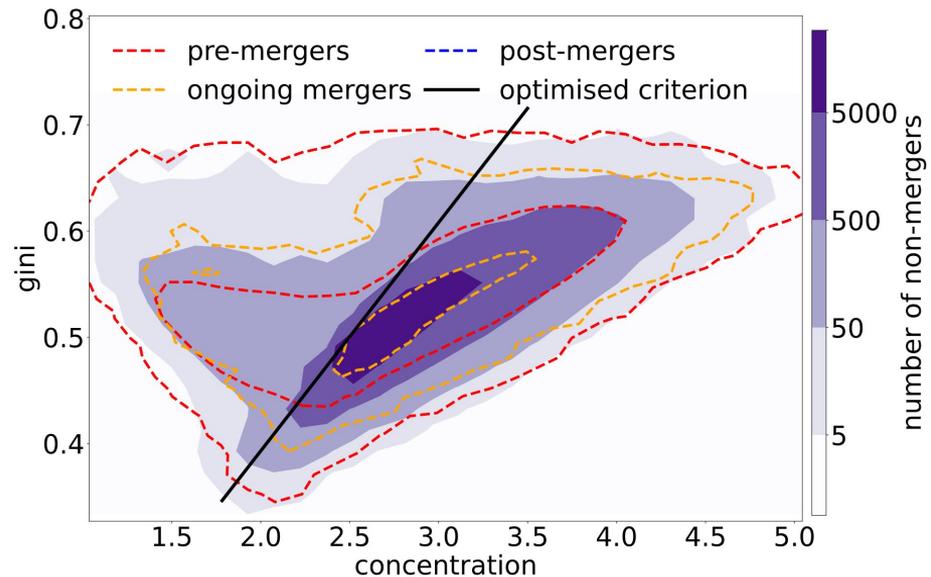
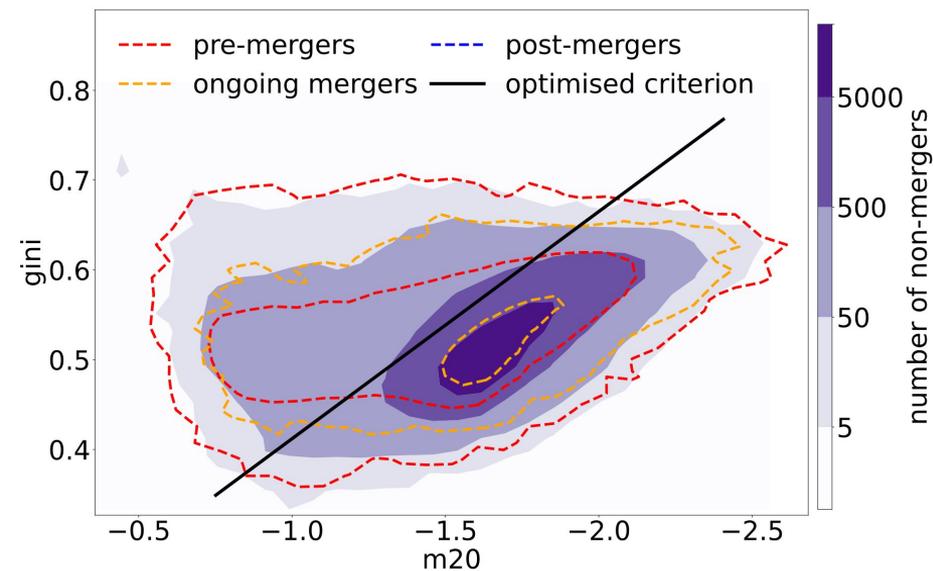
Updated merger morphologies

Can morphologies beat AI?



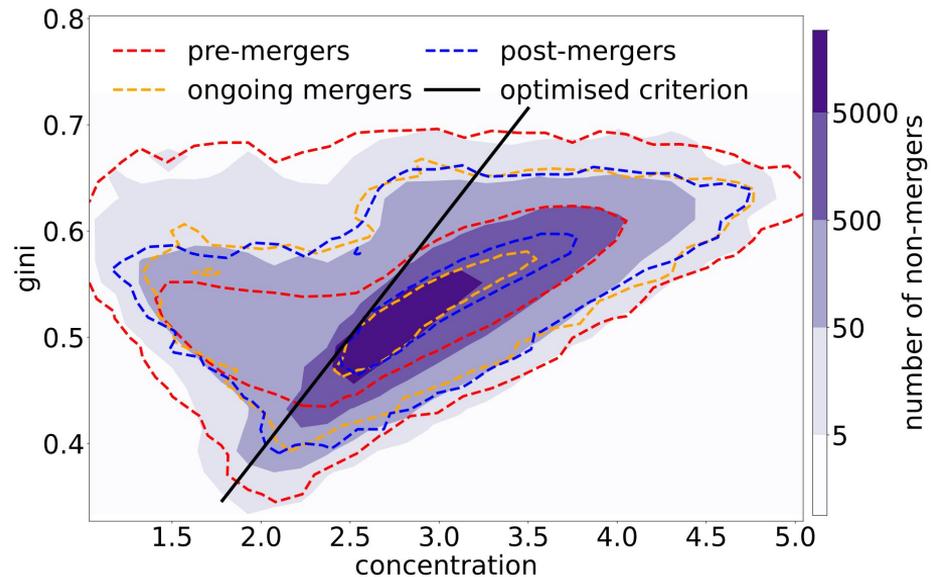
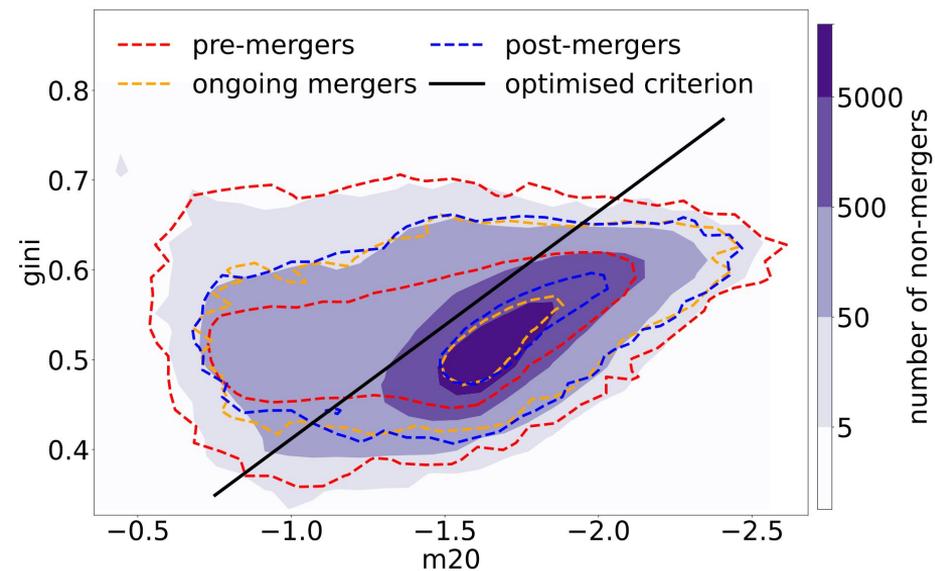
Updated merger morphologies

Can morphologies beat AI?



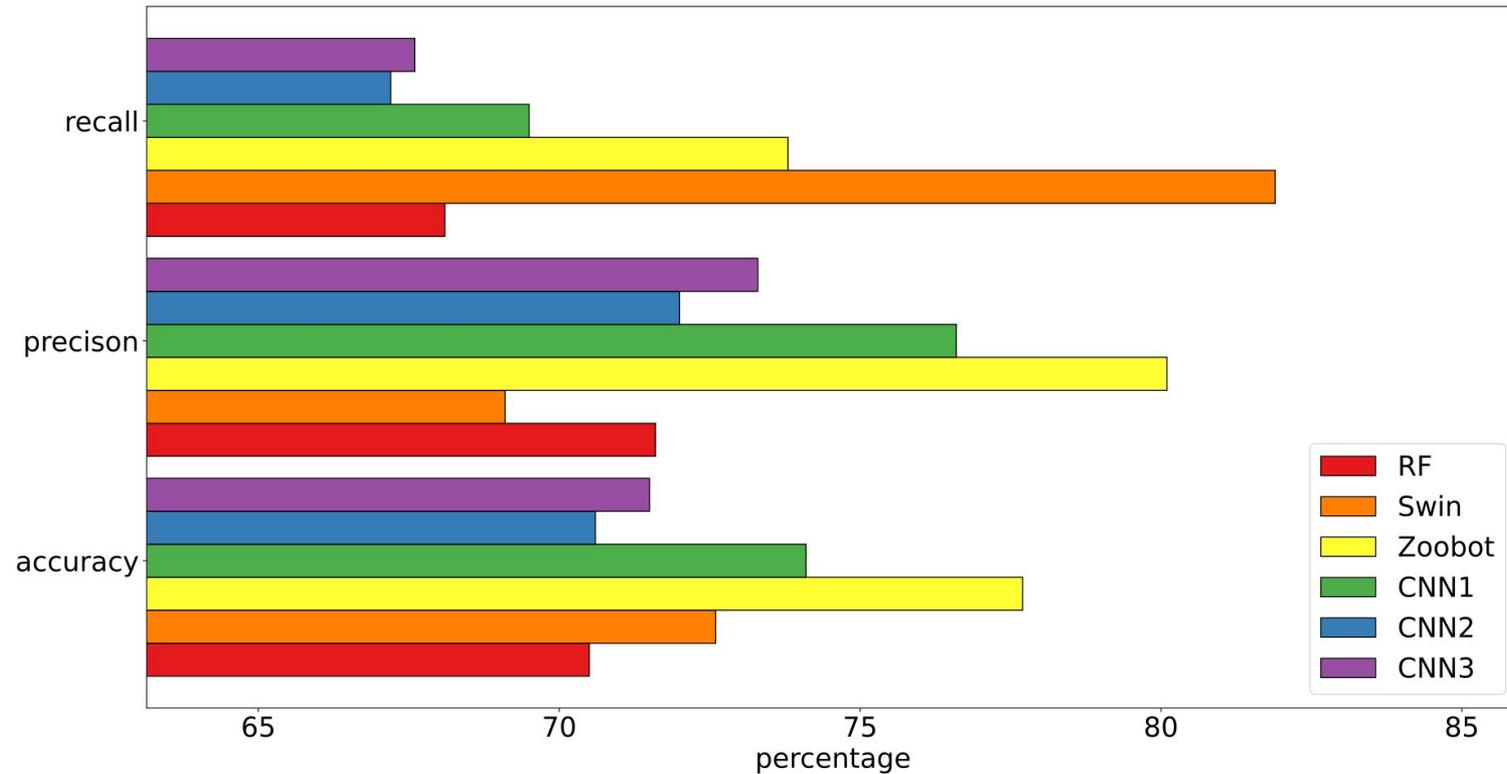
Updated merger morphologies

Can morphologies beat AI?



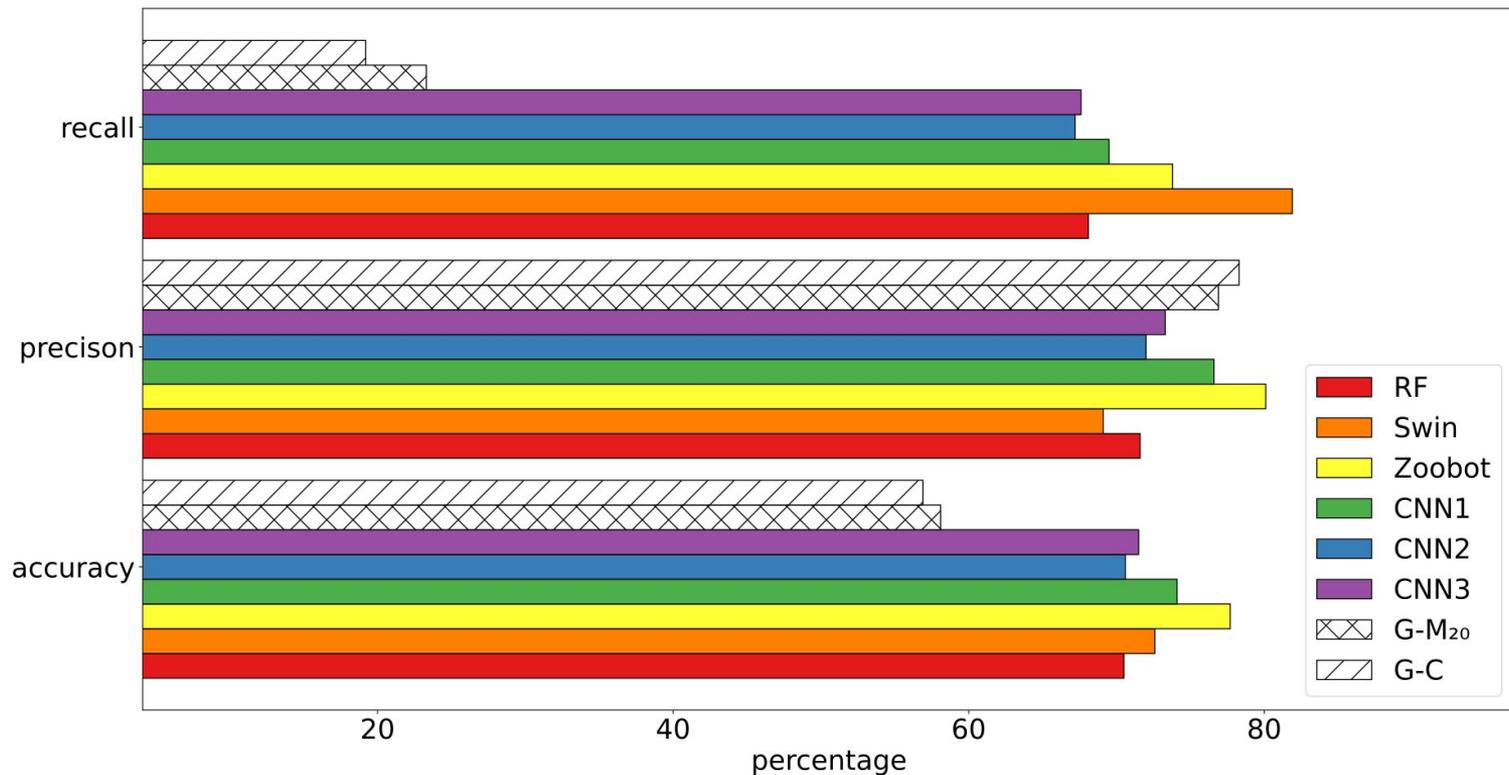
Performance comparison

Can morphologies beat AI?



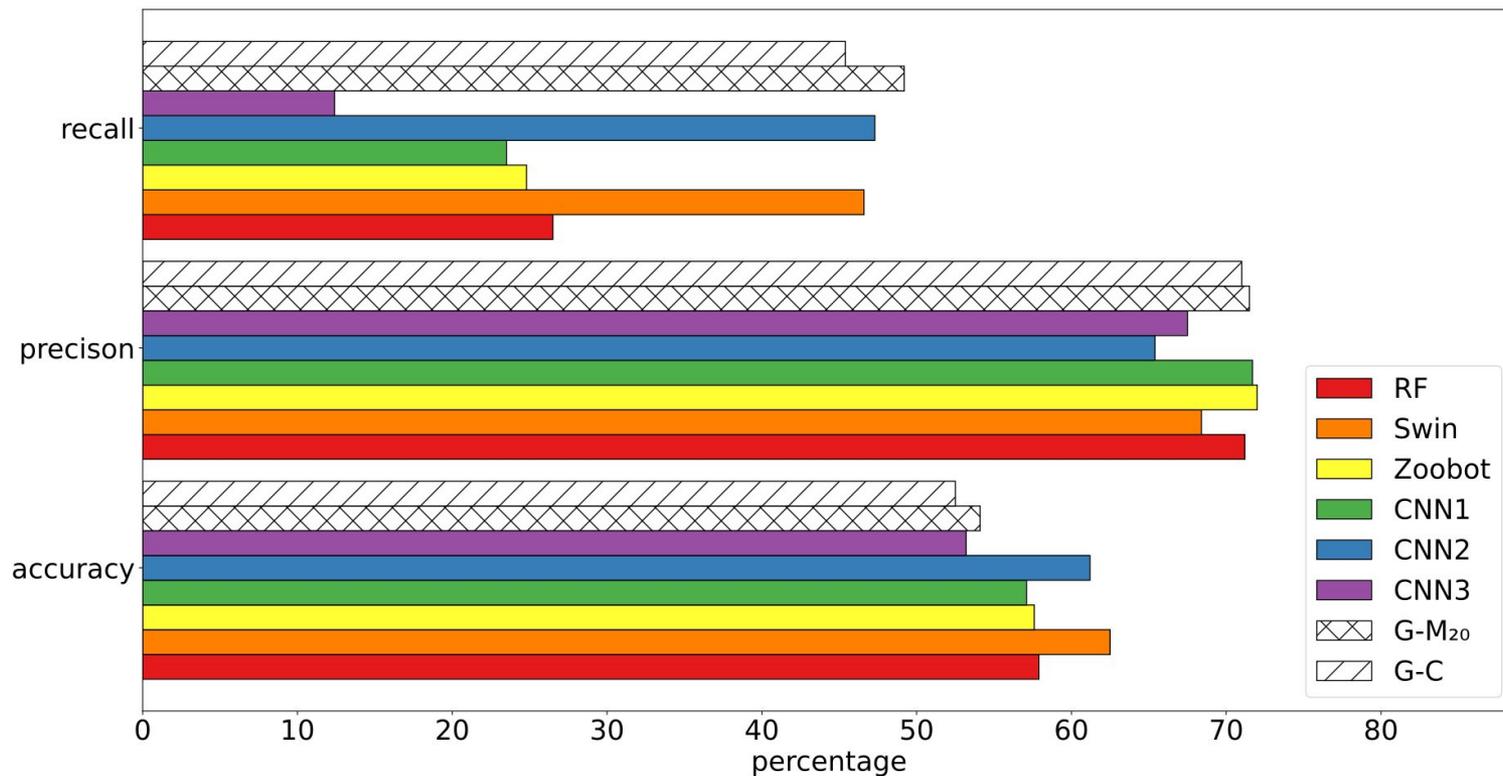
Performance comparison

Can morphologies beat AI?



Performance comparison

Can morphologies beat AI?



Using
Horizon-AGN
mock images

Introduction

Classification techniques

Can morphologies beat AI?

Future developments

Summary



Goals for new statistics

Future developments

Accurate

- Must work on ongoing and post mergers
- Should compete with some ML methods

Goals for new statistics

Future developments

Accurate

- Must work on ongoing and post mergers
- Should compete with some ML methods

Robust

- Should maintain accuracy across surveys
- Should work for poor image quality

Goals for new statistics

Future developments

Accurate

- Must work on ongoing and post mergers
- Should compete with some ML methods

Robust

- Should maintain accuracy across surveys
- Should work for poor image quality

Scalable

- Must be efficient enough to work on large surveys

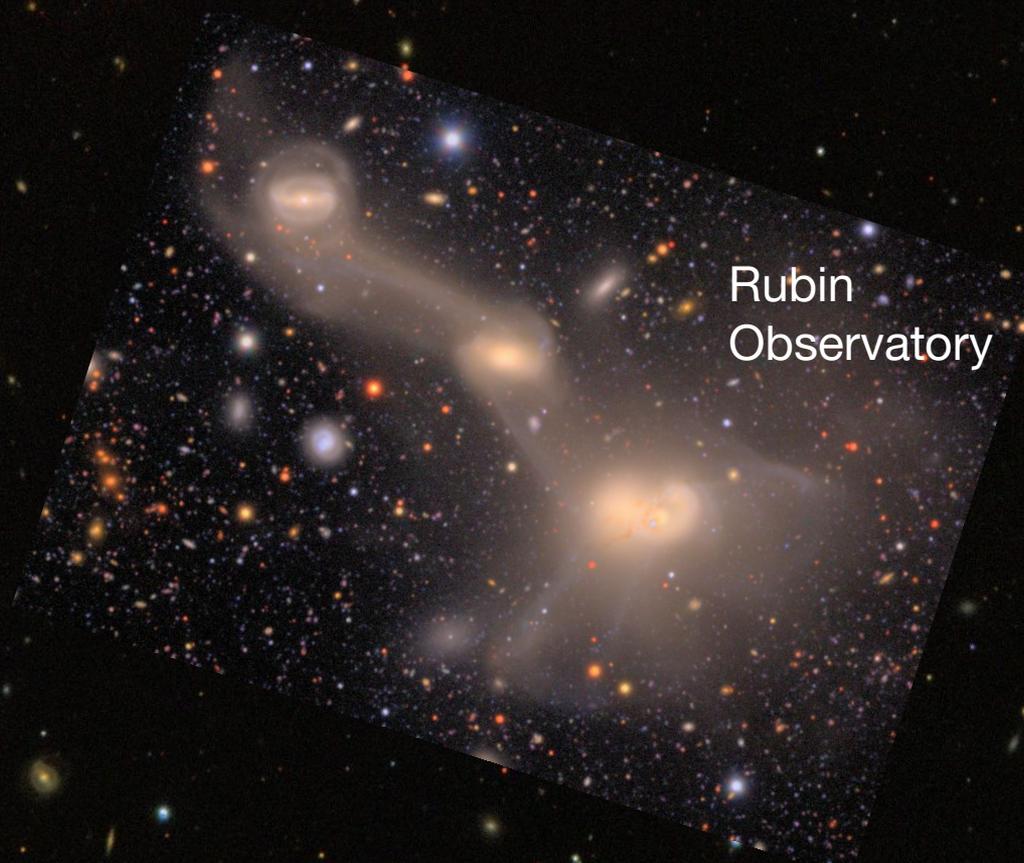
Faint structure

Future developments

Sloan Digital
Sky Survey

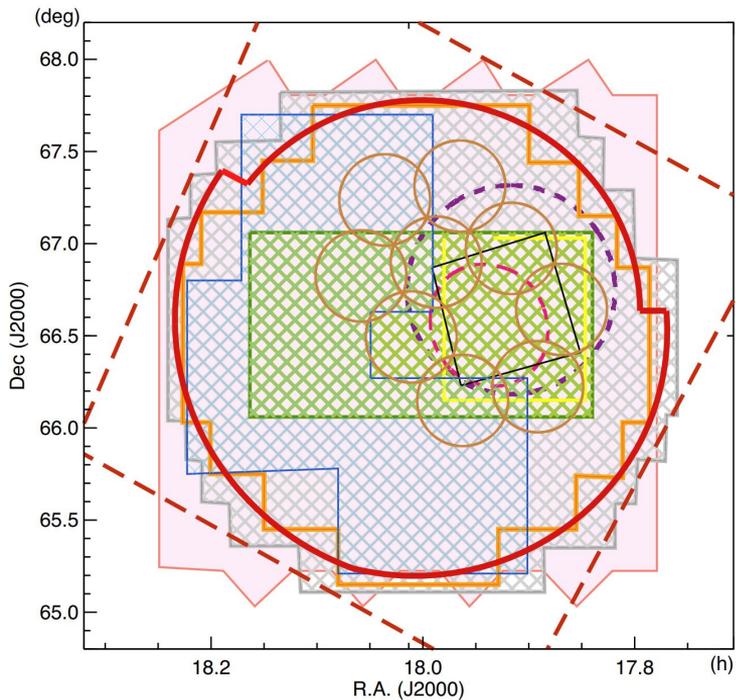
Faint structure

Future developments



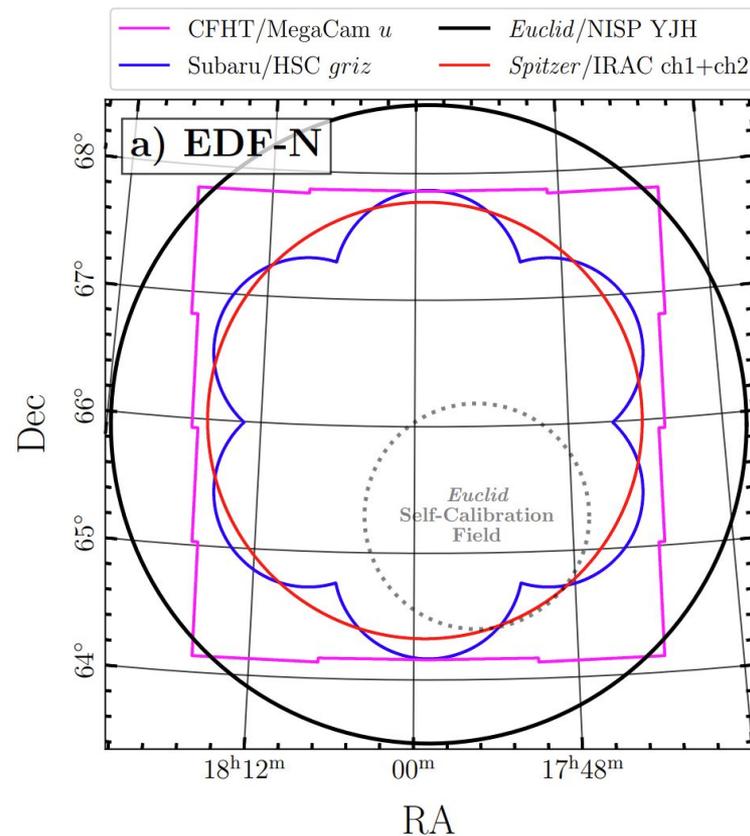
Rubin
Observatory

NEP as a testing ground

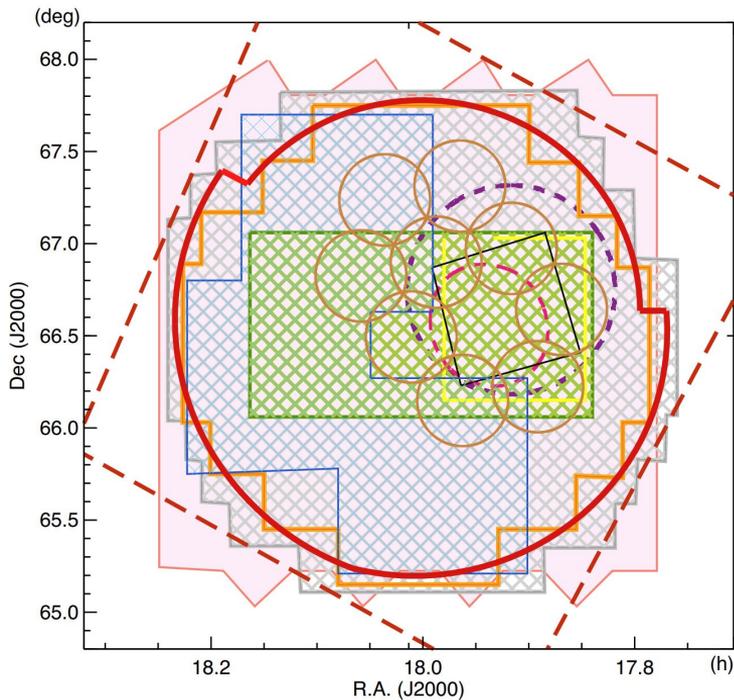


- AKARI NEP-Wide (Kim et al. 2012)
- CFHT/MegaCam (Hwang et al. 2007)
- Maidanak/SNUCam (Jeon et al. 2010)
- KPNO/FLAMINGOS (Jeon et al. 2014)
- CFHT/WIRCam (Oi et al. 2014)
- CFHT/Megaprime u (Huang et al. 2020)
- S2CLS/850μm (Geach et al. 2017)
- Spitzer/IRAC (Nayyeri et al. 2018)
- Herschel/PACS (Pearson et al. 2019)
- GALEX (Burgarella et al. 2019)
- Herschel/SPIRE (Pearson et al. 2017)
- NEPSC2/850μm (Shim et al. 2020)

Future developments



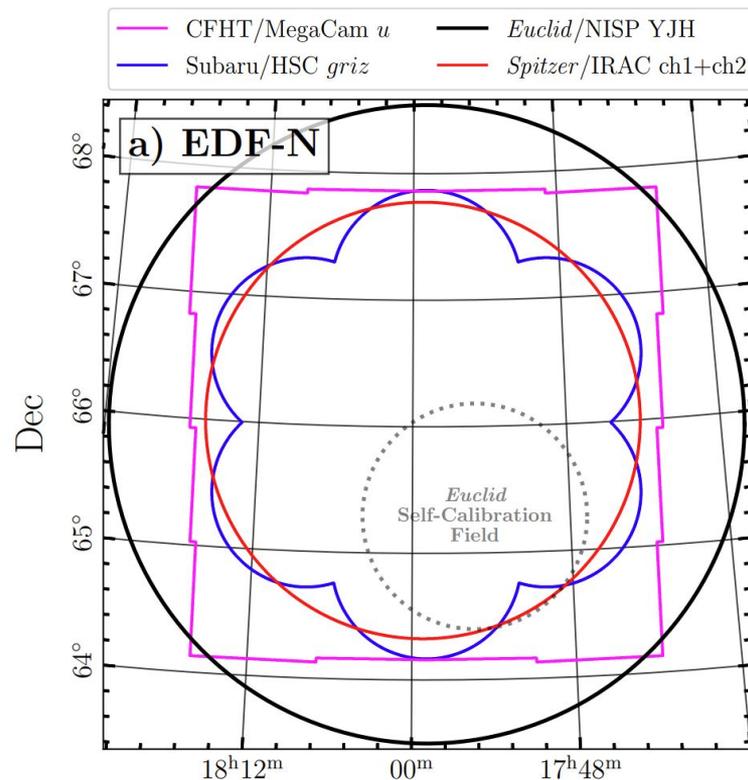
NEP as a testing ground



2020

- AKARI NEP-Wide (Kim et al. 2012)
- CFHT/MegaCam (Hwang et al. 2007)
- Maidanak/SNUCam (Jeon et al. 2010)
- KPNO/FLAMINGOS (Jeon et al. 2014)
- CFHT/WIRCam (Oi et al. 2014)
- CFHT/Megaprime u (Huang et al. 2020)
- S2CLS/850 μ m (Geach et al. 2017)
- Spitzer/IRAC (Nayyeri et al. 2018)
- Herschel/PACS (Pearson et al. 2019)
- GALEX (Burgarella et al. 2019)
- Herschel/SPIRE (Pearson et al. 2017)
- NEPSC2/850 μ m (Shim et al. 2020)

Future developments



2025

RA

Euclid collaboration 2025

30/32

Introduction

Classification techniques

Can morphologies beat AI?

Future developments

Summary



- 
- Mergers are a key part of galaxy evolution; a galaxy's morphology, star formation, AGN activity, and mass are strongly linked to its merger history
 - Classification of mergers can be done with visual classification, the close pairs method, machine learning, and morphological classification
 - Non-parametric morphologies describe the light distribution of a galaxy without assuming a model
 - Literature provides merger criteria based on Gini, M_{20} , and Asymmetry
 - Merger criteria were made using the Gini, M_{20} , and Concentration of mock images from IllustrisTNG and an MCMC method, matched ML methods in terms of precision but lacked accuracy
 - Found that existing non-parametric morphologies are ineffective for ongoing and post-mergers
 - New morphological statistics are under development to better classify mergers based on faint features

- Mergers are a key part of galaxy evolution; a galaxy's morphology, star formation, AGN activity, and mass are strongly linked to its merger history
- Classification of mergers can be done with visual classification, the close pairs method, machine learning, and morphological classification
- Non-parametric morphologies describe the light distribution of a galaxy without assuming a model
- Literature provides merger criteria based on Gini, M_{20} , and Asymmetry
- Merger criteria were made using the Gini, M_{20} , and Concentration of mock images from IllustrisTNG and an MCMC method, matched ML methods in terms of precision but lacked accuracy
- Found that existing non-parametric morphologies are ineffective for ongoing and post-mergers
- New morphological statistics are under development to better classify mergers based on faint features



Astronomy on Tap

Chat with astronomers over beers



The Taps Pub

Henryka Sienkiewicza 4

March 10th 2026

19:00

Will we ever know what
Dark Energy is?

How do stars end their
lives?

Host Speakers

Nicola Principi

Prof Wojciech Hellwing



Cosmology in crisis,
why this is actually
good news!

Dr Priscila Pessi

Stellar deaths

