Astrofit2 - 3rd Annual meeting

Mario Pasquato

ARTIStIC – ARTificial Intelligence Search for Intermediate-mass black holes in star Clusters

This project has received funding from the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement No 664931

Papers since our last meeting

with machine learning with my host

Submitted

- *Mass and star formation rate of the host galaxies of compact binary mergers across cosmic time,* Artale M. C., Mapelli M., Bouffanais, Y., Giacobbo, N., Spera, M., Pasquato M. 2019, MNRAS Submitted
- *Detecting IMBHs with machine learning: I. feature-based supervised classification on MOCCA-SURVEY Database I simulations,* Pasquato, M., Mapelli, M., Askar, A., Giersz, M. 2019, A&A Submitted [a related draft was rejected by MNRAS]
- *Towards a theory of the dynamical clock - Evolution of the A+ indicator in Plummer models,* Pasquato, M. 2019, A&A Submitted
- *Multiple Stellar Populations in NGC 2808: a Case Study for Cluster Analysis,* Pasquato, M. & Milone, A. 2019, ApJ Submitted, astro-ph:1906.04983

Accepted

- *Further properties of the dynamical clock A+ indicator in a toy model of pure dynamical friction,* Pasquato, M. 2019, RevMexAA Accepted, astro-ph:1907.11965
- *Analytical solutions for the dynamical clock A+ indicator in a toy model of pure dynamical friction,* Pasquato, M. 2019, RevMexAA Accepted, astro-ph:1907.11964
- *Radial Dependence of the Proto-Globular Cluster Contribution to the Milky Way Formation,* Chung, C., Pasquato, M., Lee, S.-Y., Di Carlo, U.N., An, D., Yoon, S.-J., Lee, Y.-W. 2019 ApJL Accepted, astro-ph: 1909.01353
- *Clustering clusters: unsupervised machine learning on globular cluster structural parameters,* Pasquato, M. & Chung, C. 2019, MNRAS Accepted, astro-ph:1901.05354

Published

- *Extended halo of NGC 2682 (M 67) from Gaia DR2,* Carrera, R., Pasquato, M., Vallenari, A., Balaguer-Núñez, L., Cantat-Gaudin, T., Mapelli, M., Bragaglia, A., Bossini, D., Jordi, C., Galadí-Enríquez, D., Solano, E. 2019, A&A, 627, A119
- *Finding Black Holes with Black Boxes - Using Machine Learning to Identify Globular Clusters with Black Hole Subsystems,* Askar, A., Askar, A., Pasquato, M. & Giersz, M. 2019, MNRAS, 485, 5345
- *Merging black holes in young star clusters* Di Carlo, U. N., Giacobbo, N., Mapelli, M., Pasquato, M., Spera, M., Wang, L. & Haardt, F. 2019, MNRAS, 487, 2947

Conferences since our last meeting

- **Invited:** *Ringberg Workshop on Machine Learning in Astronomy*, Ringberg, Germany 08/12/2019 – 13/12/2019 (upcoming)
- *Artificial Intelligence in Astronomy,* ESO Garching, Germany 22/07/2019 *Image-in science out? A proof of concept with deep learning on molecular cloud simulations.*
- *Galaxy Coffee,* MPIA Heidelberg, Germany 18/07/2019 *Applying machine learning to astronomy, beyond simple classification towards automatic science*
- Talks at Prof. H.-W. Rix (MPIA Heidelberg director), Dr. Annalisa Pillepich, and Prof. Nadine Neumeier group meetings.
- *KASI Colloquium,* Korea Astronomy and Space Science Institute, Korea 11/06/2019 *Cluster analysis for multiple populations in GCs – NGC 2808 as a case study*
- *IAU Symposium 351 & MODEST-19,* INAF Bologna, Italy, 27/05/2019 Poster: *Finding IMBHs with machine learning: encouraging results*

Outreach and other activities since our last meeting

- **European Researchers Night, Padua Astronomical Observatory, Italy, 27/09/2019** gave a `Young Researcher' talk on *Astronomy and Artificial Intelligence*
- **Outreach talk at Padua Planetarium, Italy, 06/04/2019**

Artificial Intelligence in Astronomy

Mario Pasquato

L'intelligenza artificiale nell'astronomia

- **PI of Italian Super Computing Resource Allocation C-class projects** DLSCHIMB, 200k hours (Padua, 2018); DISSM67, 400k hours (Padua, 2019)

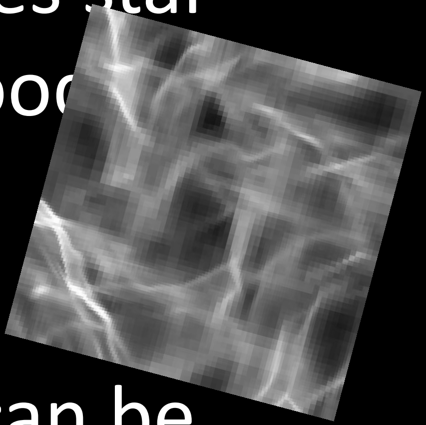
Mentoring and teaching since our last meeting

Tobia Peruzzi, Padua University, Italy, currently ongoing Masters student, Preliminary thesis title: *Visualizing AGN spectra with t-SNE*. Expected defense date in March 2020.

Piero Trevisan, UNIPD, Padua, Italy, 07/03/2019 Masters student, Thesis title: *Deep Convolutional Neural Networks in Astrophysics: a case study for gas turbulence*, Grade: 110/110 summa cum laude

Optics Laboratory, UNIPD, Italy
06/05/2019 – 26/06/2019
Esperimentazioni di fisica II

Detour: deep learning the spectral index of turbulence in molecular clouds



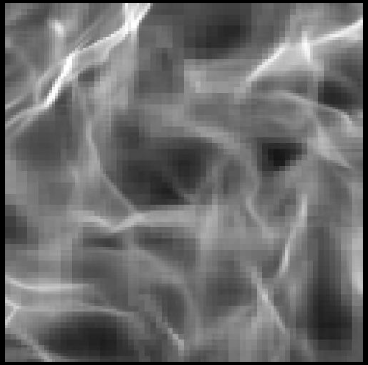
- Turbulence in molecular clouds modulates star formation, physics still not fully understood [Elmegreen & Scalo 2004, Hennebelle & Falgarone 2012]
- Velocity power spectrum of turbulence can be measured directly through e.g. line-of-sight velocity [Koch 2019]
- Can we measure it from projected density maps (images) using neural networks?

Detour: deep learning the spectral index of turbulence in molecular clouds

Hard

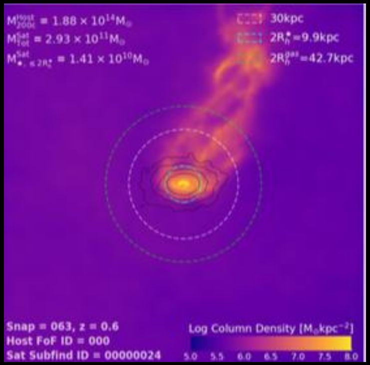
IMBH host? Yes/No

Intermediate



Kolmogorov turbulence? Yes/no

Easy



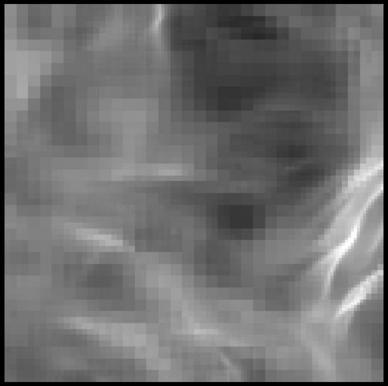
Jellyfish galaxy? Yes/No

gitlab.com/mariomario/jellyfishtng

Not so much: it is a very similar supervised classification problem to my main project... but an easier one. I am effectively preparing for it through this

Question

- Can we measure the turbulence index of simulated turbulent gas from density maps?
- In particular discriminate between Kolmogorov $P_v(k) = k^{-11/3}$ and Burgers $P_v(k) = k^{-4}$ spectra

Simulations

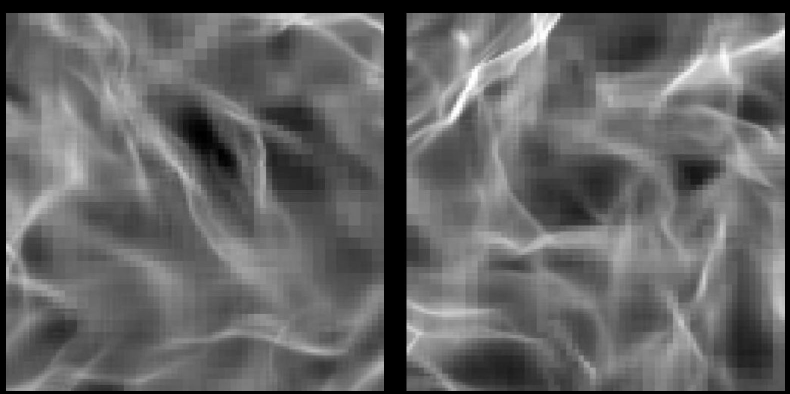
- 1000 simulations of turbulent gas with RAMSES2 [Teyssier 2002] AMR code
- 10x10x10 pc box, initially uniform density gas ($6.77\times10^{-22} g/cm^3$), total mass of $10^4 M_{sun}$.
- Gas kept isothermal at temperature T=10K
- Injected a divergence free, turbulent, supersonic (Mach 1.41) velocity field with spectrum index n=11/3 or 4
- Evolved for 0.5 Myr, solving Euler's equation with a Lax-Friedrichs Riemann Solver, periodic boundaries without self-gravity and magnetic fields

Train/test/holdout split

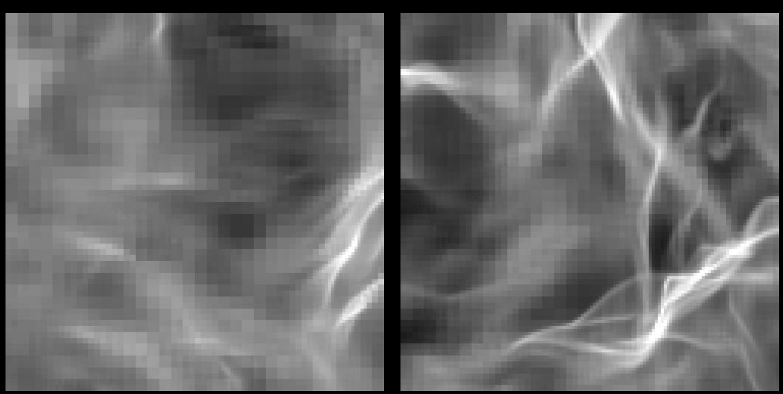
- 500 sims w. Kolmogorov index, 500 w. Burgers
- 400+400 build the train set -> 3 projections (x,y,z) X 4 flip/flop X 4-way cut = 38400 training images
- 50+50 in the test set = 4800 test images
- 50+50 never looked at (holdout set) = 4800 images

Images

- 250x250 pixels, grayscale; each image corresponds to ¼ of the box, seen in projection along an axis (x,y,z)
- Luminosity encodes log column density



Kolmogorov



Burgers

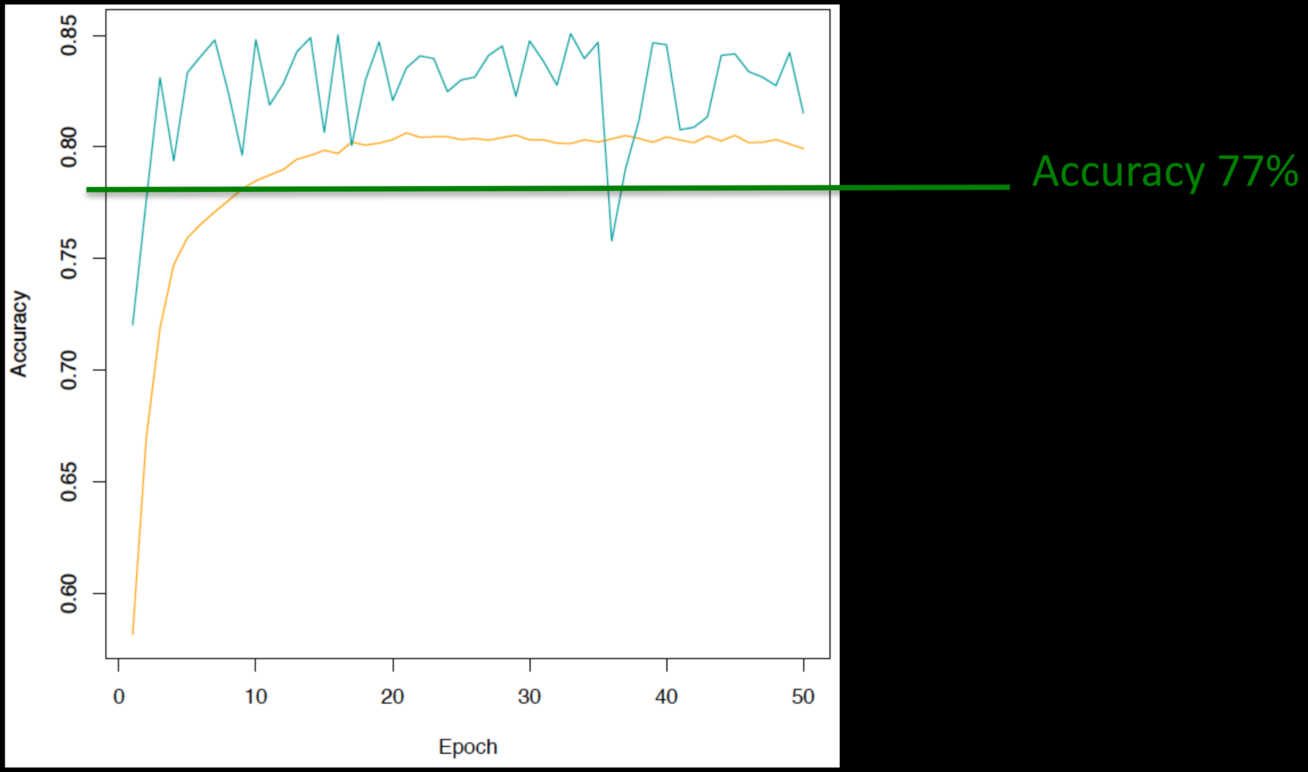
DL setup

- Keras on top of Tensorflow on workstation with a Titan V GPU
- Four convolutional layers (with max pooling) + three dense layers
- RELU activations
- Dropout regularization *
- RMSprop optimizer

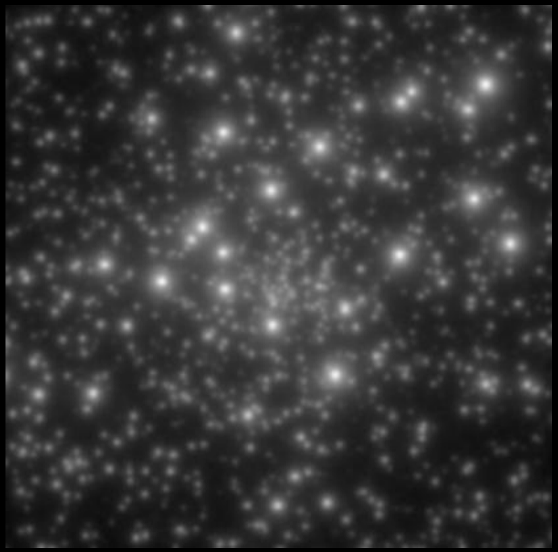
* Question for Julian

Results on holdout set

	Predicted Kolmogorov	Predicted Burgers
Kolmogorov	2113	287
Burgers	812	1588



Back to my main project: structural parameters as features = BH subsystem Yes/No?



Askar, Askar, Pasquato, Giersz
2019 MNRAS 485, 5345

2000 state-of-the art simulations of star clusters with the MOCCA code (Hypki & Giersz 2013, Giersz et al. 2013)

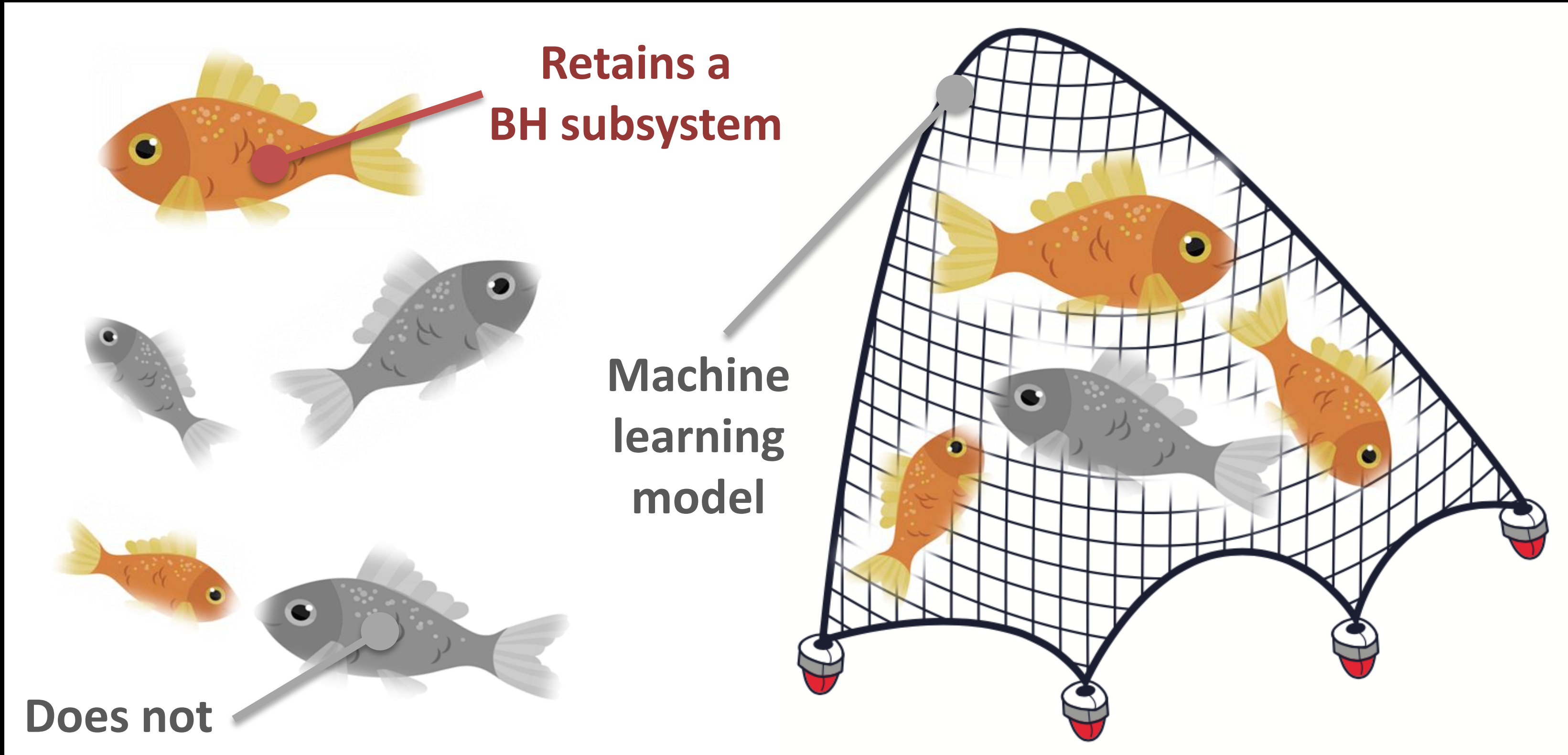
Largest realistic star cluster simulation dataset to date (Askar et al. 2017)

- Half-Light Radius
- Central Surface Brightness
- Central Velocity Dispersion
- Total Luminosity
- Relaxation Time
- Core Radius

Features

Catching BH subsystem hosts

- Which initial conditions + evolutionary history (as reflected by the structural parameters) lead to retaining a BH subsystem?

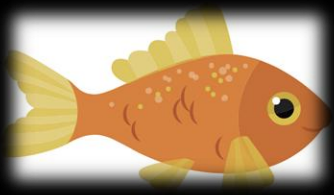


Desired model properties

- Catches (almost) only real BH subsystem hosts
- Interpretable

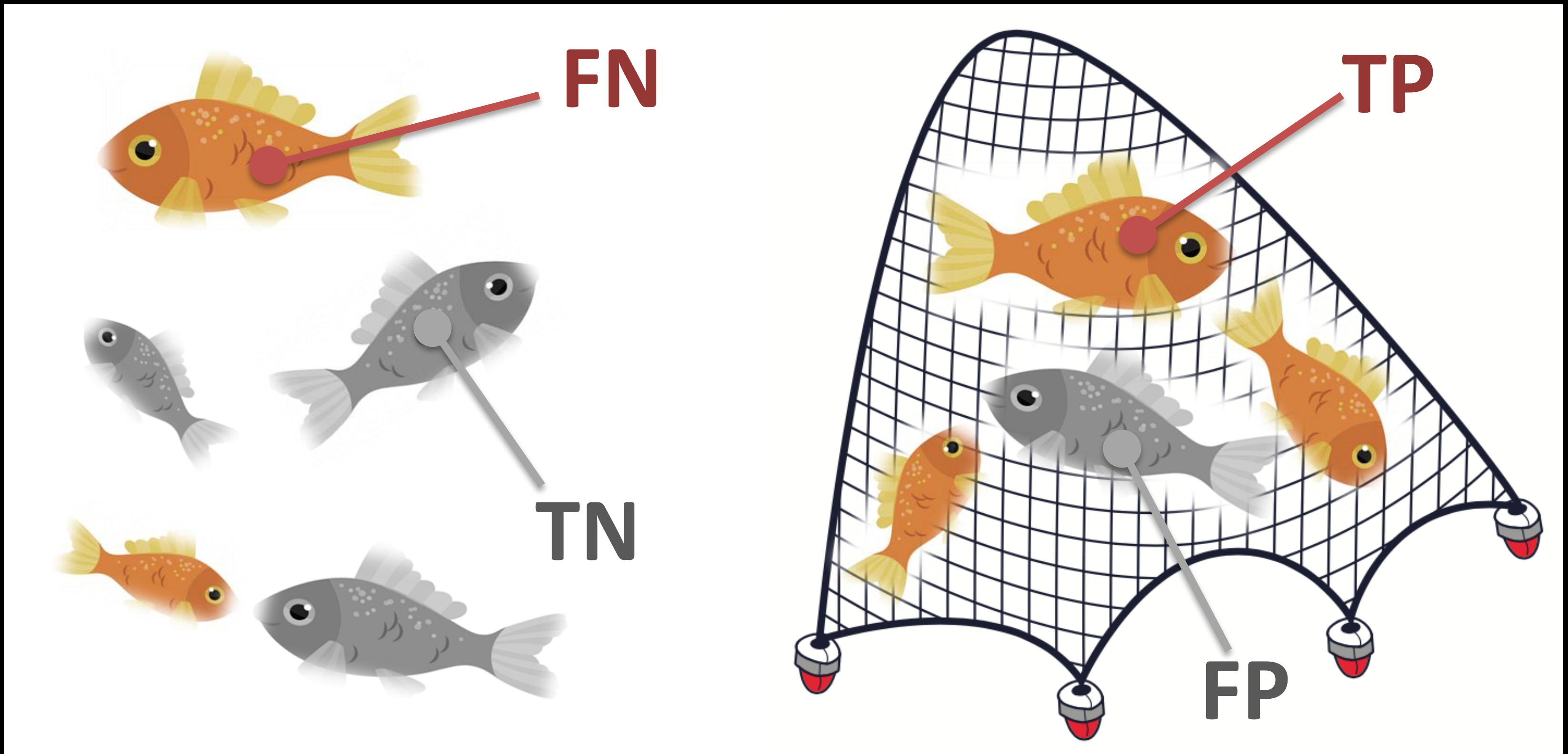
What makes a star cluster
a BH subsystem host?

Performance metric
should weigh
precision



Performance metrics

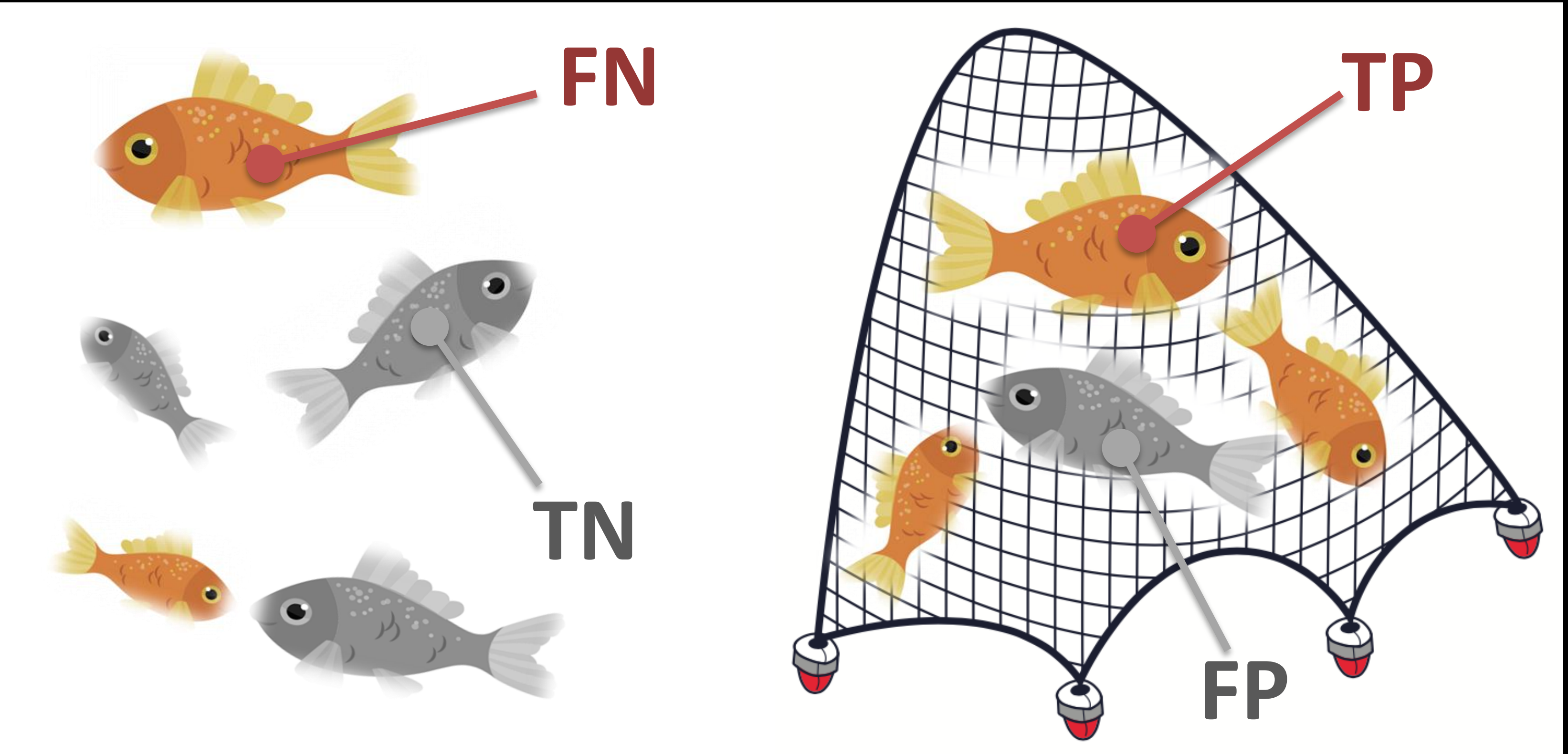
- Precision TP/(TP+FP) how clean is the catch?
- True Positive Rate or Recall TP/(TP+FN) how big is the catch?
- F-score = 1/(1/Precision + 1/Recall)



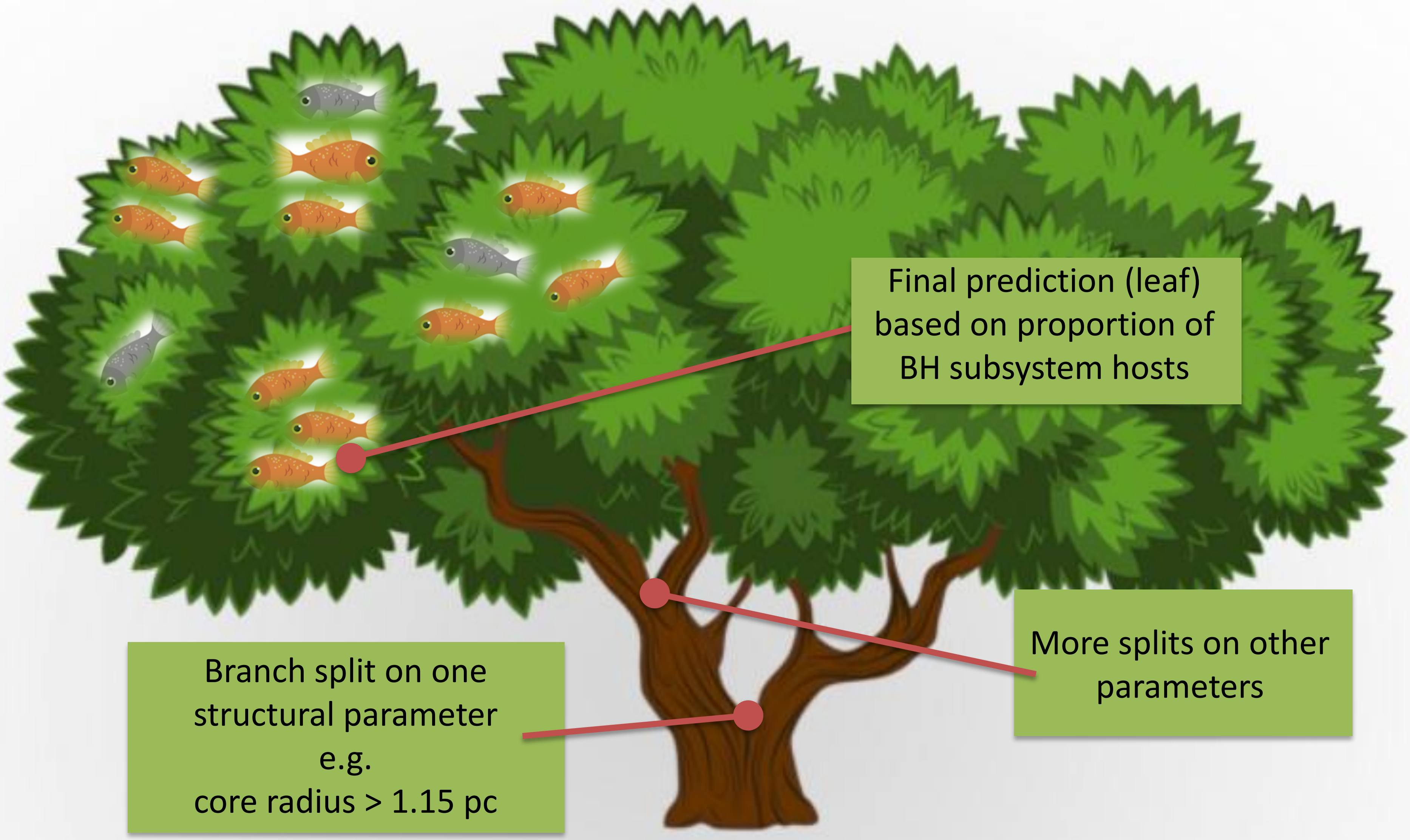
Performance metrics

- Precision 3/(3+1) = 3/4
- Recall 3/(3+2) = 3/5
- F-score = 1/(4/3 + 5/3) = 1/3

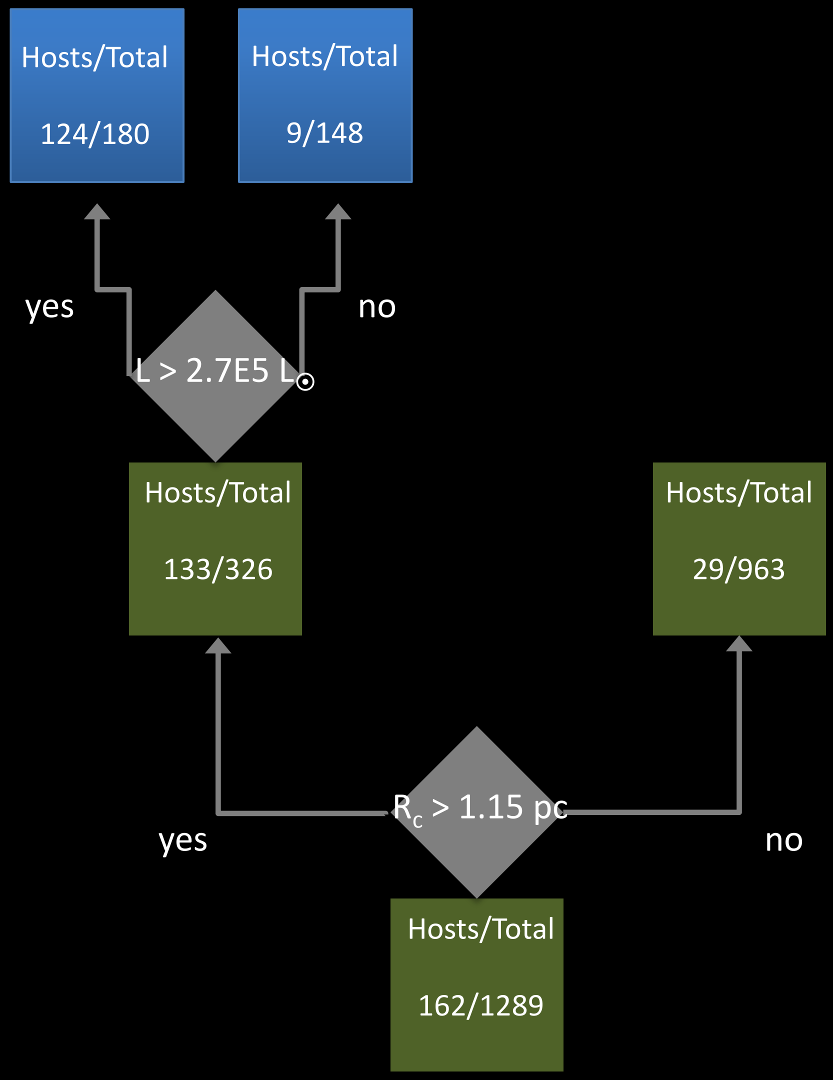
Example with numbers from this pic



Tree-based model



Physical interpretation



First few branches of the learned tree

First split is on core radius: black hole subsystem hosts have large cores due to dynamical heating

Second split on total luminosity: big clusters produce more black holes, have higher retention due to higher escape velocity

Tree models also have good performance

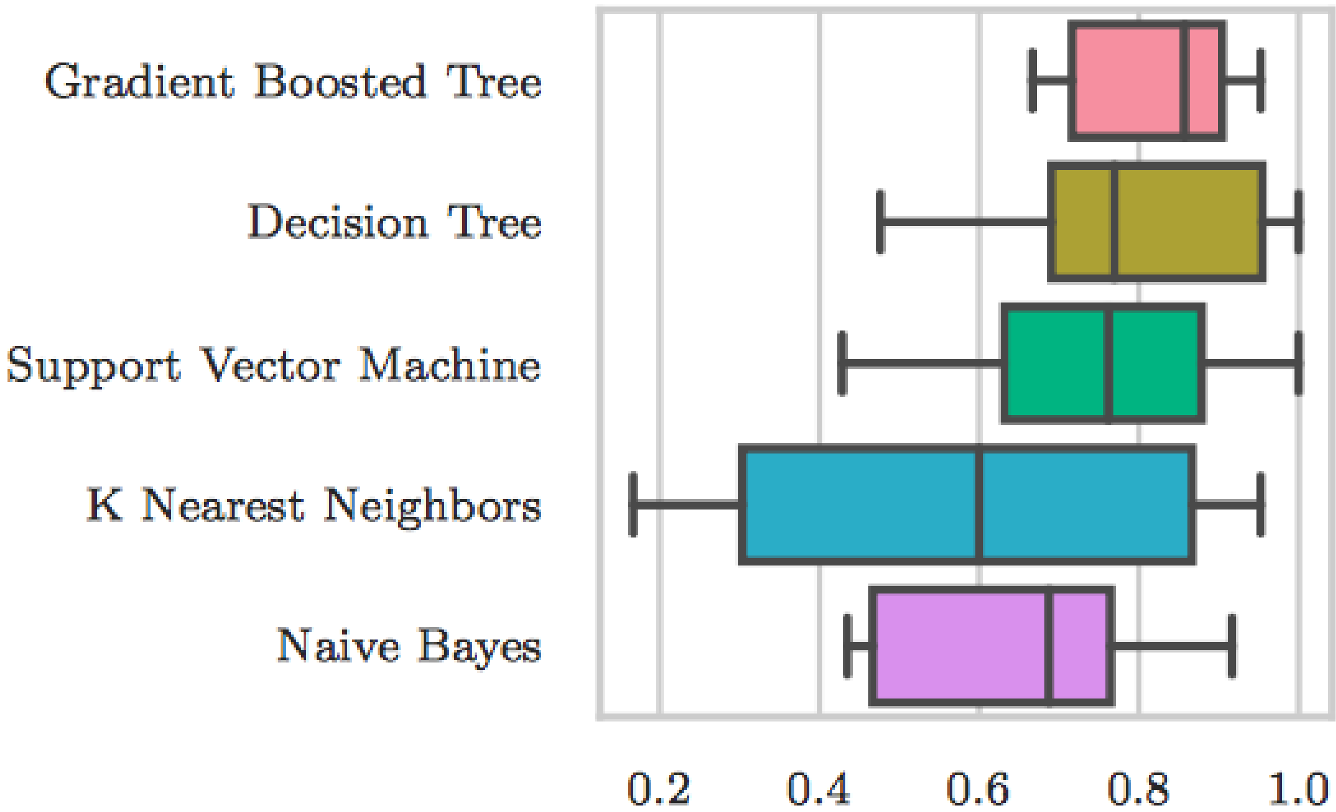


Figure 4. Comparison of each classifiers' f-score with 15-fold testing. An f-score of 1.0 is the best possible, 0.0 is the worst. The score is affected by how well the classifier can find all the BH subsystems and whether the identifications are false-positives.

Comparison with real GCs

Table 4. Predictions from the Harris (1996, *updated 2010*) and Baumgardt & Hilker (2018) datasets using the gradient boosted decision tree classifier. Entires where BHS presence was classified positively are shown. The *BHS* column represents the classifier trained on all simulation data whereas *Fallback* represents training on models where mass fallback was enabled and BH natal kicks were lower.

Green row = predicted BH subsystem host by all models e.g. NGC 288, M10

Results compare well with other methods (Askar et al. 2017 marked with *)

Cluster Name	BHS (Harris)	Fallback (Harris)	BHS (B&H)	Fallback (B&H)
IC 4499 *	✓	✓	✓	✓
NGC 288 *	✓	✓	✓	✓
NGC 3201 *	✓	✓	✗	✓
NGC 4372 *†	✓	✓	✓	✓
NGC 4590 (M68)	✗	✓	✗	✗
NGC 4833 *†	✗	✓	✓	✓
NGC 5139 (ω Cen)	✓	✓	✗	✓
NGC 5272 (M3) *	✗	✓	✓	✓
NGC 5286	✗	✓	✗	✓
NGC 5466 *	✗	✓	✗	✗
NGC 5897 *†	✓	✓	✗	✓
NGC 5904 (M5)	✗	✓	✓	✓
NGC 5927	✗	✗	✓	✓
NGC 5986 *†	✓	✓	✗	✓
NGC 6101 *†	✓	✓	✗	✗
NGC 6139 †	✓	✓	✗	✗
NGC 6144 *†	✗	✓	✓	✓
NGC 6205 (M13) *	✗	✓	✓	✓
NGC 6218 (M12)	✓	✓	✗	✗
NGC 6254 (M10)	✓	✓	✓	✓
NGC 6266 (M62)	✗	✗	✓	✗
NGC 6273 (M19) †	✗	✓	✓	✓
NGC 6287 †	✓	✓	✗	✗
NGC 6304 †	✗	✓	✓	✓
NGC 6316 †	✓	✓	✗	✗
NGC 6333 (M9) †	✓	✓	✗	✗
NGC 6356 †	✗	✓	✗	✓
NGC 6362 *	✗	✓	✗	✓
NGC 6380 †	✓	✓	✗	✗
NGC 6388	✗	✗	✓	✗
NGC 6401 *	✓	✓	✗	✗
NGC 6402 (M14) †	✓	✓	✓	✓
NGC 6426 *†	✗	✓	✗	✗
NGC 6440 †	✓	✓	✗	✗
NGC 6496 *†	✗	✓	✗	✗
NGC 6517 †	✗	✓	✗	✗
NGC 6539 (GCL 85)	✓	✓	✗	✗
NGC 6553	✓	✓	✗	✗
NGC 6569 *†	✓	✓	✓	✓
NGC 6584 *†	✓	✓	✗	✗
NGC 6656 (M22) *	✓	✓	✓	✓
NGC 6712 *	✗	✓	✓	✓
NGC 6723 *†	✓	✓	✓	✓
NGC 6760 †	✓	✓	✗	✗
NGC 6779 (M56) *	✗	✓	✓	✓
NGC 6809 (M55) *	✗	✓	✗	✓
NGC 6934 *	✓	✓	✗	✗
NGC 6981 (M72) *	✗	✓	✗	✗
NGC 7078 (M15)	✗	✗	✓	✓
NGC7089 (M2)	✗	✓	✗	✓
Pal11 *†	✓	✓	✓	✓
Terzan5 †	✗	✓	✗	✗

Conclusions

- Feature based approach is on track (paper submitted, tried and tested on a very slightly different problem – BH subsystems)
- Proof of concept with DL on images ready soon (Piero Trevisan thesis; paper in Prep.)