Classify and study Milky Way Globular cluster system based on color-magnitude diagram morphology using machine learning

Bishnu Kumar Sharma and Chow-Choong Ngeow

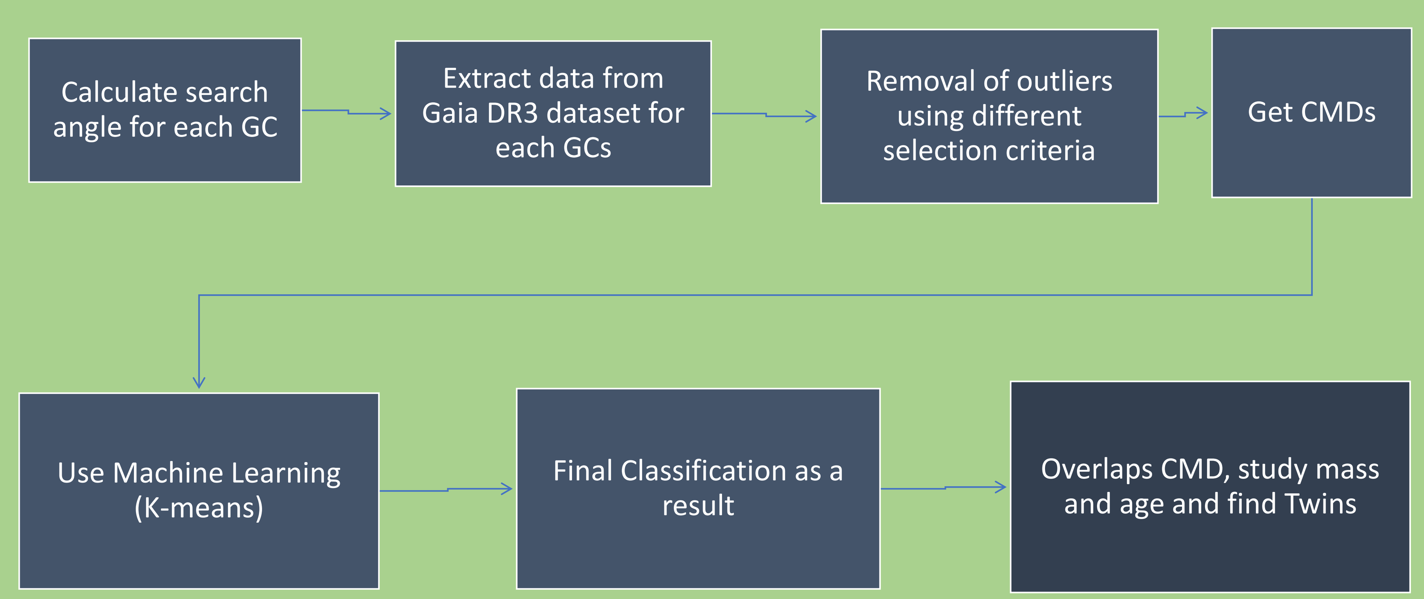
Graduate Institute of Astronomy, National Central University (Taiwan)

Email: bishnus350@gmail.com

Abstract: We used Gaia DR3 data to construct color-magnitude diagrams (CMDs) for the Milky Way Globular clusters (GCs). Our aim is to classify GCs based on their CMD morphology, such as main-sequence branch, red giant branch, and horizontal branch, using machine learning (ML) algorithm, as well as learning the properties of GCs based on ML classification. As a first step, we got CMDs using the information of position, proper motion, parallax and photometry from DR3 extracted using the VizieR Queries. We used clustering algorithm to find the most similar group of images. Then we superimposed them and study their properties like metallicity and age to find GCs twins based on their similar properties.

Introduction: We use five parametric astrometric data i.e. position (RA, Dec), proper motion (pmRA, pmDE), parallax and photometry (G_{mag}, G_{BP}, G_{RP}) from Gaia DR3 [1] , and construct CMDs and classify them using ML algorithm. We use Convolutional neural network (CNN), tensorflow.keras, and K-means clustering [5] algorithm for image classifications. K-means clustering classifies 'n' images into 'K' number of classes in which each observation belongs to cluster based on nearest distance.

Flow Chart:



Methods:

- We extracted the data from Gaia DR3 archive,
- Different selection cuts such as parallax [3], proper motion , Re-normalize unit weight error (RUWE) < 1.4 [4], epsi <= 2,
- CMD is a plot of BP-RP versus G_{mag},
- Scikit learn library, CNN, tensorflow.keras, and K-means clustering were used for image classification,
- After making groups, we superimposed individual CMDs into one CMD,
- We classified them into perfectly overlapping, perfectly non-overlapping and singular groups,
- We studied their properties such as age, metallicity, distance etc.
- Based on similar age and metallicity, we identified them as twins.

Some of the samples from different groups

Composite CMD: A plot of superimposed CMD of 14 GCs from lowest to highest metallicity.

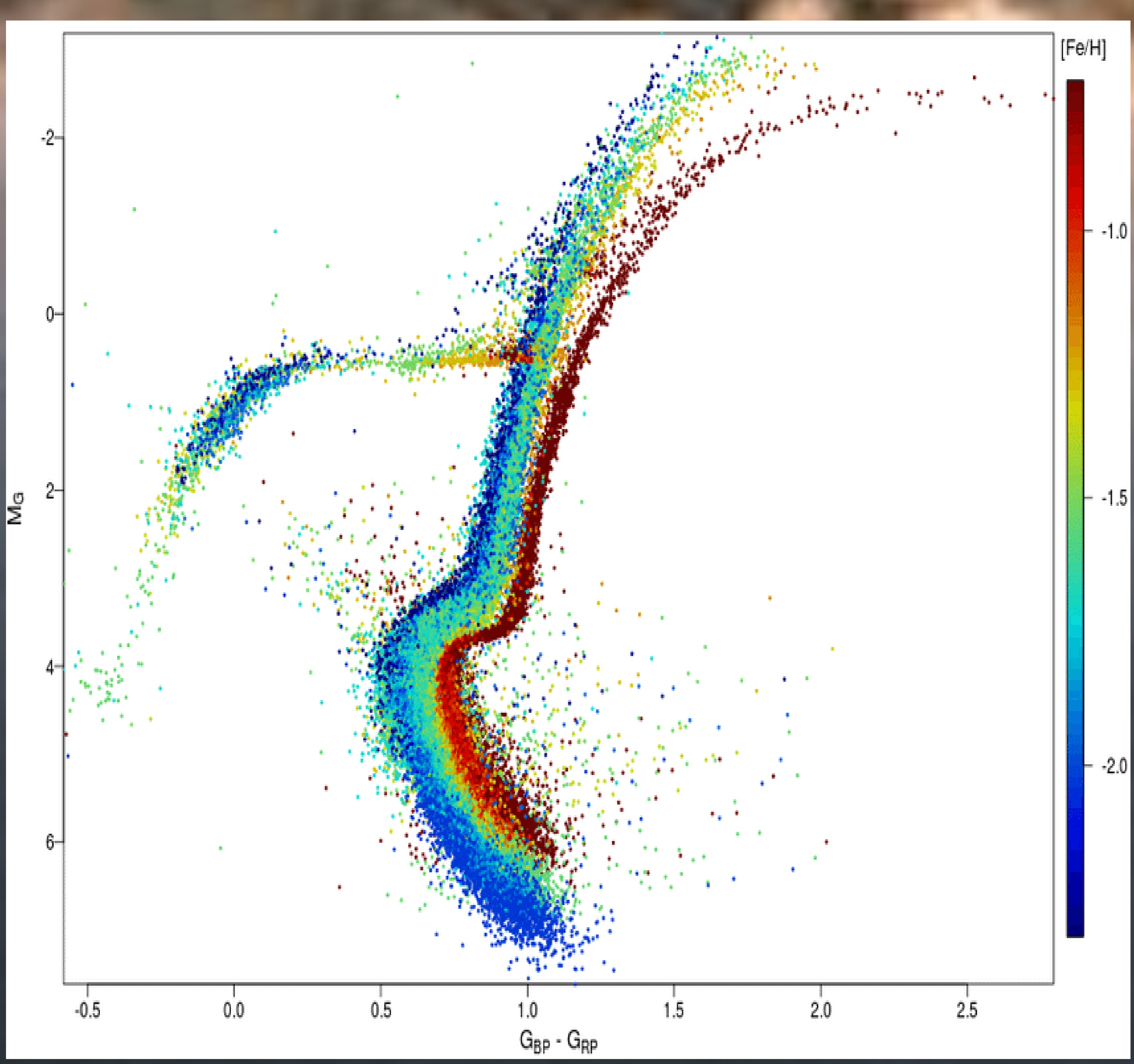


Fig1: Composite CMDs taken from[6]

Not perfectly overlapping group

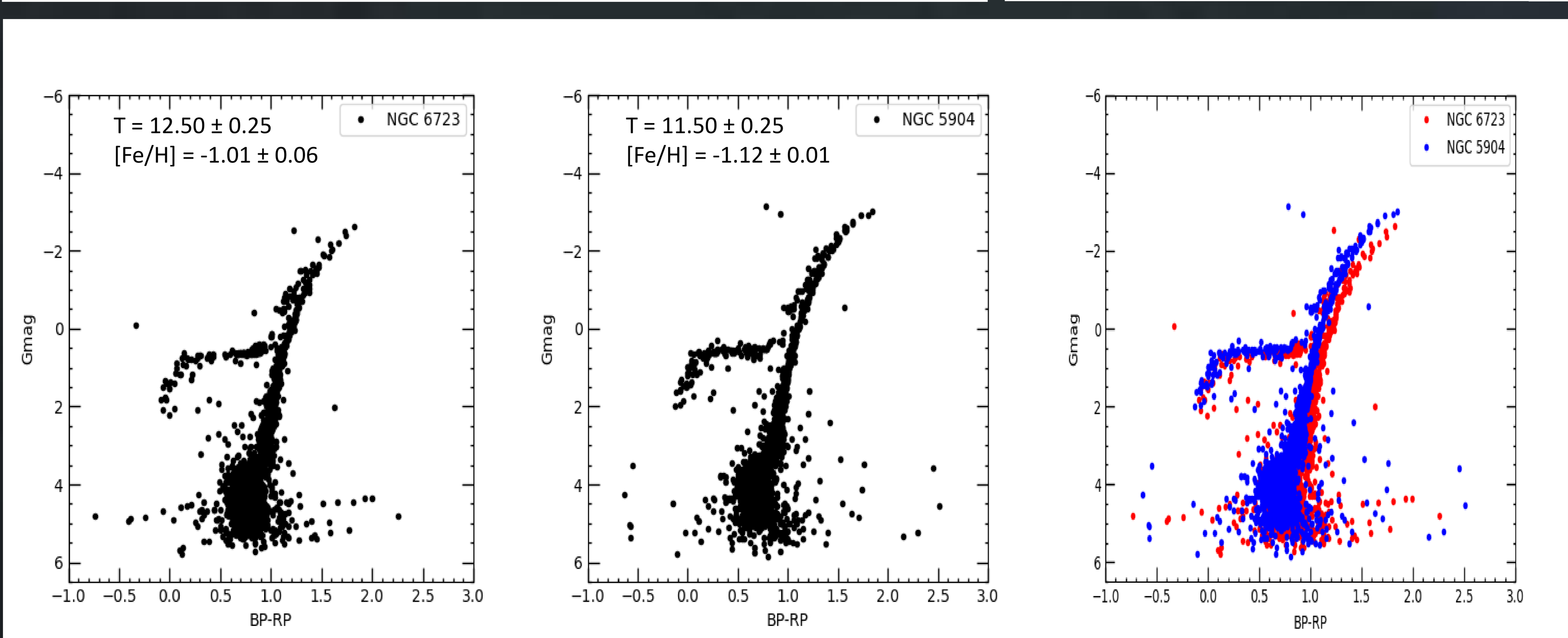


Fig2: NGC 6723, NGC 5904 and combined CMD

Results and discussions: After making different groups of CMDs based on similar morphology, we superimposed CMDs and studied their properties like age and metallicity. GCs (**NGC 6637, NGC 6624**), (**NGC 6541, NGC 6779**), (**NGC 5634, IC 4499**), (**NGC 6352, NGC 6496**), (**NGC 6539, NGC 6760**) are considered as twins.

Perfectly overlapping groups

Singular group

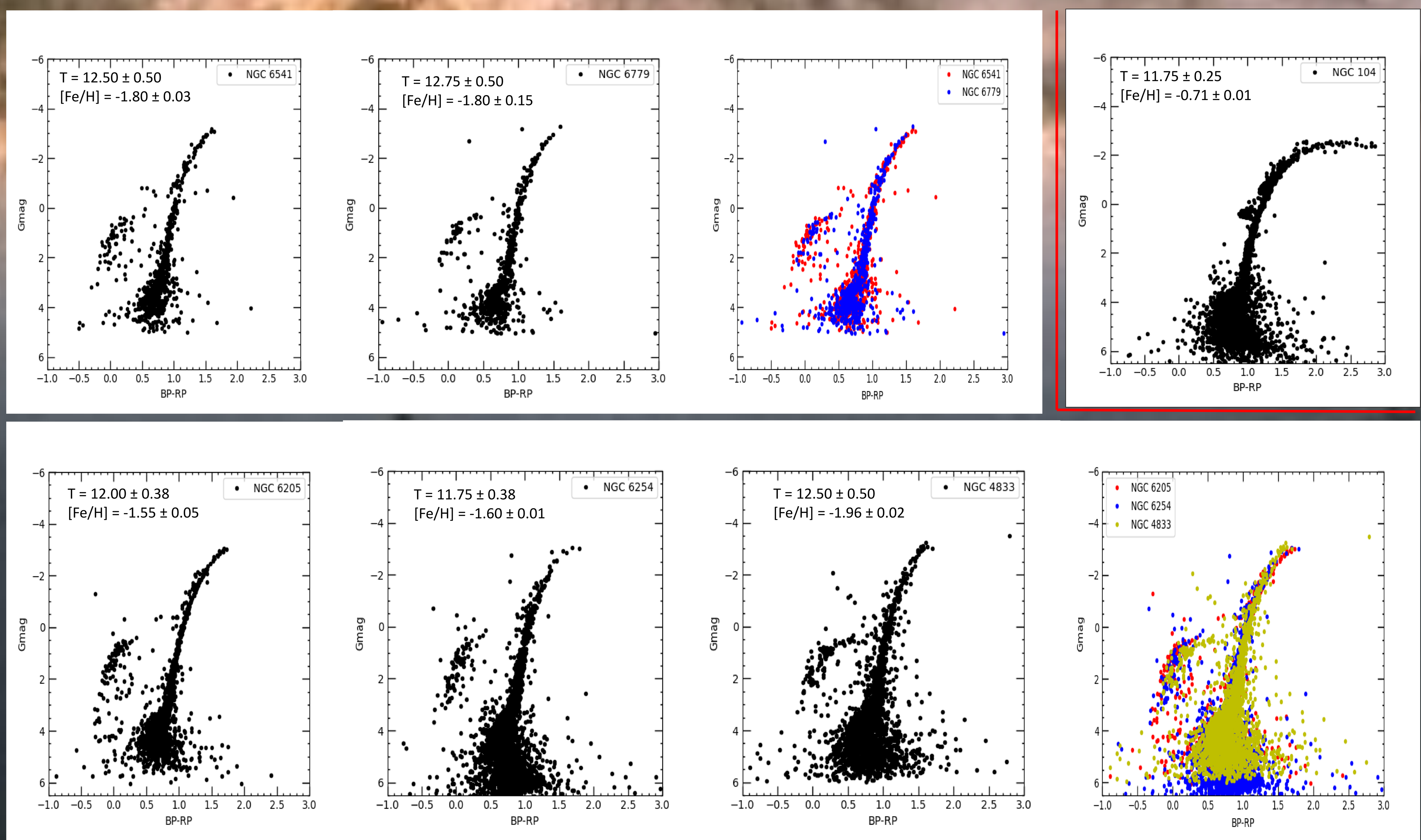


Fig3: Top row: CMD of NGC 6541, NGC 6779 and their combined CMD. **Top right** CMD of NGC 104
Bottom row: CMD of NGC 6205, NGC 6254, NGC 4833 and their combined CMD

Conclusion: We used K-means clustering for image similarity model. We got 27 groups from our image datasets. A group may contain either single or multiple CMDs. After superimposition, we got 5 CMD twins, all from perfectly overlapping groups. Age and metallicities are the main two factors that determine twins.

References:

1.Gaia Collaboration: A. Vallenari, A.G.A. Brown et. al. 2022, A&A

2.https://people.smp.uq.edu.au/HolgerBaumgardt/globular/parameter.html

3.X. luri, A.G. A Brown, L.M. Sarro et. al. 2018, Astronomy and Astrophysics

4.Gaia Collobaration, T. Antoja, P.J. McMillan, G. Kordopatis et. al. 2021, A&A

5. Hartigan, J.A. & Wong, M.A(1979) A k-means clustering algorithm. JSTOR: Applied Statistics, 28, 100—108.

6. Gaia Collaboration, C. Babusiaux, F Van Leeuwen et al 2018, A&A