New catalogue of intermediate mass Pre-Main Sequence objects in Gaia DR2 using Machine Learning

Miguel Vioque* (University of Leeds/Isdefe, Madrid), René D. Oudmaijer (University of Leeds), Deborah Baines (Quasar, Madrid), Ricardo Pérez-Martínez (Isdefe, Madrid)

*pymvdl@leeds.ac.uk

Introduction

The intermediate mass Herbig Ae/Be stars are young stars approaching the Main Sequence. They are key to understanding the differences in formation mechanisms between magnetic low mass stars and the non-magnetic high mass stars. Using Gaia, in combination with other catalogues, we studied the characteristics of the **infrared excess**, the **photometric variability** and the $H\alpha$ **emission** line typical of these Pre-Main Sequence objects (PMS, Vioque et al. 2018); resulting in the largest homogeneous analysis of these sources to date. We now present our programme to discover new Herbig Ae/Be stars from Gaia DR2 using those results by means of a neural network, which is trained with the known objects of the class using their main distinctive characteristics.

Features/Observables

We applied Principal Component Analysis to a set of distance independent observables to select 13 features (combinations of these) to feed the network. These observables are:

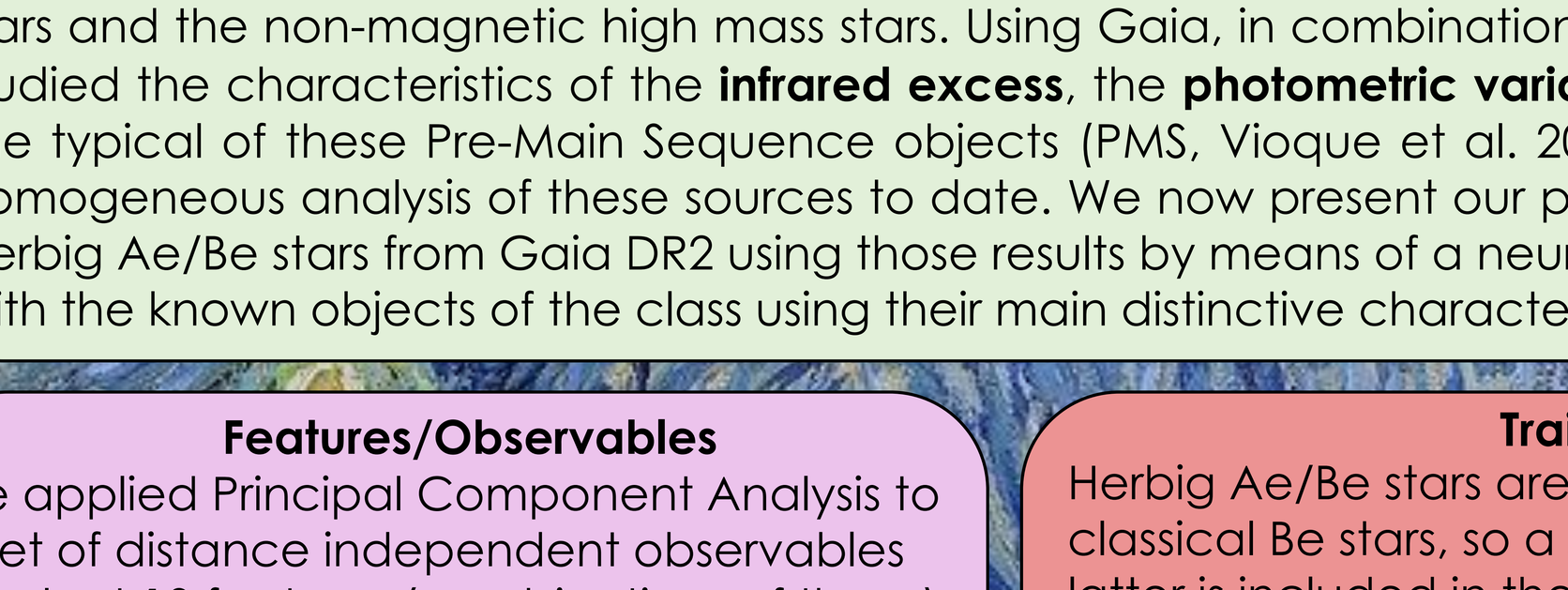
- From **Gaia DR2**: All Gaia colours and 2 variability indicators
- From **AllWISE:** All AllWISE colours
- From **IPHAS** & **VPHAS+**: $r-i,\ r-H_\alpha,\ H_\alpha-i$

Training Set

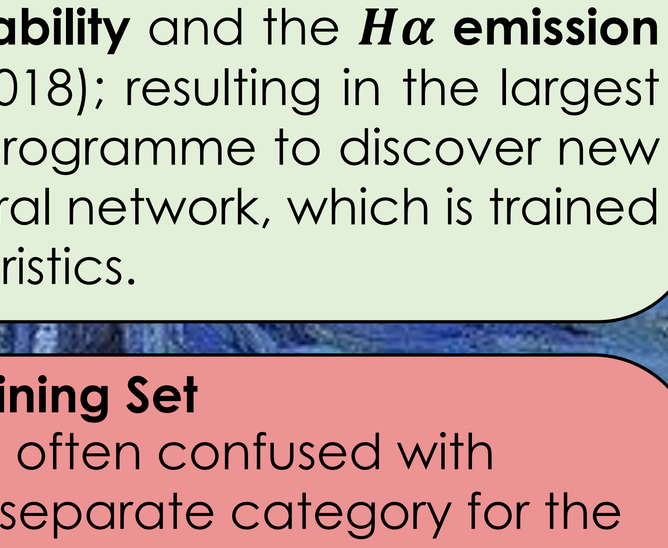
Herbig Ae/Be stars are often confused with classical Be stars, so a separate category for the latter is included in the training set:

- 847 Pre-Main Sequence objects (all available)
- 775 Classical Be stars (all available)
- A subset, large enough to be representative, of other objects present in the used catalogues. This subset is inevitably contaminated by undiscovered PMS objects.

Preliminary results

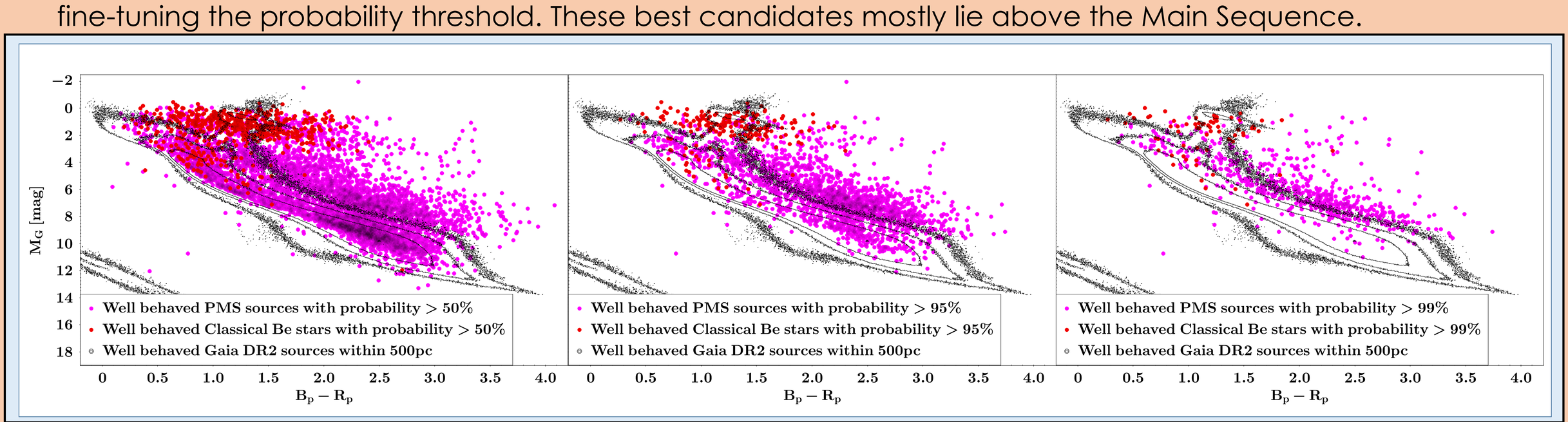


Sky Location



Rosette Nebula

- The algorithm can only be applied to sources with all the characteristics used to train it (~4 million)
- Evaluation on test set gives a recall of 90%.
- We retrieve 15,914 new Pre-Main Sequence objects, 5,768 with more than 95% certainty.
- We retrieve 1,534 new Classical Be stars, 497 with more than 95% certainty.
- New PMS objects tend to cluster, often tracing nebulosities. Classical Be stars are mostly uniformly distributed.
- The HR diagram, which was not used for training, can be used to further select the strongest PMS candidates by fine-tuning the probability threshold. These best candidates mostly lie above the Main Sequence.



References

- Vioque, M., Oudmaijer, R. D., Baines, D., Mendigutía, I., & Pérez-Martínez, R. 2018, A&A, 620, 128

Acknowledgements

This project has received funding from the European Union's Horizon 2020 research and innovation programme under MSCA ITN-EID grant agreement No 676036.