New Herbig Ae/Be and Classical Be candidates using Gaia and Machine Learning

Miguel Vioque
University of Leeds

R. D. Oudmaijer (University of Leeds, UK), M. Schreiner (Desupervised, Denmark), D. Baines (ESAC, Spain), and R. Pérez-Martínez (Isdefe, Spain)
Star Formation in a nutshell

\[ \lambda \xrightarrow[\mu m]{} \text{Pre-Main Sequence Object} \]
Star Formation in a nutshell

van der Marel, et al. (2019)
However, high mass Pre-Main Sequence accretion is still largely not understood.
However, high mass Pre-Main Sequence accretion is still largely not understood
Star Formation in a nutshell

Herbig Ae/Be stars
Intermediate-mass T-Tauri stars
Spectral types ~F5 to B0
Mass: $2 - 10 M_{\odot}$

Star Formation in a nutshell

Star Formation in a nutshell

Around 260 known to date

Herbig Ae/Be stars
Spectral types ~F5 to B0
Mass: $2 - 10 M_\odot$

Intermediate-mass T-Tauri stars

T-Tauri stars

Star Formation in a nutshell
Star Formation in a nutshell

All known Herbig Ae/Be stars

Star Formation in a nutshell

All known Herbig Ae/Be stars

Star Formation in a nutshell

Break in accretion properties between Herbig Ae and Herbig Be

Wichittanakom, et al. in prep

Some other open questions

Clustering

Testi et al. (1999)

Do all massive stars form in clusters?
Some other open questions

Clustering

Do all massive stars form in clusters?

Testi et al. (1999)

Planet Formation

Planet forming or planet hosting disks - when do we start seeing signatures of planet? Class 1? Class 0?

Pinte et al. (2019)
Some other open questions

Clustering

Do all massive stars form in clusters?

Testi et al. (1999)

Disk Structure and Evolution

• Is the spiral structure caused by planets or some other effect?

Garufi et al. (2018)

Planet Formation

Planet forming or planet hosting disks - when do we start seeing signatures of planet? Class 1? Class 0?

Pinte et al. (2019)

• Why do disks around Herbig Ae/Be stars dissipate faster but some seem to persist for very long times?
Looking for new **Pre-Main Sequence (PMS)** objects in Gaia!

Main characteristics of PMS objects:
- Infrared excesses
- $H\alpha$ emission
- Photometric variability
Looking for new Pre-Main Sequence (PMS) objects in Gaia!

Main characteristics of PMS objects:
- Infrared excesses
- Hα emission
- Photometric variability

“Easy” to do for low-mass objects:
*eg., Ksoll et al. (2018), Marton et al. (2019)*

*Cánovas et al. (2019)*
Main characteristics of PMS objects:
- Infrared excesses
- Hα emission
- Photometric variability

High mass PMS objects (Herbig Be stars) are very similar to Classical Be stars

... and supergiants, B[e] stars, ...

Looking for new Pre-Main Sequence (PMS) objects in Gaia!
Looking for new Pre-Main Sequence (PMS) objects in Gaia!

Perform an homogeneous selection, distance and position independent!

High mass PMS objects (Herbig Be stars) are very similar to Classical Be stars

... and supergiants, B[e] stars, ...
Neural Network

1st Step: The algorithm learns from labeled data (Training Set) through minimizing error in successive iterations.
Neural Network

2nd Step: The algorithm improves itself by finding non-trivial patterns hidden within the data

Credit: 3Blue1Brown
3rd Step: Trained algorithm can be applied to data it hasn’t seen before (and hopefully it is good at classifying it)
Before training:
- Training Set
- Set of characteristics
- Set of categories

Algorithm is trained with known labeled data

After generalizing:
- Each category gets a probability
- Efficiency of the algorithm

The best architecture is selected
Neural Network

Selection of the characteristics:
Neural Network

Selection of the **characteristics:**

- Infrared excesses
Selection of the **characteristics:**

- Infrared excesses

**AllWISE (WISE+2MASS)**

\[
W_1, W_2, W_3, W_4 \quad \leftrightarrow \quad J, H, K_s
\]
Neural Network

Selection of the characteristics:

- Infrared excesses
- Hα emission

AllWISE (WISE+2MASS)

\[ W_1, W_2, W_3, W_4 \]

\[ J, H, K_s \]
Neural Network

Selection of the characteristics:

- Infrared excesses:
  - AllWISE (WISE+2MASS) $W_1, W_2, W_3, W_4$
  - IPHAS
  - VPHAS+

- $H\alpha$ emission
  - $J, H, K_s$
  - $r - H\alpha$
Neural Network

Selection of the characteristics:

- Infrared excesses
- $W_1, W_2, W_3, W_4$
- $J, H, K_s$
- $r - H_\alpha$

- Photometric variability

AllWISE (WISE+2MASS)

IPHAS

VPHAS+

H$\alpha$ emission
Selection of the characteristics:

- Infrared excesses
- Hα emission
- Photometric variability

AllWISE (WISE+2MASS)

\[ W_1, W_2, W_3, W_4 \]

\[ J, H, K_s \]

IPHAS

VPHAS+

\[ r - H_\alpha \]

Gaia

2 variability indicators

\[ B_p, G, R_p \]
Selection of the characteristics:

Create **all possible colours**

*Distance and position independent!*

Remove all linear dependency (PCA)

12 dimensions (**characteristics**)
Selection of the **categories:**
Selection of the categories:

PMS category
Selection of the categories:

PMS category

Classical Be category
Selection of the **categories**:

- PMS category
- Classical Be category
- *Other* sources
Selection of the **Training Set:**

- PMS category
- Classical Be category
- *Other* sources
Selection of the **Training Set**: 

- AllWISE
- IPHAS
- VPHAS+
- Gaia

\[ \text{AllWISE} + \text{IPHAS} + \text{VPHAS+} + \text{Gaia} = 4,151,538 \text{ sources} \]

- **PMS category**
- **Classical Be category**
- **Other sources**
Selection of the **Training Set:**

**PMS category**
- 848 Pre-Main Sequence objects (163 Herbig Ae/Be)
- 775 Classical Be stars
- 471,111 random sources

**Classical Be category**

**Other sources**

\[ 4,151,538 \text{ sources} \]
Selection of the **Training Set**:

- **AllWISE**
- **IPHAS**
- **VPHAS+**
- **Gaia**

\[ \text{4,151,538 sources} \]

- **848** Pre-Main Sequence objects (163 Herbig Ae/Be)
- **775** Classical Be stars
- **471,111** random sources
Training the Neural Network
Trained Neural Network

\[ \text{AllWISE} + \text{IPHAS} + \text{VPHAS+} + \text{Gaia} = 4,151,538 \text{ sources} \]
Probability Map

- Other objects candidates
- PMS candidates
- Classical Be candidates
- Either PMS or Classical Be candidates
636 Classical Be candidates

Evaluation on Test Set

PMS Completeness 78.8 ± 1.4%

Classical Be Completeness 85.5 ± 1.2%

8452 PMS candidates

4,140,629 other
Selection of the **Training Set:**

\[
\text{AllWISE} + \text{IPHAS} + \text{VPHAS+} + \text{Gaia} = 4,151,538 \text{ sources}
\]

- **PMS category**
  - 848 Pre-Main Sequence objects (163 Herbig Ae/Be)
- **Classical Be category**
  - 775 Classical Be stars
- **Other sources**
  - 471,111 random sources
Selection of the **Training Set**:

- **AllWISE**
- **IPHAS**
- **VPHAS+**
- **Gaia**

Total **4,151,538** sources

- **848** Pre-Main Sequence objects (163 Herbig Ae/Be)
- **775** Classical Be stars
- **471,111** random sources

There is a large contamination between categories!
This algorithm cannot assess itself, we need a **totally independent** analysis.

### Sources

- **PMS category**
  - 848 Pre-main sequence objects (163 Herbig Ae/Be)

- **Classical Be category**
  - 775 Classical Be stars

- **Other sources**
  - 471,111 random sources

There is a large contamination between categories!
Gaia HR diagram

PMS candidates

Training Set

- $B_p - R_p$
- $M_g$ [mag]
- Probability PMS

- T Tauri
- Classical Be
- Herbig Ae/Be
Gaia HR diagram

Classical Be candidates

Training Set
3131 potential high mass (682 with good Gaia solution)
Coordinates
Coordinates
Coordinates
Physical behaviour

Variability vs. Masses (lower limits)

Candidates

Training Set
Variability vs. Masses (lower limits)

Candidates

Vioque et al. 2018
Caveats

Hα emission

Main Sequence

Reddening

Candidates

PMS candidates
Caveats

Candidates

Hα emission

Main Sequence

Reddening
Caveats

Candidates

Planetary Nebula!
Future work

Past and future observations
- INT
- 2.2m Calar Alto
- NTT
- MEAVE
- GaiaESO

Populate HR diagram
- 260 objects
- \( \sim 3000 \) objects
Results

• We retrieve 8452 new PMS candidates. 3131 (682) potential Herbig Ae/Be stars. Completeness 78.8 ± 1.4%

• We retrieve 636 new Classical Be stars candidates. Completeness 85.5 ± 1.2%

• We retrieve 1266 candidates of belonging to either one of the two categories.