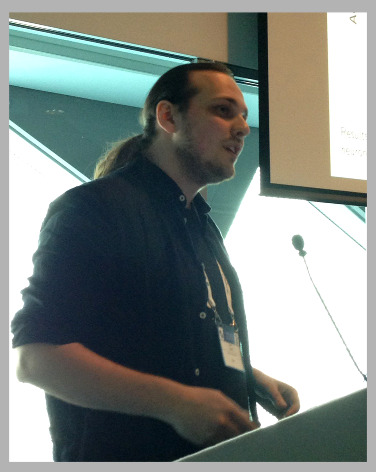


Selection of Spitzer Young Stellar Object candidates using Deep Learning classifiers

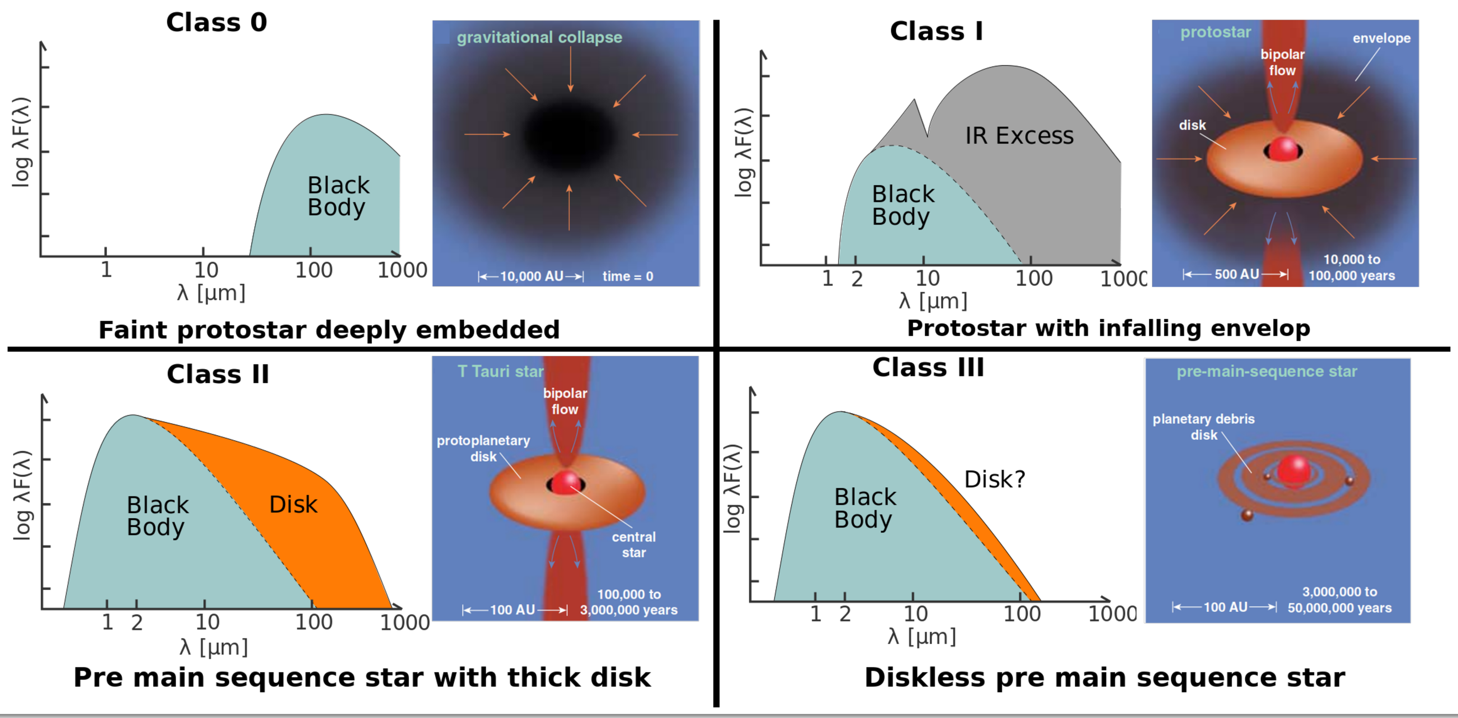
David Cornu and Julien Montillaud

UTINAM, Univ. Bourgogne Franche-Comté, Besançon, France

linkedin page

1 - Introduction

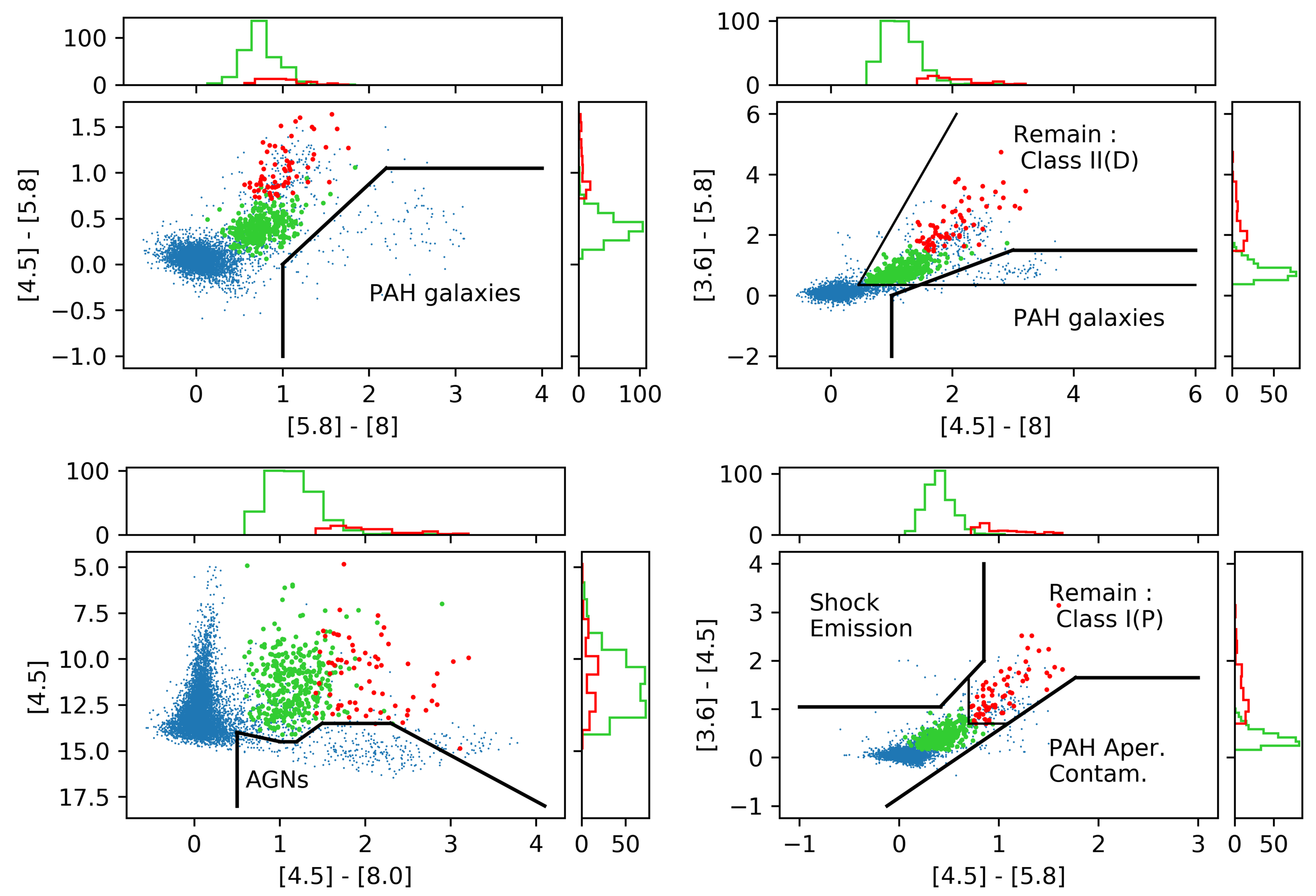
Observing **Young Stellar Objects** (YSOs) enables us to characterize the star forming regions, e.g. by their star formation efficiency and history or by their density distribution. They are classified into evo lutionary stages (class 0, I, II) using their infrared (IR) Spectral Energy Distribution (SED). **The Spitzer telescope** can be used for such classification using his IRAC bands at 3.6, 4.5, 5.6, 8 μm and MIPS 24 μm, allowing us to distinguish IR excess in class I or the disk emission of the class II from other kind of objects/contaminents.



2 - Usual Method and Limitations

The most used classification scheme is described by Gutermuth+ 2009 [1] and perform straight cuts in color-color (C-C) and color-magnitude (C-M) diagrams. This method extracts contaminents (PAH, AGN, Shock, more evolved field stars ...) in multiple phases, and retrieves class II and class I objects. It performs well on big and nearby star forming regions, but suffers from overlap due to the intrinsic limitations of straight cut methods.

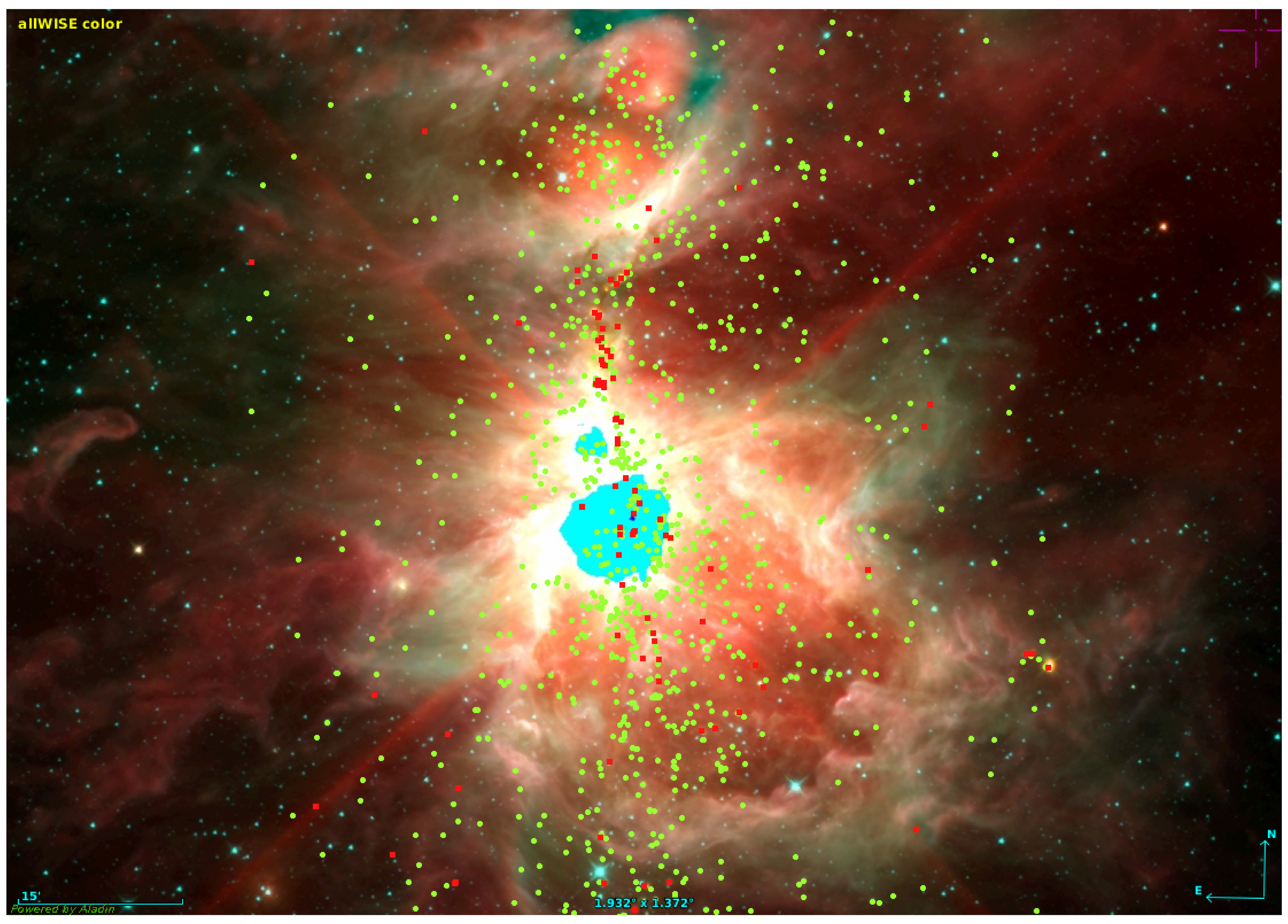
However, since Spitzer-like surveys are getting more accurate and cover large sky areas, we have access to large and diverse datasets. **Machine learning** appears then as one of the best solutions by provinding adaptive non-linear and statistically learned classification in any number of dimensions (example with SVM in Marton+ 2016 [2] and in-prep Marton+ 2019).



Usual C-C & C-M diagrams used for this classification. Retrieved class II, class I candidates and other/field stards

5 - Conclusion

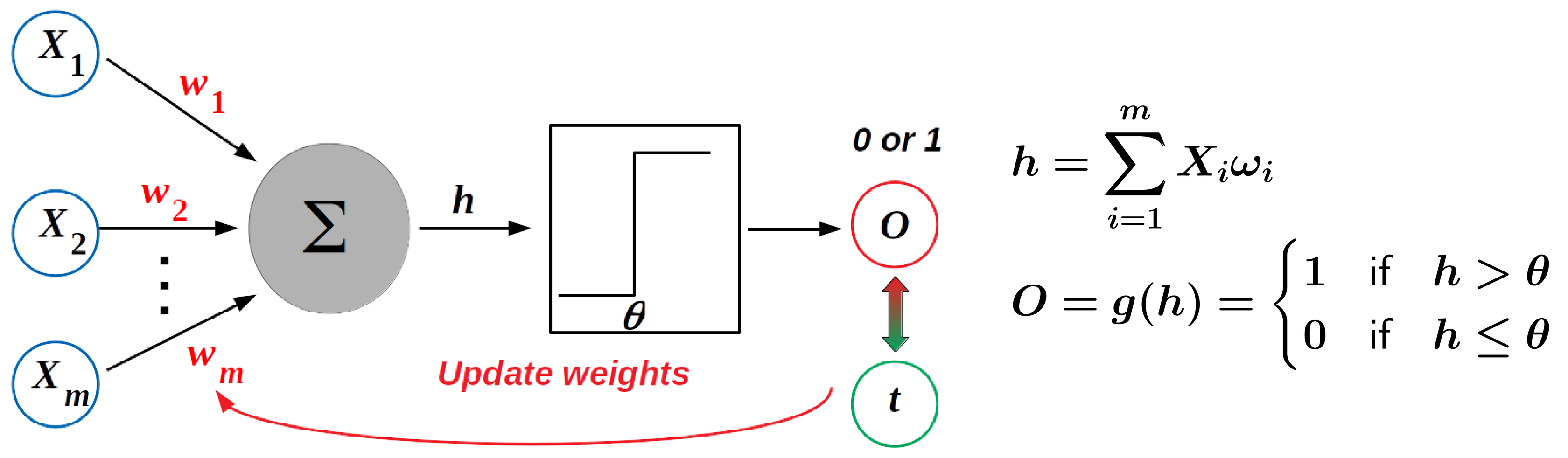
With YSO classification we can recover **structural information** about the star-forming regions since this is the place where they are born. It gives us information about their activity and evolution. With the **GAIA DR2** mission we can measure accurately the distance and motion of these stars, allowing us to recover the 3D structure and dynamics of such regions. We have an example of this application in Großschedl+ 2018 [5] where they identified two different component in Orion A using the Megeath YSO candidates catalog and GAIA DR2.



Orion cloud with WISE background (RGB : 22, 4.6, 3.4 μm). Overplot retrieved class II and class I candidates

3 - Deep Learning

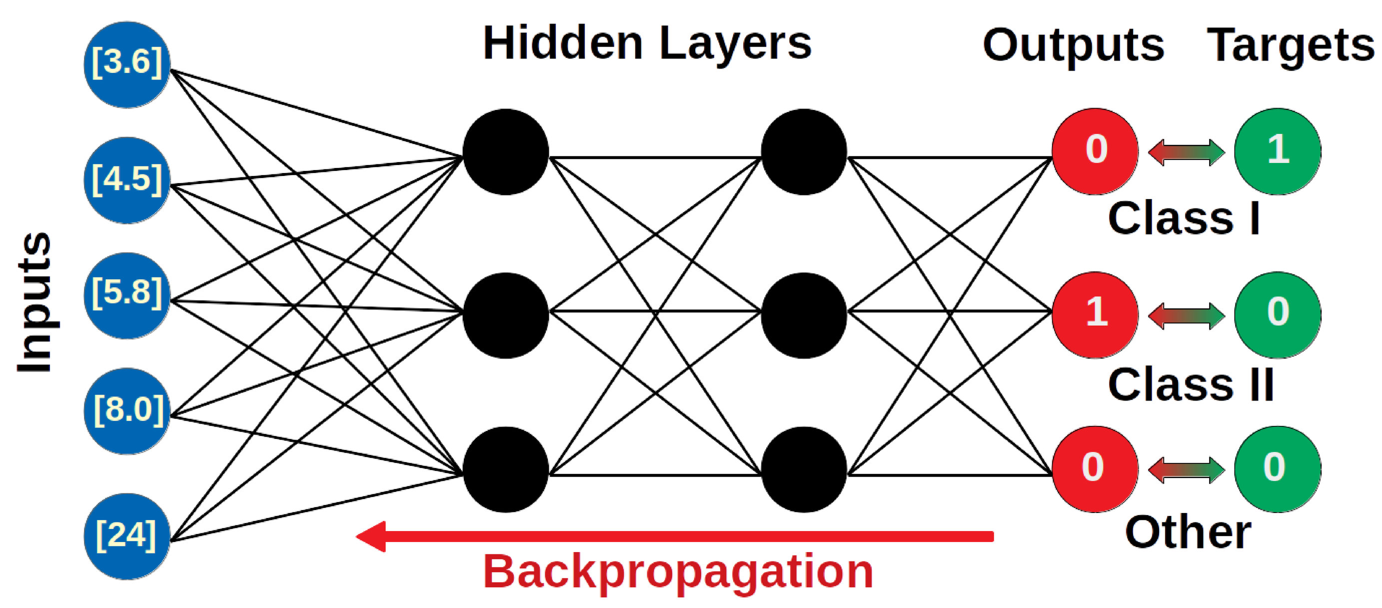
In our case, we rely on one of the most famous and widely used supervised-learning approach : **Artifical Neural Networks** (ANN). The neurons in ANN are defined as a mathematical model performing weighted sums of an input vector (which represents different dimensions of the same object to classify). It then compares the result to a threshold activation function and changes the state of the neuron to 1 or 0 (firing or not). The learning process consist in comparing the expected value (target) with the result obtained by the activation of the neuron, and then modify the weights according to the error gradient, hence learning anti/correlations.



$$h = \sum_{i=1}^{m} X_i \omega_i$$

$$O = g(h) = \begin{cases} 1 & \text{if} \quad h > \theta \\ 0 & \text{if} \quad h \leq \theta \end{cases}$$

Performing complex classification requires more than one neuron. We can add them in one layer, fully connecting each neuron to all the inputs, and we can add multiple layers by making some neurons take the output of the previous layer as their own inputs. This is the **Multi Layer Perceptron** (MLP) which is a universal function approximator, able to perform classification, regression, compression, time series prediction, ...



$$\frac{\delta E}{\omega_{\zeta\kappa}} = \frac{\delta E}{\delta h_\kappa} \frac{\delta h_\kappa}{\delta w_{\zeta\kappa}} \qquad \delta_\kappa \equiv \frac{\delta E}{\delta h_\kappa} = h'(o_\kappa) \sum_\zeta \omega_{\zeta\kappa} \delta_\zeta$$

Since one has the expected answer for the ouput layer, it is possible to propagate the error through the network, using a **gradient descent scheme**. Then one can decides how to encode the output by having one neuron for each class we want to recover (class I, class II, Other/Contaminants) and tell the network that the result should be a 1 for the expected class and a 0 for the others (using the so-called SOFT-MAX activation function).

4 - Training and Results

Since ANN are supervised methods, a **training set** has to be defined. We have built our sample with the 1 kpc young stellar cluster survey from Gutermuth+ 2009 [1], the Orion catalog from Megeath+ 2012[3], and Mon Ob1/NGC 2264 from Rapson+ 2014 [4], for a total of $\approx$ 28000 objects. We then performed a modified Gutermuth classification to define the targets and train the MLP with 2/3 of the data using 20 hidden neurons. We developed our own **neural network framework** coded in C and using CUDA GPU computation. It allow us to tune and fully control the network behavior, with fast training (convergence in $\approx$ 2-3 minutes on a GTX 780 for $\approx$ 30 000 objects seen 1M times each).
The results are in the form of a **confusion matrix**.

Predicted

Actual

IN \ OUT	YSO CI	YSO CII	Other	Recall
YSO CI	749	32	16	93.98%
YSO CII	25	3957	28	98.68%
Other	13	127	24127	99.42%
Precision	95.17%	96.14%	99.82%	

Each row corresponds to the objects of a target class and each cell tell where the network has classified them, which defines the recall (recovery rate).

In contrast, each column defines what the algorithm gave and each cell shows the original class of the training set, which defines the precision.

These results are competitive with other similar studies like Marton+ 2016[2] and his upcomming one (compared the preliminary results) with a way better purity on observational proportions. We are currently working on the application of this training on large survey catalog like **GLIMPSE** which give encouraging results so far. We are also preparing the publication in Astronomy & Astrophysics for the next month. Next to this we plan to improve our framework with more recent **semi-supervised ANN** (like Deep Belief Network) and add convolutional layers for image analysis.

References

[1] Gutermuth et al. *A Spitzer Survey of Young Stellar Clusters Within One kpc of the Sun.* ApJS, 2009.

[2] Marton et al. *An all-sky support vector machine selection of WISE YSO.* MNRAS, 2016.

[3] Megeath et al. *The Spitzer Space Telescope Survey of the Orion A and B Molecular Clouds.* ApJ, 2012.

[4] Rapson et al. *A Spitzer View of the Giant Molecular Cloud MON OB1 EAST/NGC 2264.* ApJ, 2014.

[5] Grossschedl et al. *3D shape of Orion A from Gaia DR2.* A&A, 2018.

This thesis work is founded by the CNES and the "Région Bourgogne Franche Comté". This results have been presented in a talk at EWASS 2018 Liverpool/UK and one at the GCC 2018 meeting Besancon/FR. One can find the visual support for this presentation on the "Astrophysics Source Code Library" blog at ascl.net. Or the last up to date slides can be download from the GCC webpage.