Robust detection of astronomical sources using convolutional neural networks

M. Paillassa[1] *, E. Bertin[2] *, H. Bouy[1]

[1]Université de Bordeaux; [2]Institut d'Astrophysique de Paris
maxime.paillassa@u-bordeaux.fr, herve.bouy@u-bordeaux.fr, bertin@iap.fr
github.com/mpaillassa/MaxiMask

Introduction

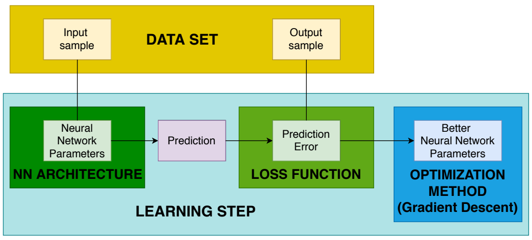
Extracting reliable source catalogues from images is a crucial step for further analysis in astronomical surveys.

A widely used method to perform source detection is SExtractor [3]: after having background subtracted the image, it consists of matched-filtering it with an appropriate filter (usually the PSF of the image) to enhance places where sources are. Then a threshold is applied to get the source pixels and a deblending procedure based on mathematical morphology is eventually run.

The two main limitations of this method are the following:

- Filtering, thresholding and deblending are heuristic-based and need parameter tuning.
- It lacks of robustness regarding contaminants.

This is why we propose to design new tools to detect sources in astronomical images. In order to avoid these limitations, we propose to address the problem using machine learning techniques, especially supervised learning with convolutional neural networks (CNNs) [5]. These methods enable to design models that can learn to map inputs to outputs using a data base of examples and have proven to be very efficient in various computer vision tasks such as image classification (assigning label(s) to images) [6] or image segmentation (assigning label(s) to pixels) [1].



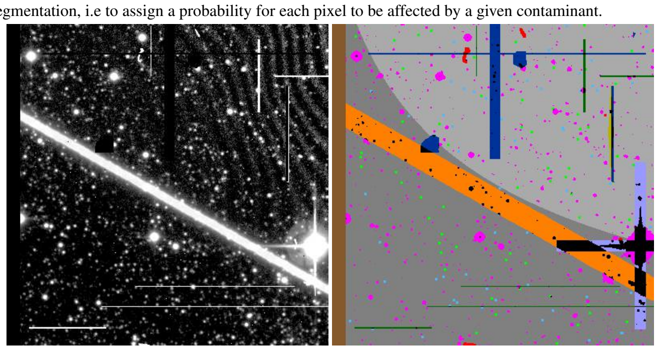
Scheme showing the principles of supervised learning. The learning procedure consists of iterating learning steps: at each learning step, the neural network makes a prediction from an input using its parameters. As the ground truth prediction is known, a loss function can be designed to quantify the prediction error of the neural network. Using an optimization method, all the parameters of the neural network can be updated to make better predictions in the future.

In addition to be promising methods to solve our source detection limitations, CNNs are also of interest with the view to design a universal detector which could encompass a wide range of conditions (optic, electronic detector type, image sampling, seeing, stellar density), providing that this diversity is represented in the training set.

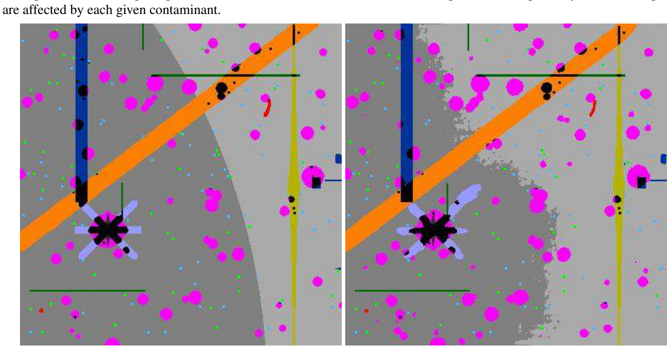
In more details, CNNs are stacks of convolutions using the convolutional kernels as learnable parameters. Doing so, CNNs can be optimized to learn directly from the data to identify the useful features (edges, sharp shapes and more and more complex features) to better map inputs into outputs. In the end the idea is to naturally evolve from an empirically designed linear filter and thresholding (SExtractor) into a non linear complex filter automatically learnt from data (CNNs).

MaxiMask for contaminant identification

As the first limitation (heuristic based processing) is naturally solved by the use of machine learning techniques, we decided to solve the second limitation (robustness regarding contaminants) in a separate process by identifying contaminants before focusing on the astronomical sources. To do so, we built a data base using mainly Cosmic Dance archives [4] to train MaxiMask, a CNN to perform semantic segmentation, i.e to assign a probability for each pixel to be affected by a given contaminant.



Example of built training sample. We added ourselves contaminants in clean images so that we perfectly know which pixels are affected by each given contaminant.



Example of prediction on a test image after training. Left: ground truth. Right: prediction.

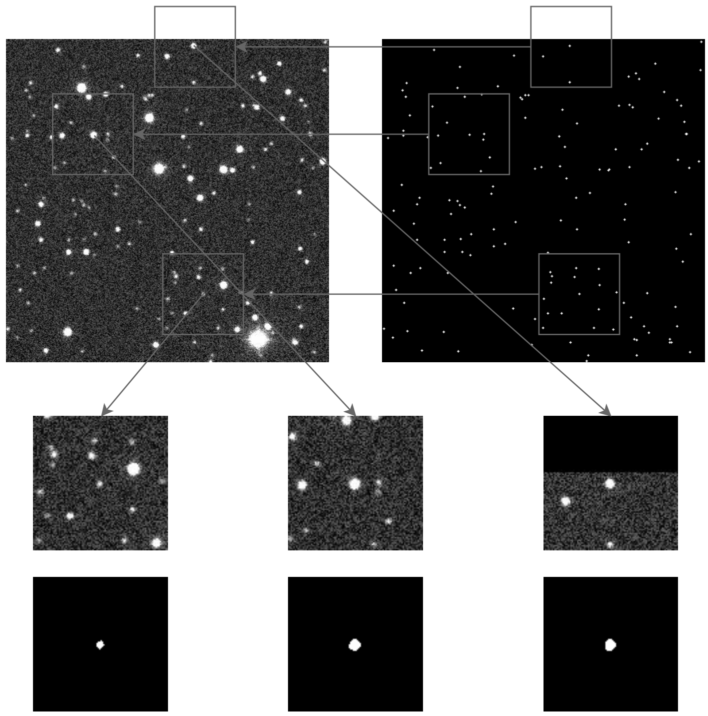
Source detection

Once all the contaminants and artifacts are identified, the source detection itself can be made. The combination of two problems is making the source detection problem particularly challenging:

- The detection must be instance aware, i.e each object must be detected individually.
- The detection must handle blending sources, i.e pixels can belong to several sources.

Having a system that solves both problems in a efficient way is quite difficult; the literature being often limited to mutual exclusive problems (pixels belong to a single object), tackling occlusion (object hiding another object) more than blending (object merged with another object).

Our first approach consists of detecting each source by identifying its centroid. In this framework, each individual star can be identified without ambiguity, even if it is blended with one or several others. In order to test different source detection prototypes, we simulate a simple data base of stellar images using SkyMaker [2].



Example of training sample for source detection. In a first place, a CNN predicts all the source centroids. Then, a stamp centered on each centroid is extracted so that another CNN can segment each corresponding source.

Further work includes making a more complete data set where images could contain contaminants, masked contaminants and galaxies.

References

[1] Vijay Badrinarayanan, Alex Kendall, and Roberto Cipolla. Segnet: A deep convolutional encoder-decoder architecture for image segmentation. *arXiv preprint arXiv:1511.00561*, 2015.

[2] E. Bertin. SkyMaker: astronomical image simulations made easy. , 80:422, 2009.

[3] Emmanuel Bertin and Stephane Arnouts. Sextractor: Software for source extraction. *Astronomy and Astrophysics Supplement Series*, 117(2):393–404, 1996.

[4] H. Bouy, E. Bertin, E. Moraux, J.-C. Cuillandre, J. Bouvier, D. Barrado, E. Solano, and A. Bayo. Dynamical analysis of nearby clusters. Automated astrometry from the ground: precision proper motions over a wide field. , 554:A101, June 2013.

[5] Yann LeCun, Yoshua Bengio, et al. Convolutional networks for images, speech, and time series. *The handbook of brain theory and neural networks*, 3361(10):1995, 1995.

[6] Karen Simonyan and Andrew Zisserman. Very deep convolutional networks for large-scale image recognition. *arXiv preprint arXiv:1409.1556*, 2014.

Acknowledgements

INITIATIVE université de Bordeaux D'EXCELLENCE
erc European Research Council Established by the European Commission
European Commission