Cats vs Dogs, Photons vs Hadrons

Francesco Visconti[1]
for the ASTRI Project[2]

[1]*Osservatorio Astronomico di Roma, INAF*
[2]*http://www.astri.inaf.it/en/library/*

ABSTRACT

In gamma ray astronomy with Cherenkov telescopes, we need machine learning models help us decide what kind of particles we observed, their energies and directions. I focused on the classification task only training a simple convolutional neural network, using as input uncleaned images generated by Montecarlo data for a single ASTRI telescope. Results show an enhanced discriminant power with respect to classical random forest methods.

Network Architecture

A five layers Convolutional Neural Network has been trained with ASTRI calibrated images from a Montecarlo production. Network architecture is a very basic one, not optimized, meant for prototyping.

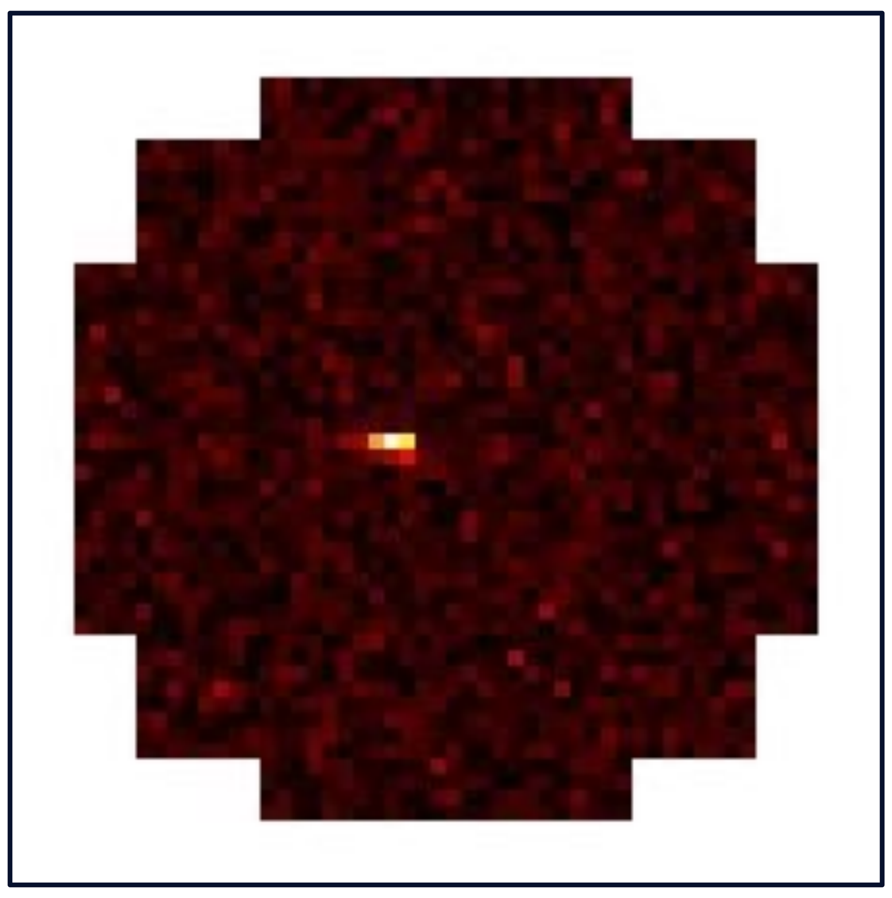

Fig. 1:
ASTRI calibrated image, not cleaned

Input Data

Single telescope images have been generated starting from a Montecarlo ASTRI data production, calibrated via the dedicated software developed in Rome by our group. In order not to introduce bias, I selected the data before the cleaning process, keeping the full background. I used 70k images for each class in the training step, 30k for validation during training, and finally made prediction on 100k files for each class.

ACKNOWLEDGEMENTS

This work was conducted in the context of the ASTRI Project

Results

The training took 1 hour on a K20 NVIDIA GPU, stopped after 11 epochs monitoring the **validation accuracy** improvement. With regard to the performance, the max validation accuracy reached was **96.7%**. For a quick comparison, the standard literature method (Random Forest on Hillas parameters) do not exceed 86%.

The good result is confirmed by the g-h separation plot, where we see an enhanced discrimination power for the CNN network on calibrated data.

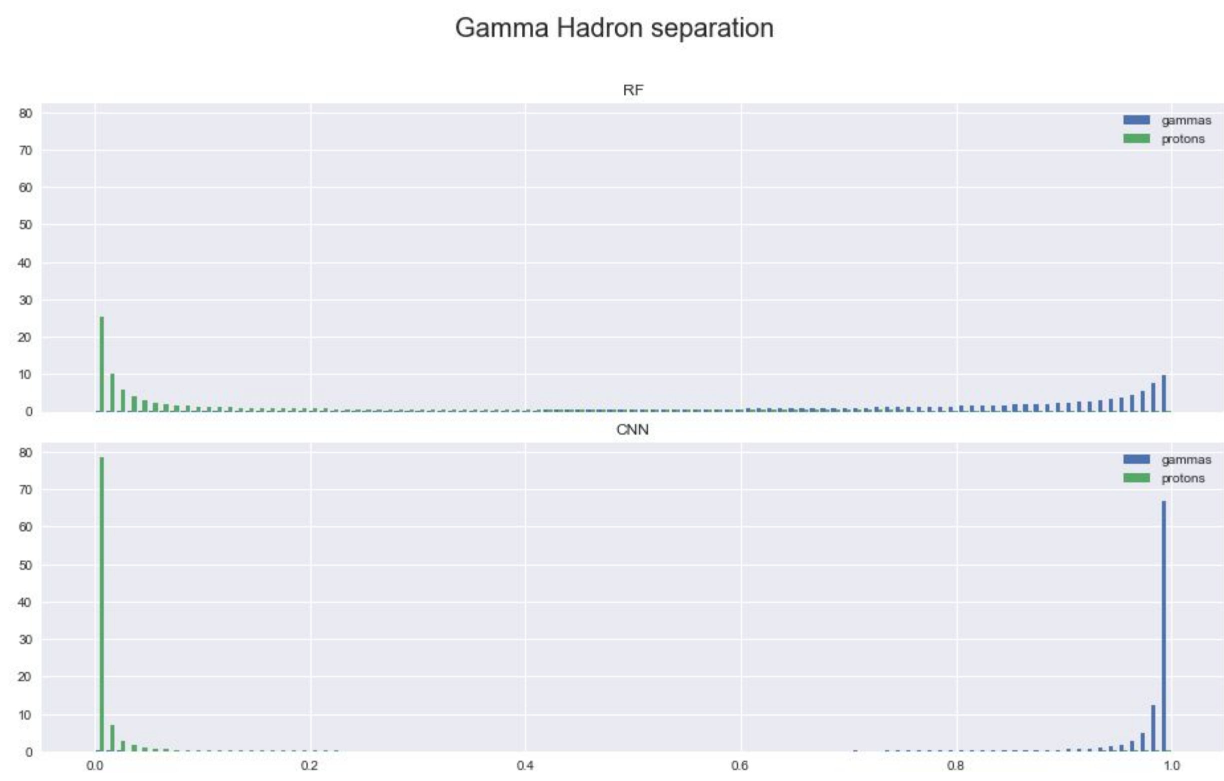

Fig. 2:
g-h separation: comparison with RF

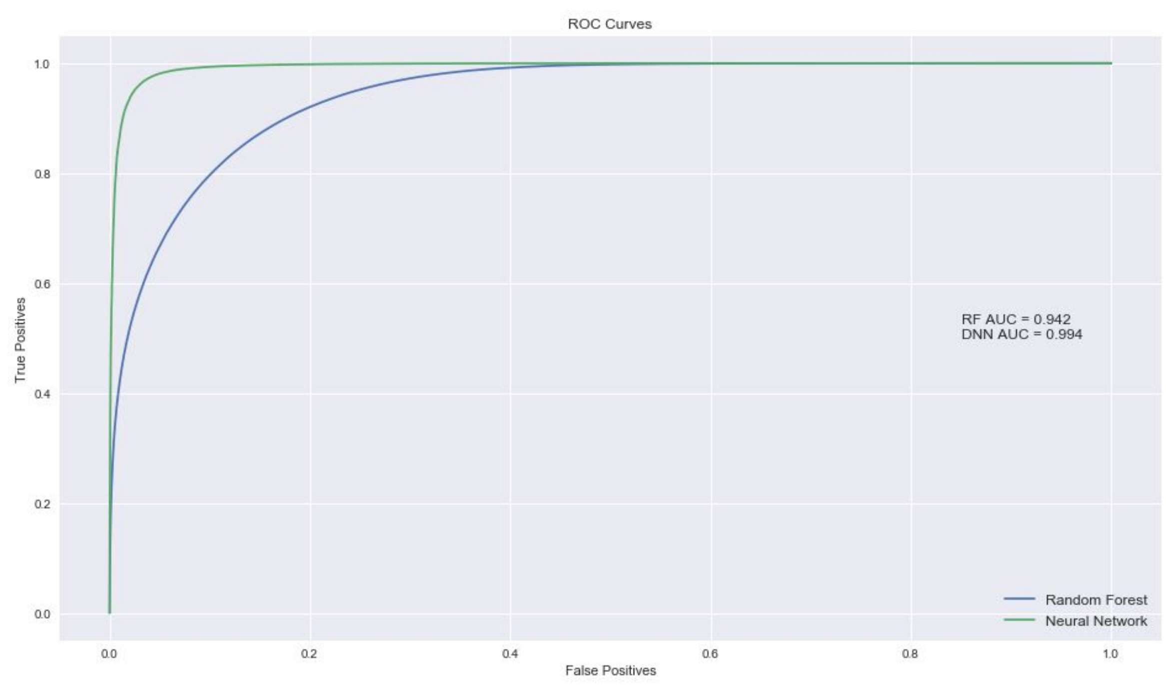

Fig. 3:
ROC curve for CNN and RF

ROC curves feature true positive rate on the Y axis and false positive rate on the X axis. This means that the top left corner of the plot is the "ideal" point. The larger the area under the curve (AUC), the better.
CNN performs significantly better than RF.

Comments

A simple experiment (reproducted recently on a more modern GPU board) produced an interesting result, enhancing the discriminatory power for the gamma hadron separation problem. This encourages in exploring these techniques, suggesting the **bias introduced by event parametrization is very important.**

References:
https://doi.org/10.1016/j.astropartphys.2018.10.003, https://doi.org/10.22323/1.395.0730 PoS