Trigger proposal

Igor Pains
with Giovanni Mazzitelli and Rafael Nóbrega
Joining the group:
Gabrihel Silva and Jordan Venâncio

19/10/2023
Analysis & reconstruction meeting

1.
Introduction

Introduction

- Motivation: reduce data to manageable levels by selecting only events of interest, saving storage and processing resources.
  - Each run may need up to 2 Gb to be stored after the compression.
    - ~1 Tb per day considering the current frequency.

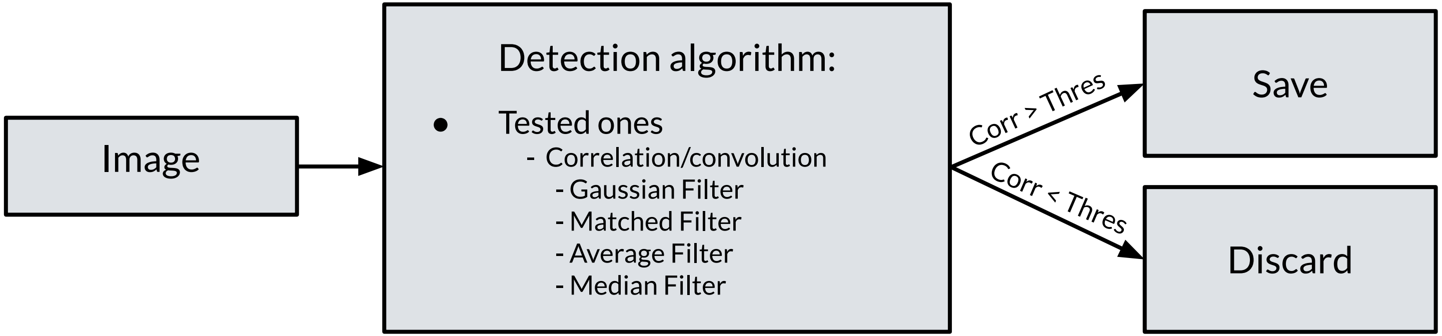
Proposal

- Develop algorithms to be tested as online trigger to decide whether to save or not images taken by the detector.
  - *on going* → Convolution of the image with several kernels: look for high correlation points. Link of the last presentation
  - Explore Machine Learning methods

2.

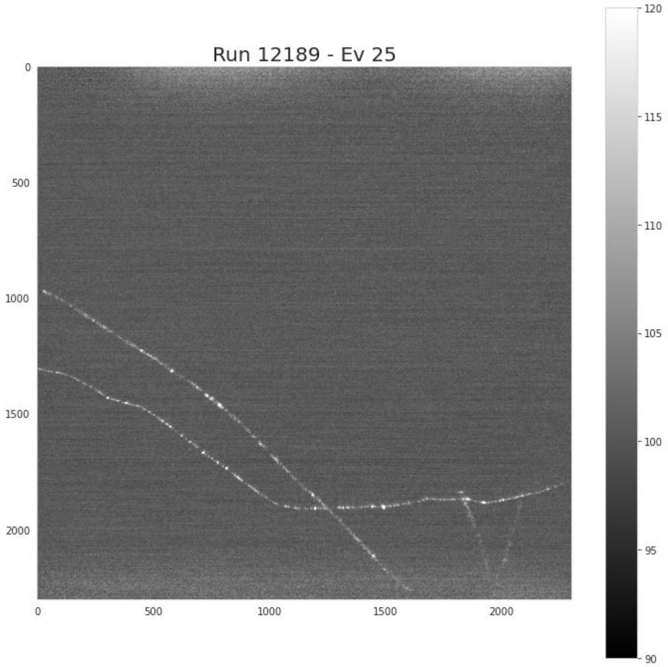
Analysis

Methodology

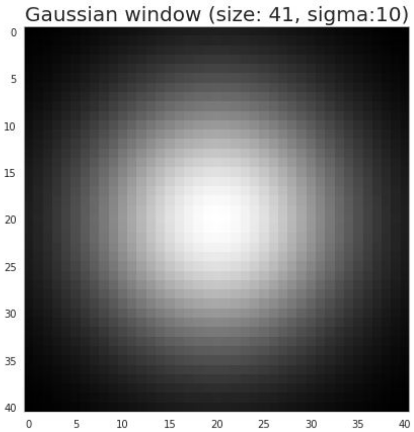


- A large set of parameters was used during the training for each filter (window size and sigma if needed)
- The filters had a slightly better performance using pedestal subtraction method.
- The best filter was the **Gaussian with window size equal to 19 and sigma 5.5**
  - close to what could be achieved by fitting the data with a 2D gaussian function

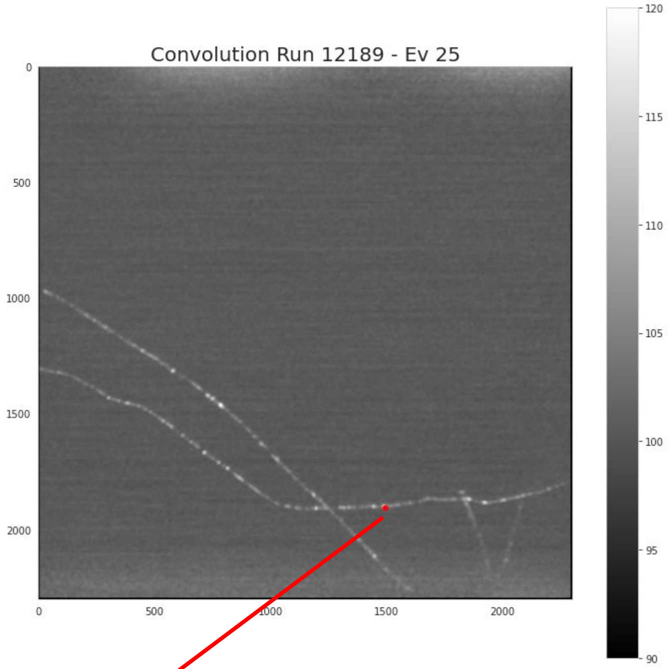
Correlation/Convolution



*



=



Raw image

Normalized for area equal to 1.

Pixel with highest correlation

Datasets

- **Datasets**
  - **Training:**
    - Noise dataset: 300 images from pedestal runs (Run 2 underground).
    - ER signal simulation: 300 images containing 0.5 keV signals added to pedestal runs.
  - **Test (reconstruction was also used for comparison):**
    - Noise dataset: 300 images (different from training dataset)
    - ER signal simulation: 300 images containing 0.5 keV signals added to real pedestal runs.
    - NRAD run: 405 images (run 12189)
    - NR simulation: 219 images containing NR simulated signals added to pedestal runs.

not shown in this presentation

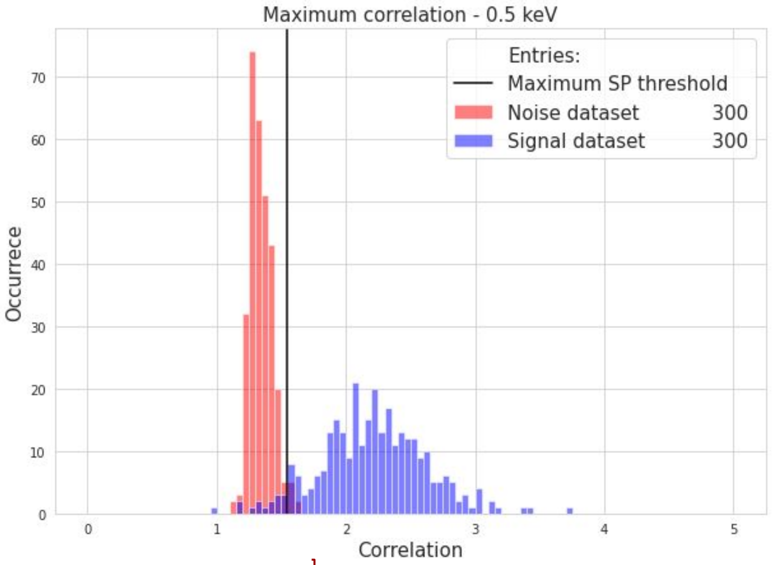
We need low energy Nuclear Recoils

Training *(scanning methods and their parameters)*

- The convolution was performed and the highest correlation was stored for each dataset.

- A threshold that best separates the two datasets was chosen (using the SP metric).

$$SP = \sqrt{\sqrt{DET_{sig} DET_{noise}} \left(\frac{DET_{sig} + DET_{noise}}{2} \right)}$$



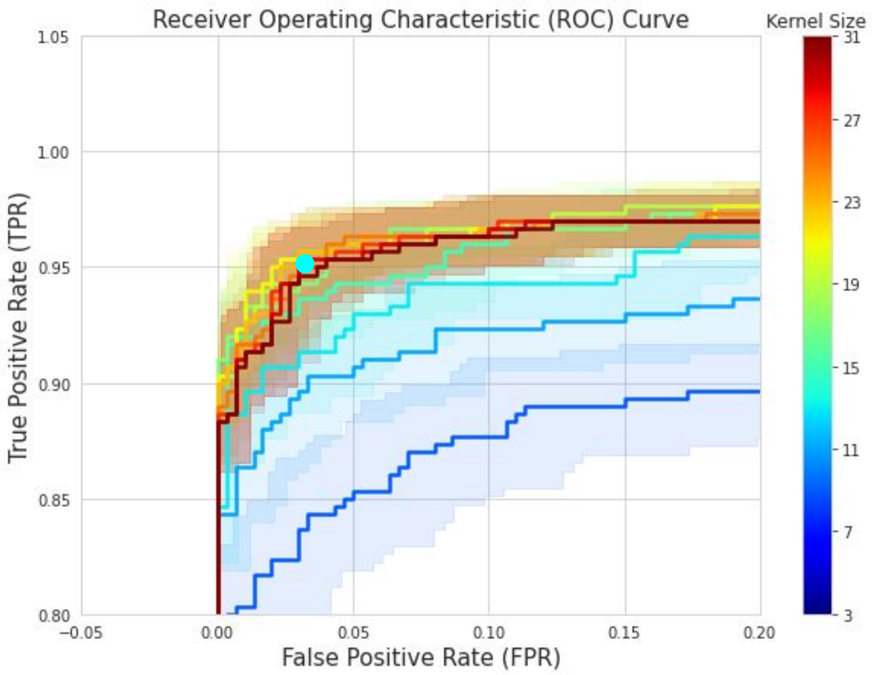
Gaussian with window size equal to 19 and sigma 5.5

DET_{sig}: (95.3±2.0)% (Percentage of signal elements above threshold)

DET_{noise}: (97.7±1.4)% (Percentage of noise elements below threshold)

SP: (96.5±1.7)%

Point of Operation

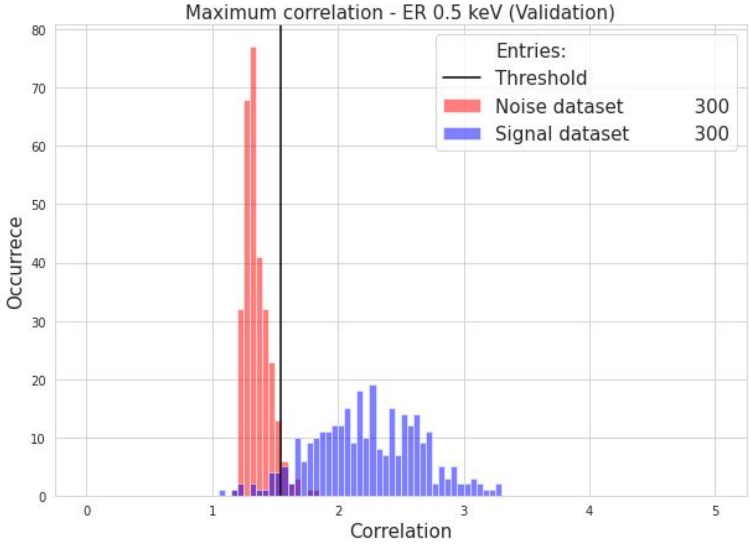


Gaussian with varying window size to sigma = 5.5

Chosen window size

Validation

- The threshold of the training was used on the validation dataset:
  - DET_{sig}: (94.7±2.1)%
  - DET_{noise}: (95.7±1.9)%
  - SP: (95.2±2.0)%

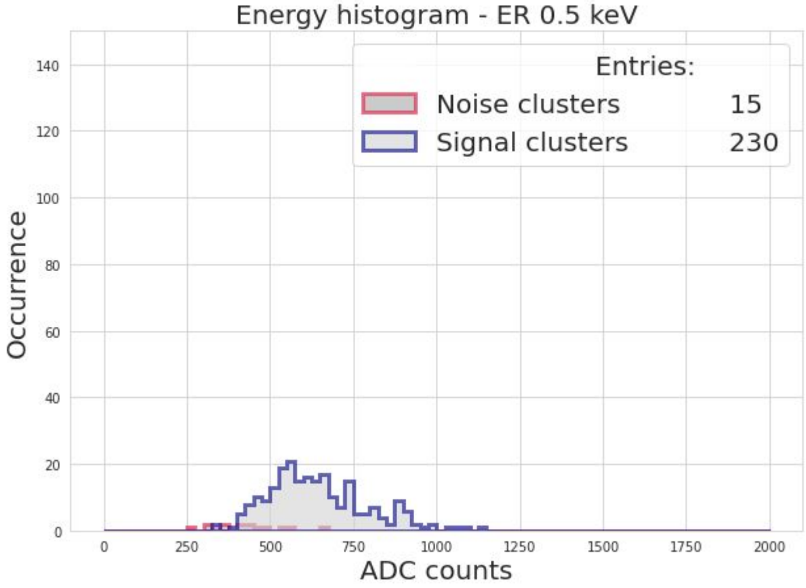


Gaussian with window size equal to 19 and sigma 5.5

Reco file (ER 0.5 keV)

- The reconstruction was used on the signal test dataset (cuts were applied on the borders due to the noise)

- The position of the clusters was compared with the truth information to check which clusters were actually signal.
    - DET_{sig}: (76.7±4.0)%

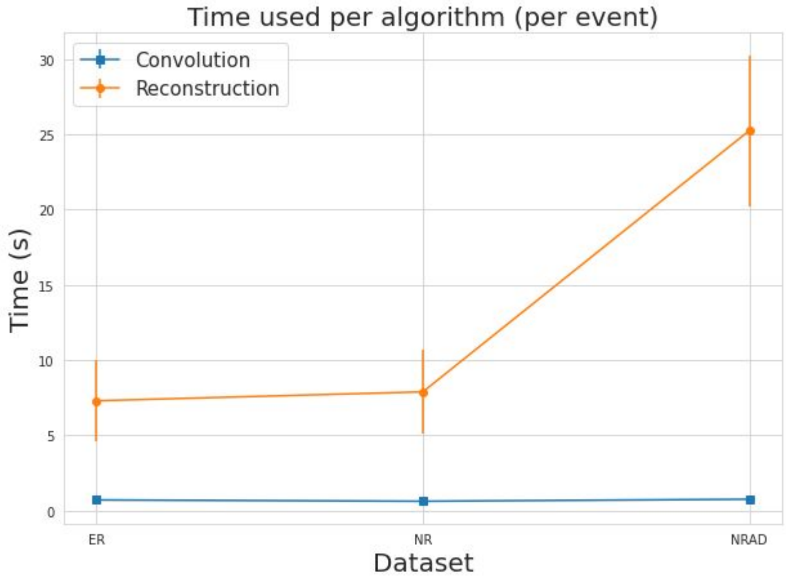


Reco version: Winter23

Time analysis

- ▷ The convolution method needs in average **0.7 seconds** per image (using fft and ifft method).
    - *Invariant to image occupancy*

- ▷ The reconstruction code may need up to 25 seconds per image (depends on the occupancy of the image).



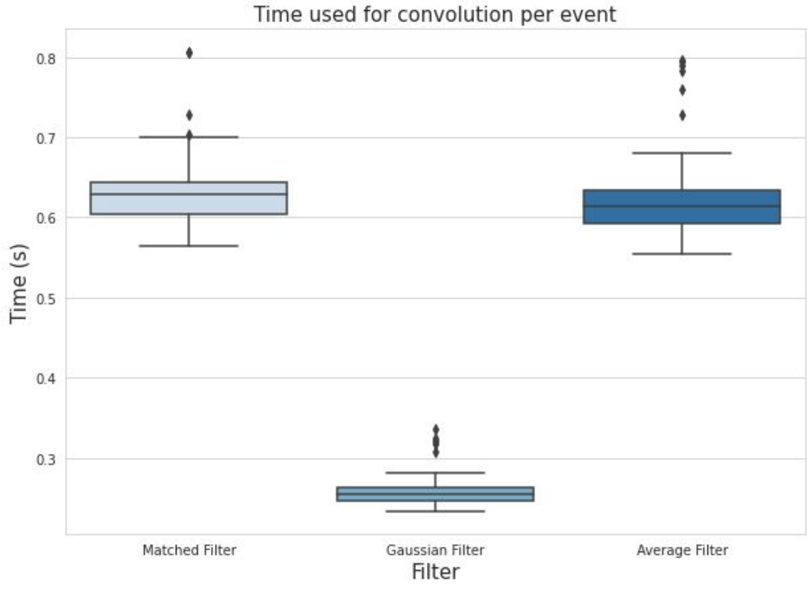
Gaussian with window size equal to 19 and sigma 5.5

Reco version: Winter23

Time analysis

- The Gaussian Filter can be applied with less than **0.3 seconds** using a dedicated function *(that takes advantage of the symmetric nature of the mask).*

- All the times were measured on the cloud (not using the condor queue).



Conclusions and next steps

- In the chosen operation point (SP based):
  - ~95% of noise rejection and
  - ~95% of signal detection efficiency
- Processing time smaller than 1 second
  (independent of the number of tracks present on the image)
- New masks will be tested to improve the method
  - Invariant correlation to position and rotation (see paper)
  - Other Machine Learning approaches will be implemented