Deep Learning techniques for the NEWSdm experiment

ESR:

Artem Golovatiuk

Supervisor:

Giovanni De Lellis

Co-supervisor:

Andrey Ustyuzhanin

Insights is funded by the European Union's Horizon 2020 research and innovation programme, call H2020-MSCA-ITN-2017, under Grant Agreement n. 765710

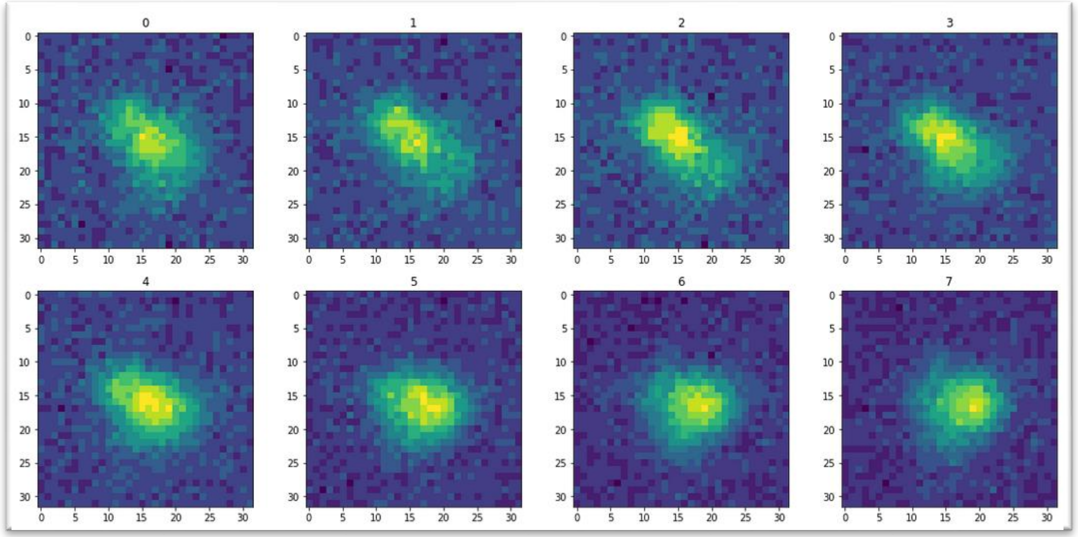
The NEWSdm experiment

- Peculiarities of this experiment:
  - Too complicated to simulate all the processes leading to the track images with Monte Carlo → simulating it with real experimental data. The signal events are produced by exposing emulsion to a Carbon ion beam with the fixed energy, modeling the nuclear recoils from WIMPs. Background samples are exposed to a specific type of background.
  - Barycenter shift analysis with polarized light allows us to go beyond the elliptical fit approximation and acquire additional information about events indistinguishable in unpolarized light.
  - Required background rejection power of $O(10^4)$ with the current status of background reduction techniques used, which is not achievable by conventional methods.
- Machine learning approach:
  - Capable of detecting complex features directly in pixel images from the microscope.
  - Has a variety of algorithm types for different possible applications, including event classification, image denoising or processing the images from the optical microscope to acquire some information.

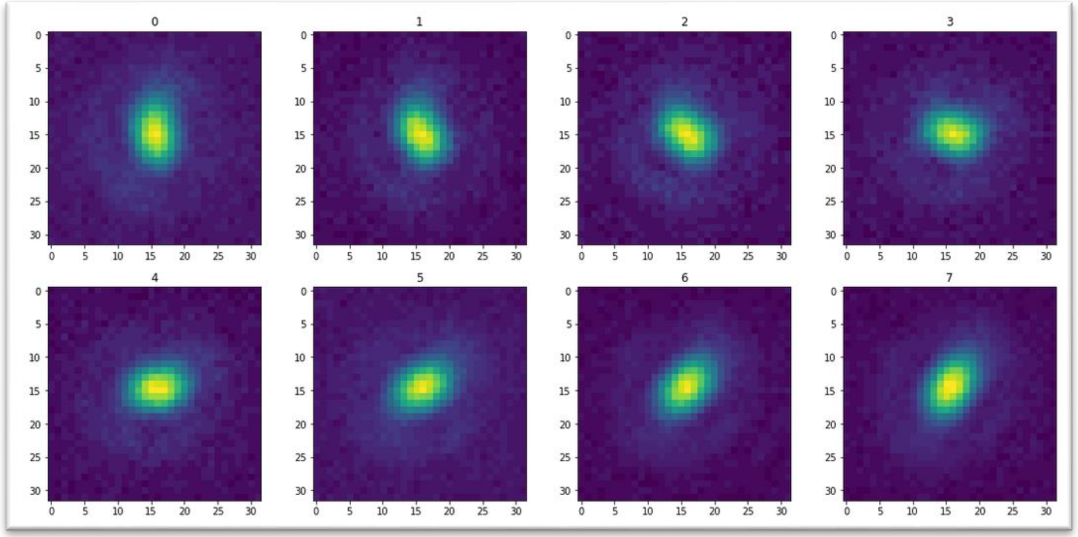
The experimental data

- Signal samples:
  - Exposed to Carbon ion beams with fixed energy
  - C100kev, C60keV, C30keV
  - Low bckg contamination due to small exposure time
- Background:
  - Gamma exposure, simulating intrinsic electrons from C14 decays (via Compton scattering)
  - Random fog: thermal excitations of the crystals

- Current goal: $> 10^5$ events of each type.



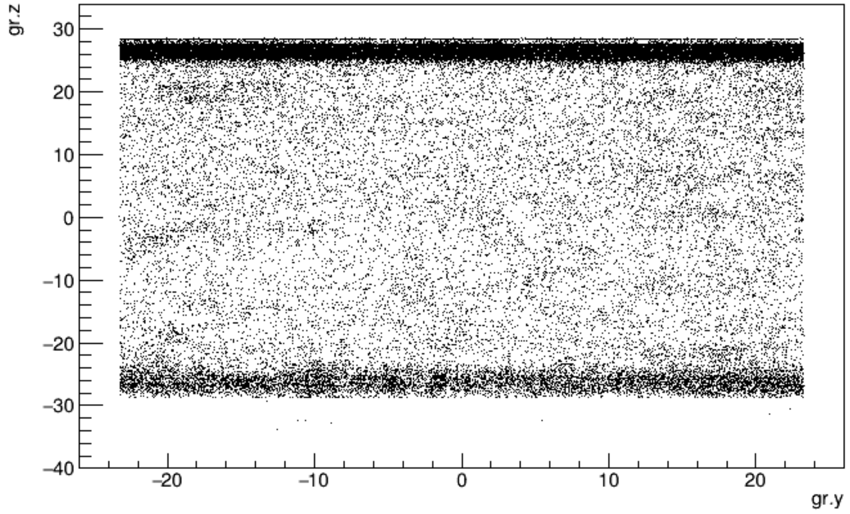
C100keV



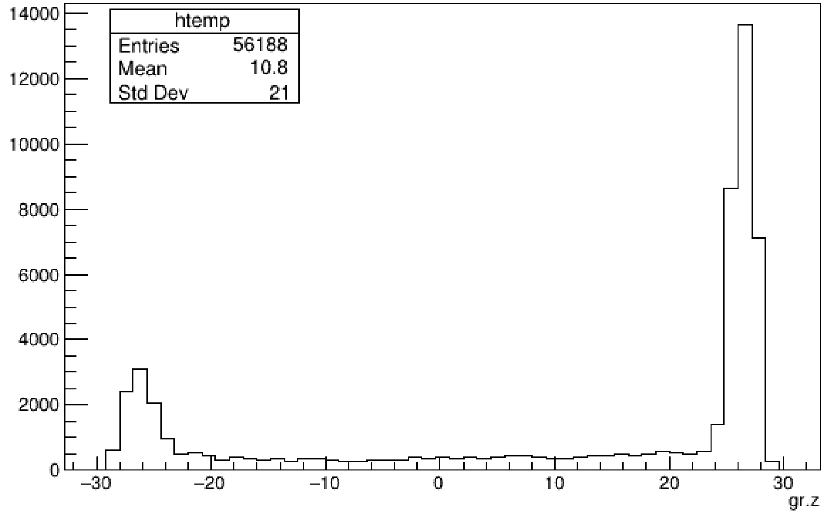
Fog

Data from the emulsions

- Emulsions usually contain different tracks or crystals apart of the main subject of our study.
  - Silver nanoparticles on the top and bottom surfaces to configure the microscope during scanning
  - Radioactive source can produce extra decay products (i.e. α-particles) that have to be discarded.
- Emulsion sample needs to be clearly understood before starting the Machine Learning analysis.



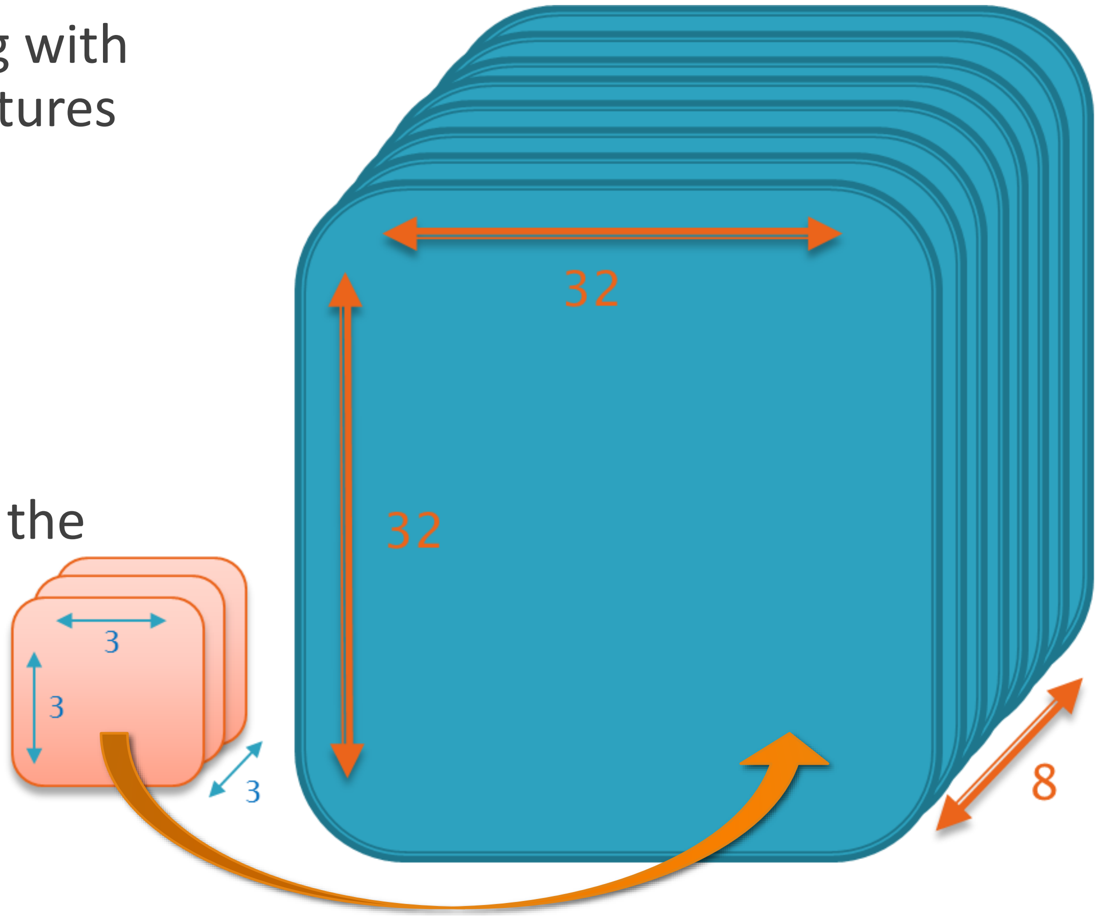
Fog sample with silver nanoparticles on the surfaces



Events distribution over Z axis in the Fog sample

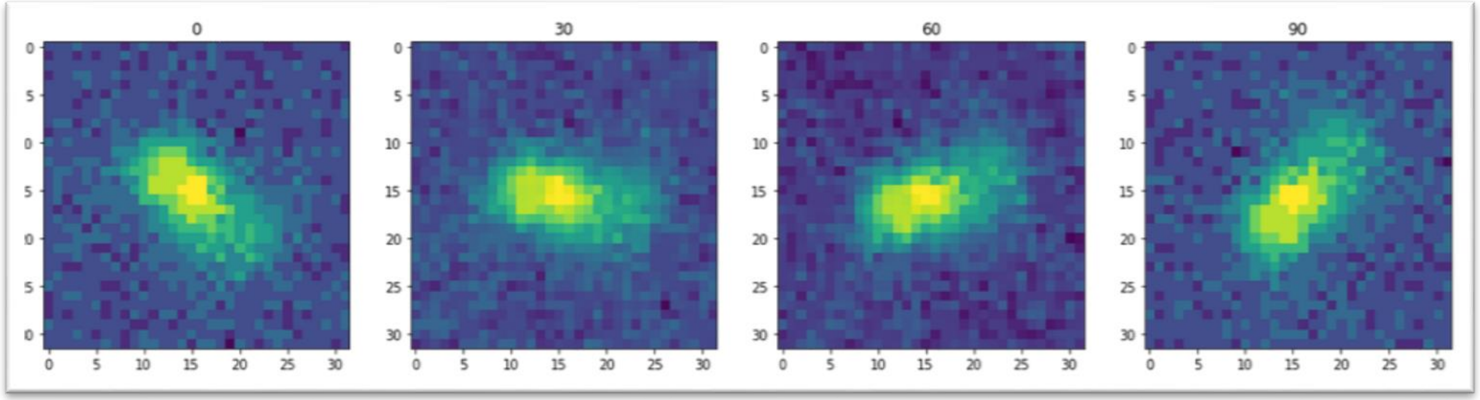
3D Convolutional Neural Networks

- Convolutional approach is designed for working with images. It is capable of discovering complex features of the images and gaining high performance.
- Stacking together images for different light polarizations to obtain a 3D image.
- Empty polarization images are filled with zeros.
- Network “scans” not only plain image, but also the “polarization” axis.
- Allows Network to learn correlations between features of different polarization images.



Track image rotations

- Carbon ion beam has a specific direction, while background is mostly isotropic.
- To force Network to find more complex and less obvious features we want to make both background and signal isotropic.
- Solution: random rotations of a small subsets of data.
- Positive side-effect: it can significantly increase the size of the dataset.
- Risks: can produce some artifacts, can lead to overfitting the training images if misused.

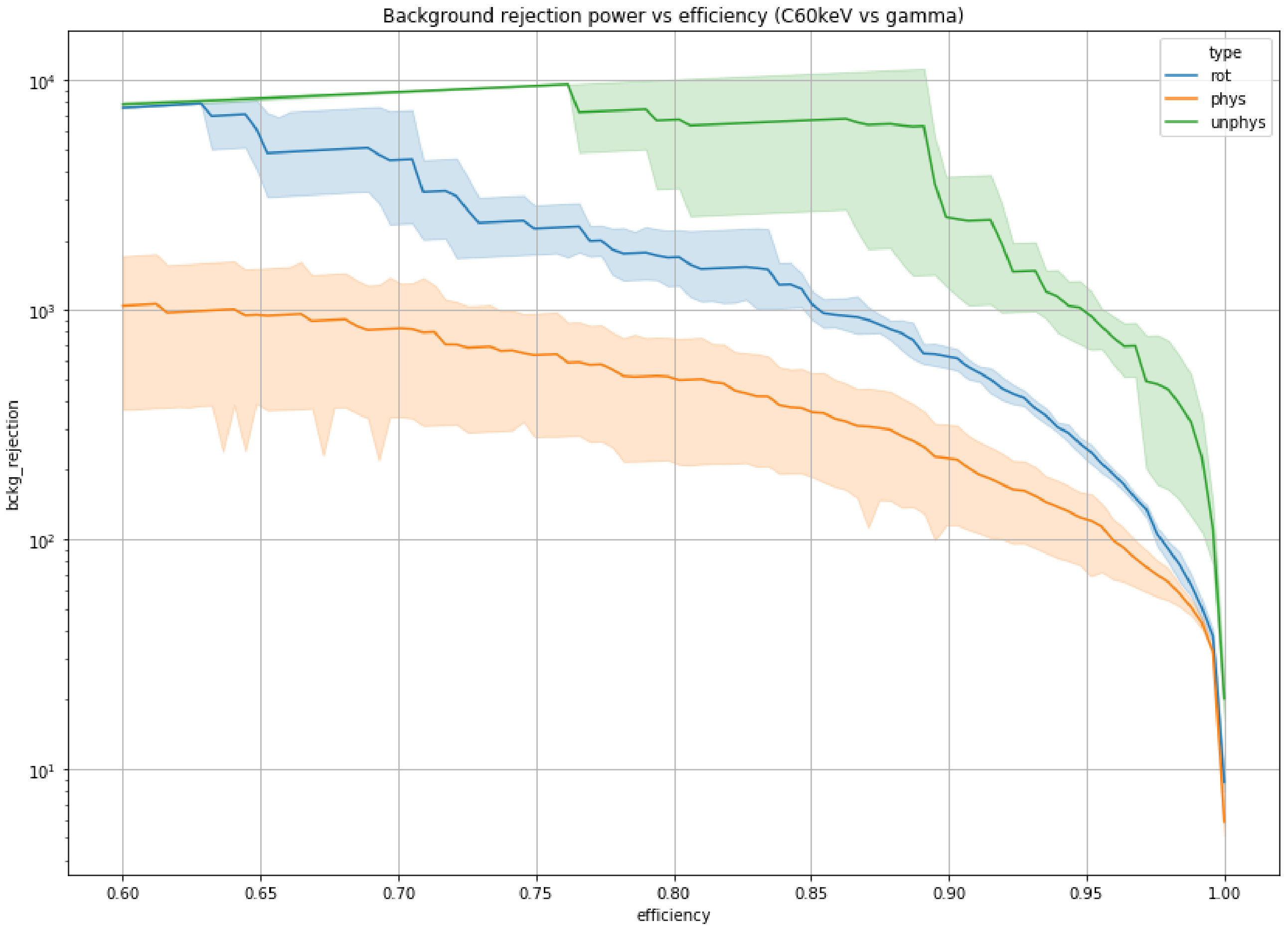


Rotations of C100keV track

Preliminary results (C_{60keV} vs $\gamma_{+\alpha}$)

- We compared several approaches by the way of treating the training data:
  - "Physical" – using only physically motivated data (limited size of the dataset)
  - "Unphysical" – adding to physically motivated data the samples we are not interested in, like silver nanoparticles images on the surface of emulsion (images themselves still look very similar)
  - "Rotational" – enlarging the physical dataset by random rotations several time (2 for now)

- The validation data used to check the models performances contains only intact physical data, which is not augmented or modified.



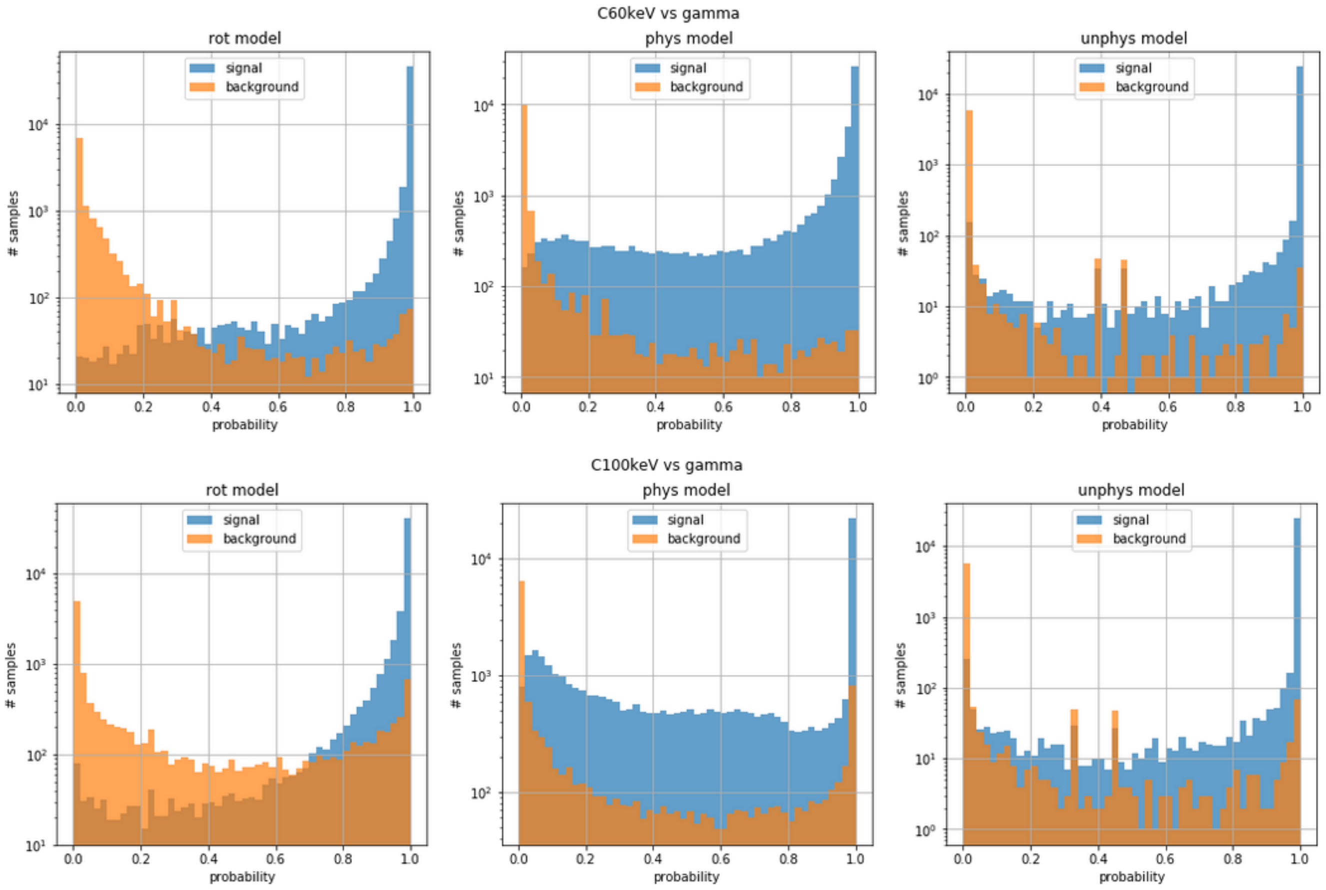
Preliminary results (C vs $\gamma_{+\alpha}$)

- Gamma sample seems to remain contaminated with some alphas:
  - C100keV is more similar with background than C60keV, which might be due to the “tail” of alpha distribution that could not be discarded by cuts.

- Bigger training dataset has significant impact on the models performance.
  - Current physical gamma background is $\sim 10^4$ events

- Feasible solution: scan another clean gamma sample and acquire more tracks.

(already brought from Japan)

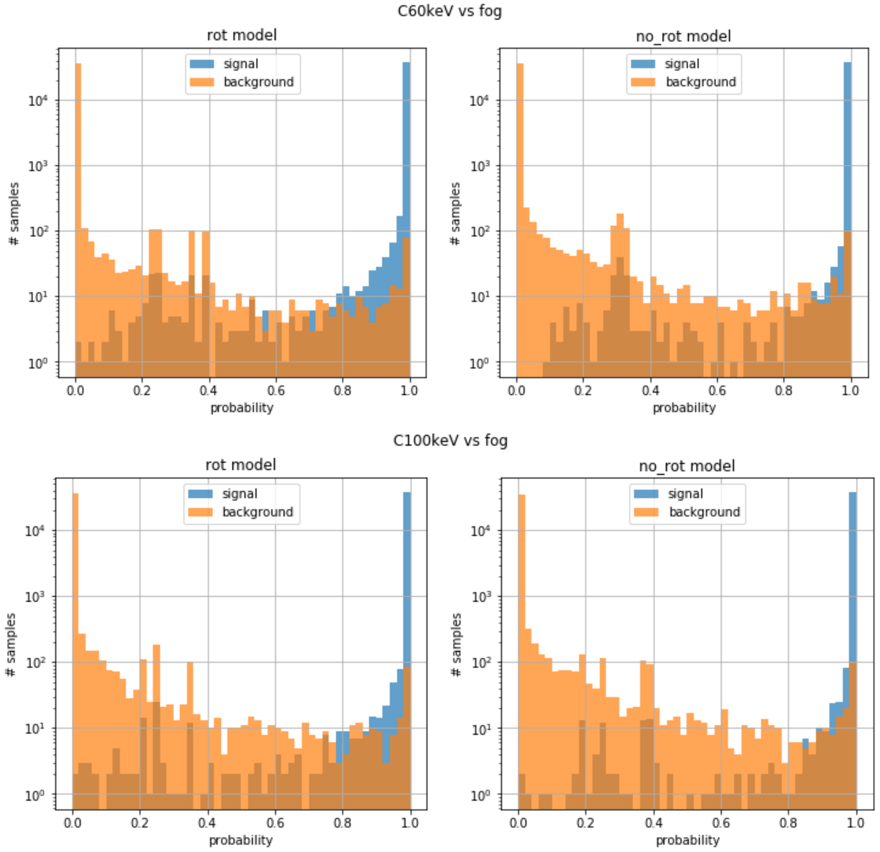


Histograms of the models’ outputs

Preliminary results (C vs fog)

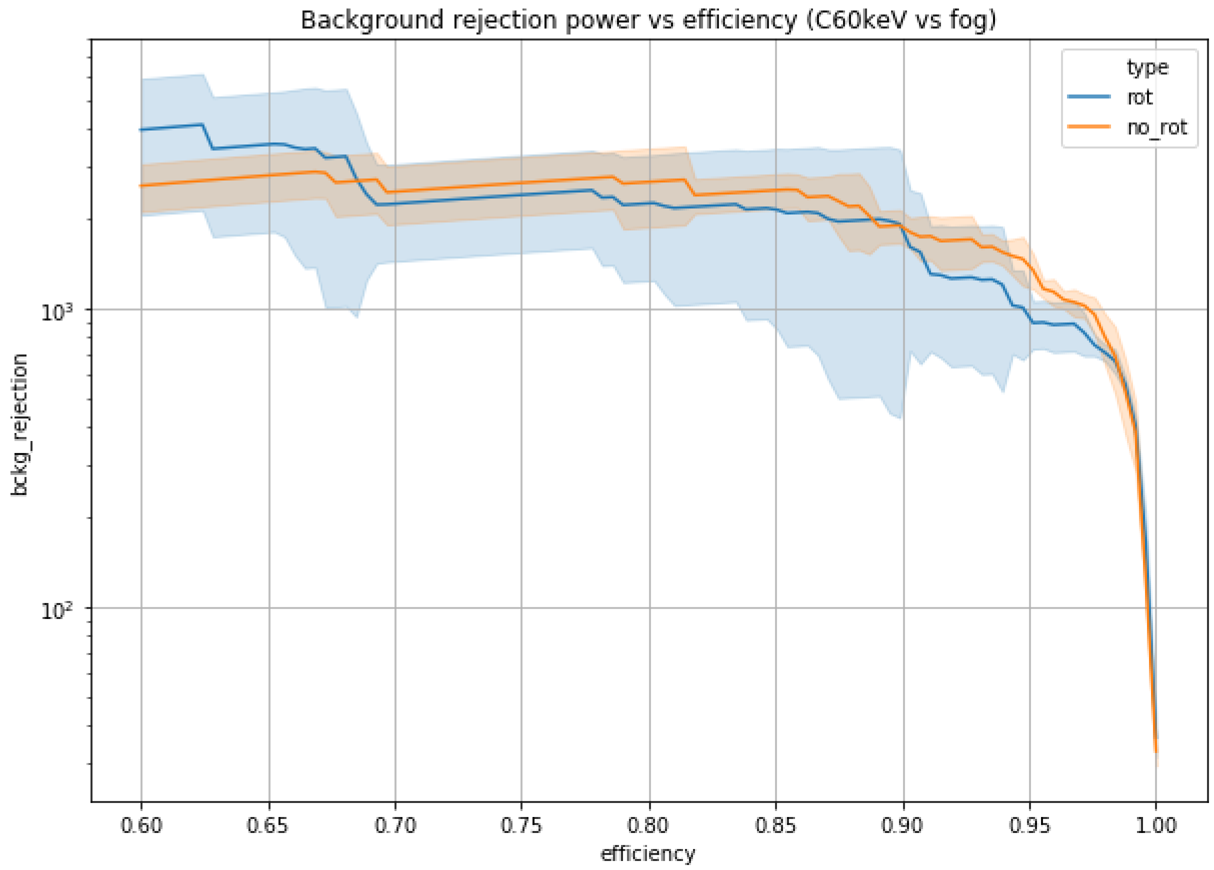
- Large enough clean dataset with $\sim 10^5$ events.
- No such significant difference between C ions with various energies.
- No obvious difference between models with rotations and without either.
- ✓ (more visualization on the next slide)

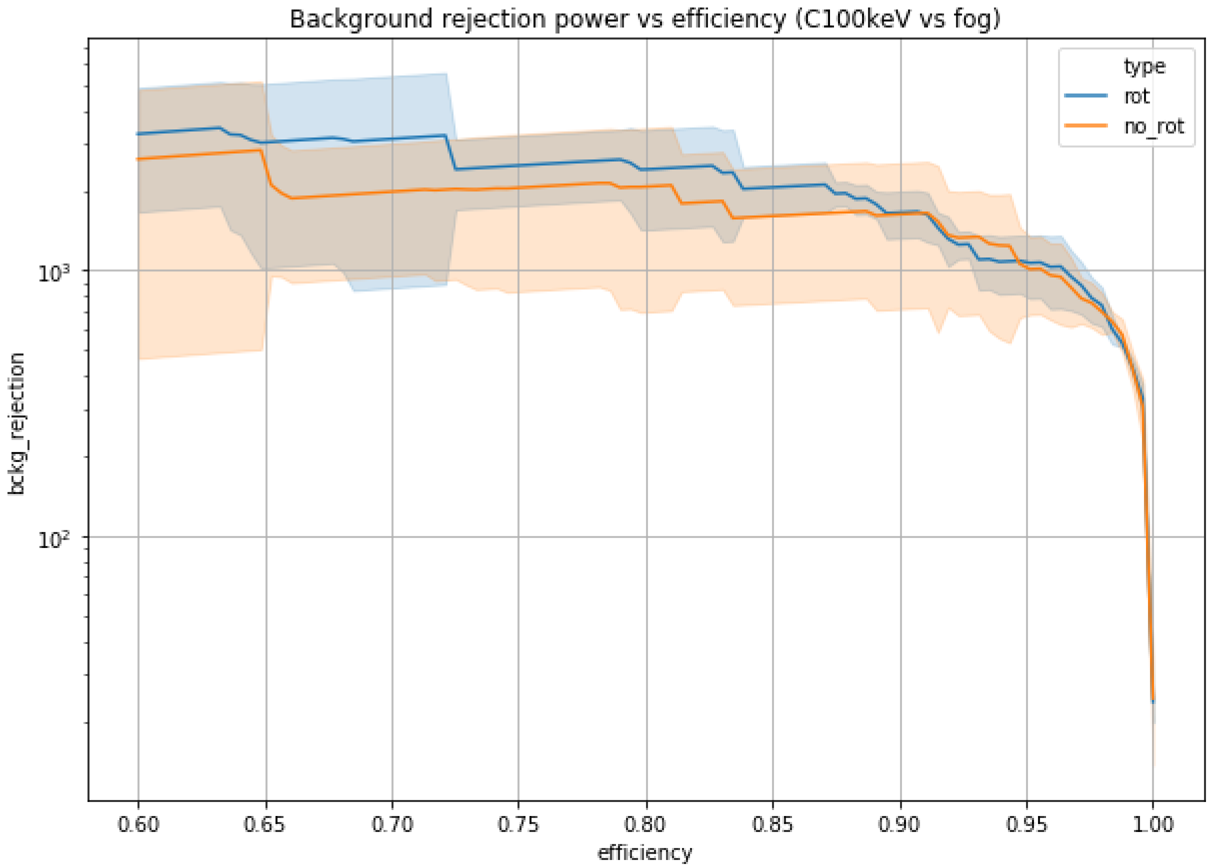
- Image rotations imply isotropy, while do not hurt nor improve the performance significantly.
- Might be a sign that further enlargement of the dataset is not so crucial.



Histograms of the models' outputs

Preliminary results (C vs fog)





Summary

Conclusions:

- Clear understanding of the emulsion sample prior to application of Deep Learning techniques is crucial.
- Both physically motivated data and redundant contaminations of the emulsions might be used in the model training.
- Random rotations do not harm the performance, while implying important condition of isotropy.
- Deep Learning algorithms allows us to achieve background rejection power of 10^3 with still some room for improvement.
- The increasing dataset to 10^6 events is not guaranteed to provide crucial enhancement in performance.

Further plans:

- Carry on the study for “fog” and “gamma” background samples.
- Deeper study of the models performance on unseen data (separate scan or sample).

- Adding information about the barycenter shift to the algorithm’s input.
- Using images from the color camera, since scattered wavelength depends on the form of the silver grains.