The 8th International Conference on Deep Learning in Computational Physics,
June 19-21, 2024

GAMMA/HADRON SEPARATION IN THE TAIGA EXPERIMENT WITH NEURAL NETWORK METHODS

Elizaveta Gres[1], A. Kryukov[2], P. Volchugov[2], A. Demichev[2], D. Zhurov[1], Ju. Dubenskaya[2], S. Polyakov[2], S. Polyakov[2], A. Vlaskina[2]

1 - Applied Physics Institute of Irkutsk State University, Irkutsk;
2 - Lomonosov Moscow State University, Skobeltsyn Institute of Nuclear Physics, Moscow.

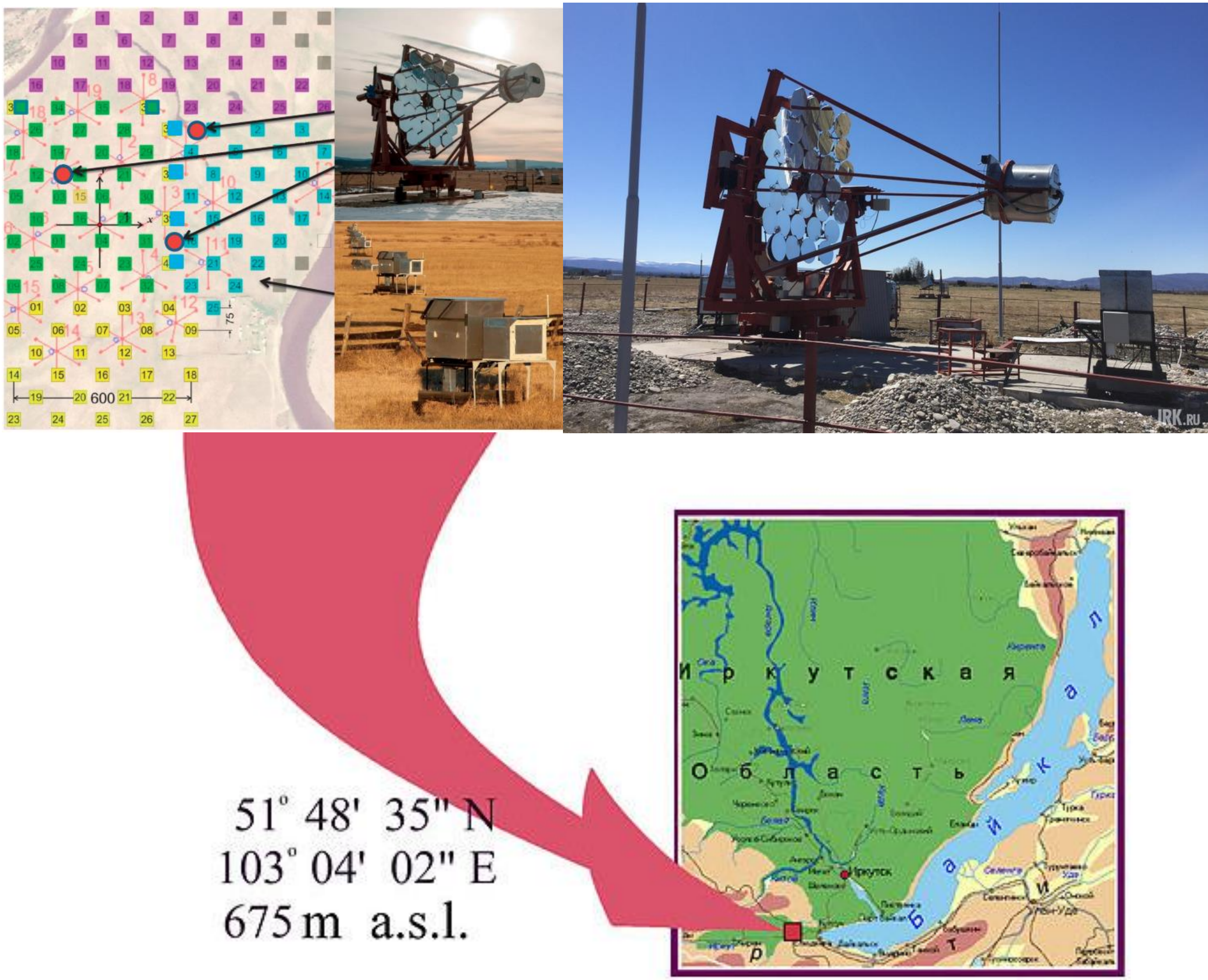
TAIGA AND TAIGA-IACT

- Imaging Air Cherenkov Telescopes (IACT) are located in Tunka Valley (50 km from Lake Baikal), Republic of Buryatia, at the astrophysical gamma observatory TAIGA of the API ISU.

- They are telescopes-reflectors with a 4-meter segmented spherical mirror. In its focus there is a camera, representing as matrix of ~600 PMTs.

- IACTs detect Cherenkov light from Extensive Air Showers (EASs), originating from the interaction of cosmic or gamma radiation with the atmosphere.

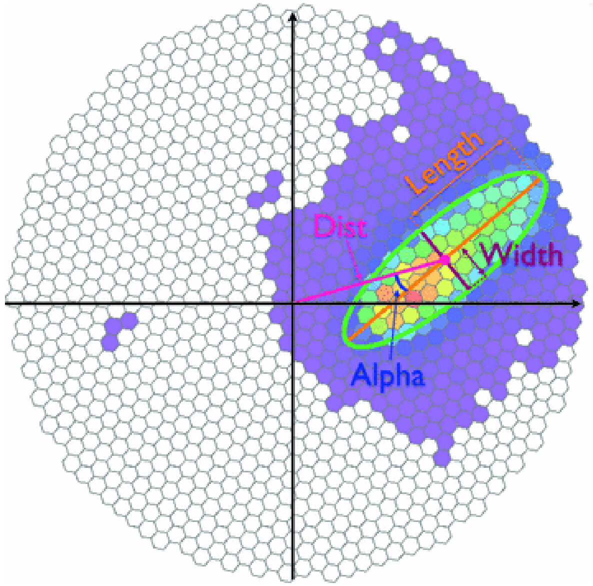
- This type of detectors helps to select gamma-ray events from the hadron background based on obtained EAS images.

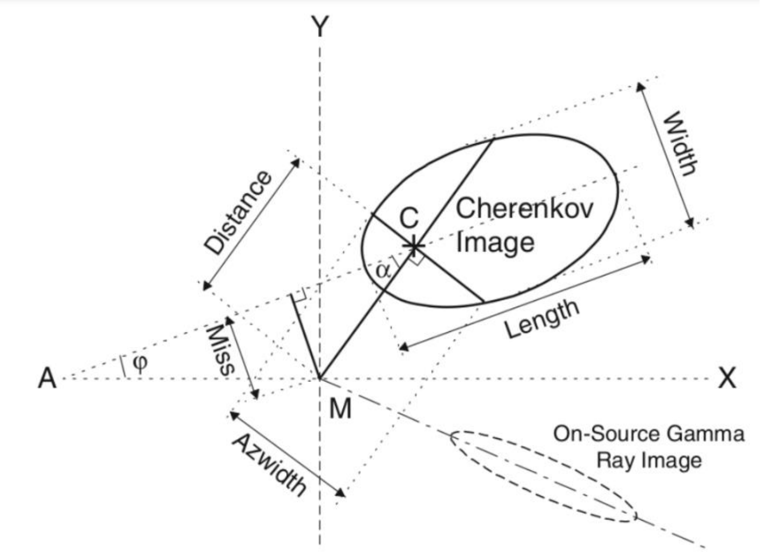


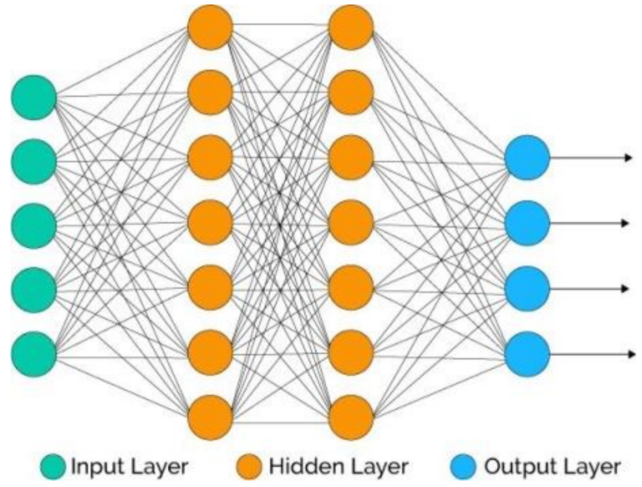
IACT DATA PROCESSING

- I. Hillas Parameter cuts - Standard image processing method
  - Calculation of image first, second momentums named Size, Width, Length, etc.;
  - Search dependences and optimal criteria in these parameters to distinguish gamma-events from hadron background;
  - Application those cuts and estimation of the signal from gamma-source.

- II. CNN image classification – machine learning method
  - Creation of MC training and validation set according to experimental data;
  - Set the neural network model - CNN;
  - Image modifications for CNN;
  - Assessment of neural network in the classification problem;
  - Application CNN and estimation of the signal from gamma-source.







PROBLEM AND AIM OF THE RESEARCH

- TAIGA-IACT detect gamma EAS in TeV energy range – the ratio of gamma to hadron fluxes is *$1:10^4$*. Hillas Parameter cuts are not allow the researchers to separate gamma-events from hadron events completely.
  - In the standard method, cuts are not unified: there could be different combinations of Hillas Parameters cuts in the task of IACT events classification.
    - Standard method is the only method of analysis of TAIGA-IACT experimental data now.

- **The aim of the research**: development of neural network methods for processing and analyzing data from the Cherenkov telescopes TAIGA.
- Tasks:
  - To collect required MC data, train and validate the training of the neural network model.
  - To find the solution for neural network method for high class imbalance for TAIGA-IACT data analysis.
  - To test CNN method in the case of Crab Nebula observation;
  - To compare obtained results with standard method, based on Hillas parameters cuts.

CNN ARCHITECTURE

- **Number of weights**: 7 millions
- **Loss function**: binary crossentropy.
- ***Training***: 50 400 MC gamma events & 52 400 thousands experimental hadrons
  - Generated by CORSIKA and with TAIGA detector modeling software.
  - Energy range: 1-200 TeV
  - Zenith angle: 30-40°
  - ***Validation***: 16900 gammas & 17300 experimental hadrons

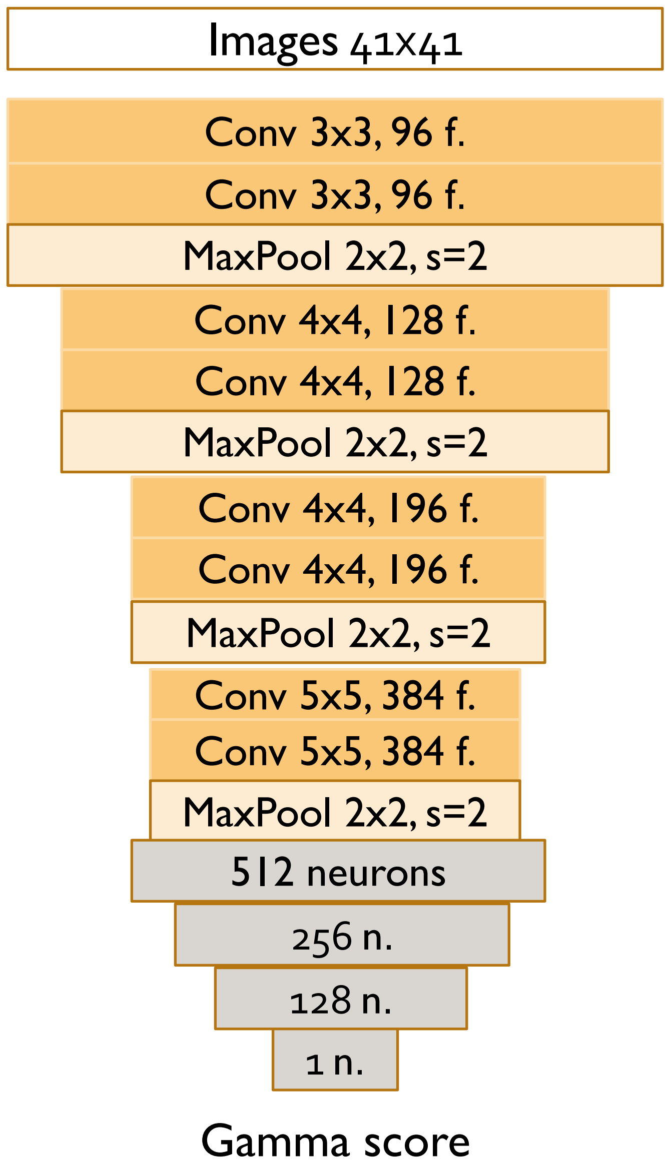
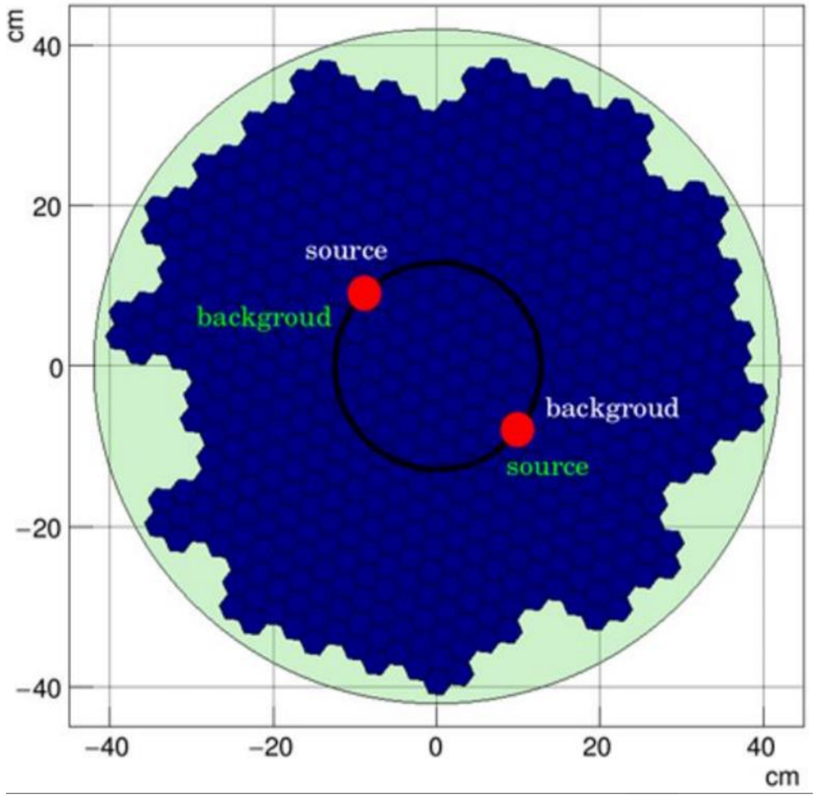
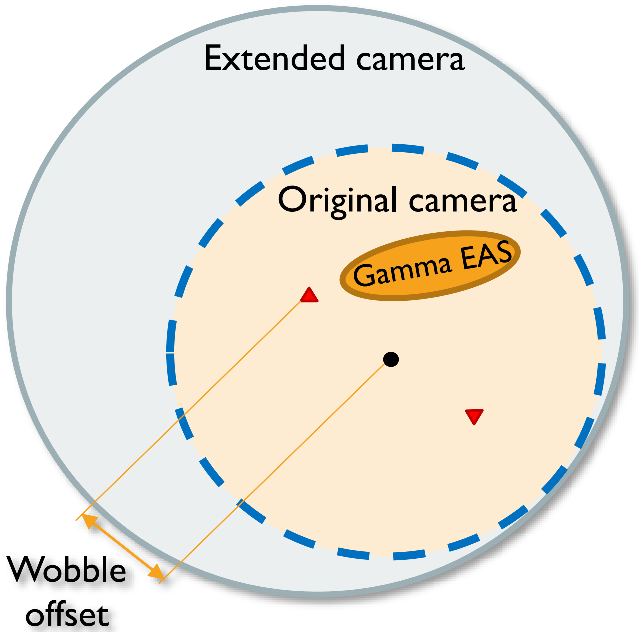
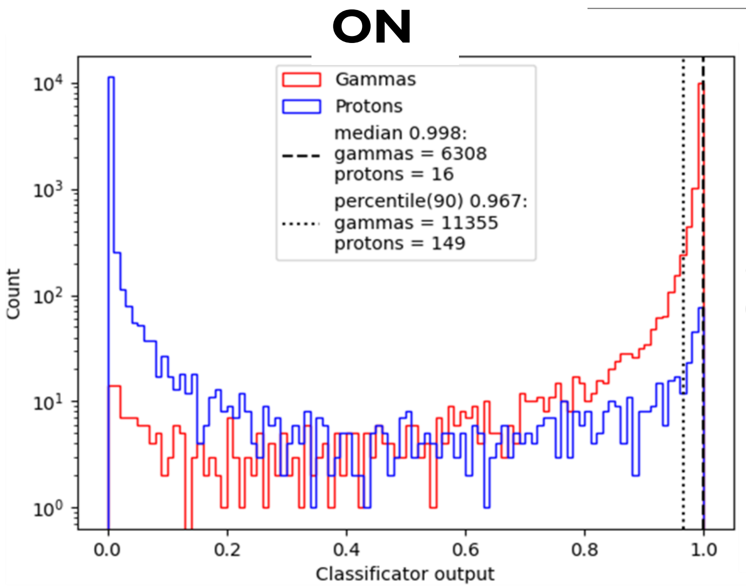


IMAGE PREPROCESSING

- **I.** Image cleaning;
- **II.** Image modification for Wobble mode of observation:
  - Wobble mode - IACT operating mode, in which the gamma radiation source is located *not in the center of the camera*, but *at a fixed distance from it.* In this mode the source periodically shifts from its position to another (opposite) offset position







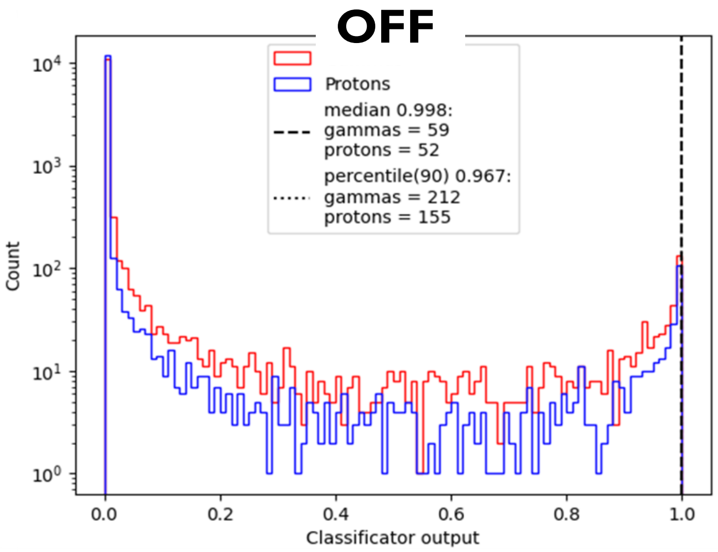
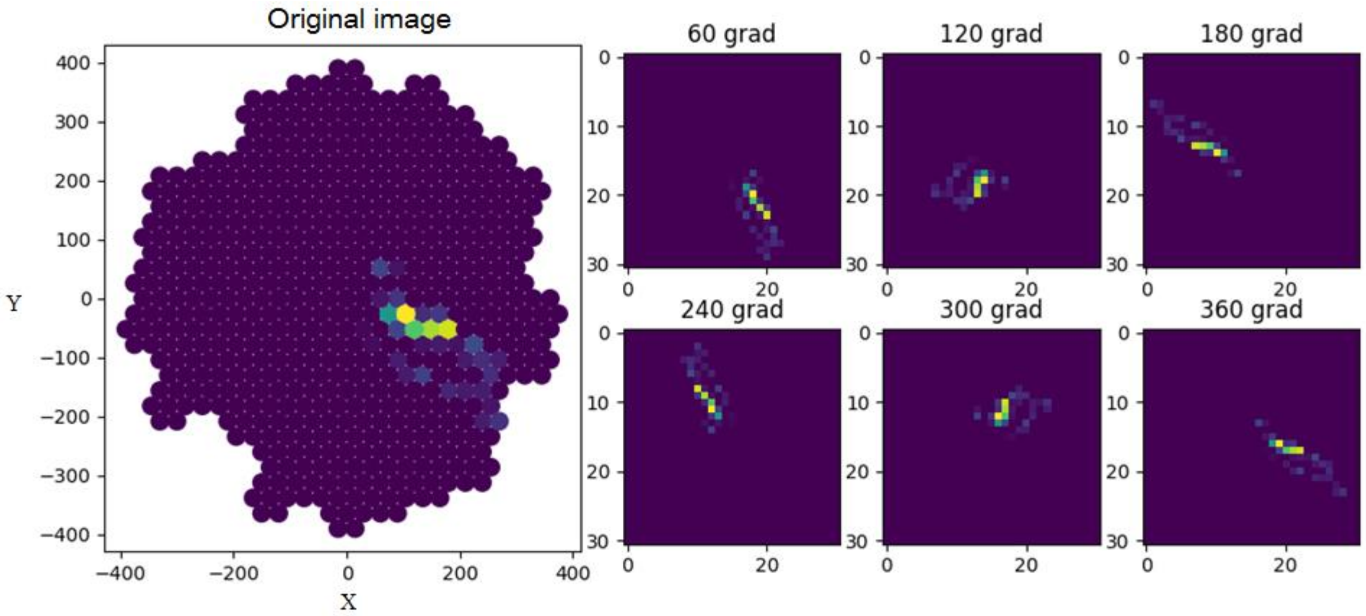
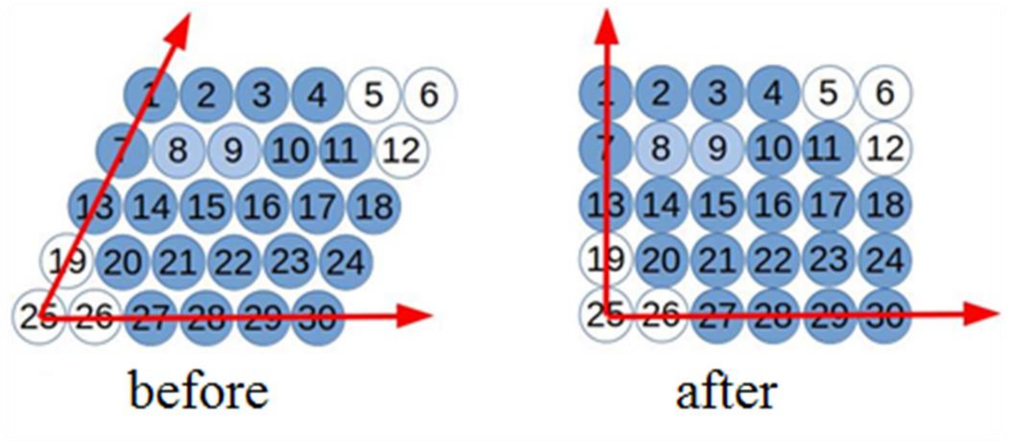


IMAGE PREPROCESSING

- **III.** Train set expansion by 60° rotation (not applied in validation and test).
- **IV.** Transformation to a square image with the axial method;
- **V.** Logarithmic amplitude normalization:

$$\frac{\log(1+x)}{4.5}$$, x - pixel amplitude;

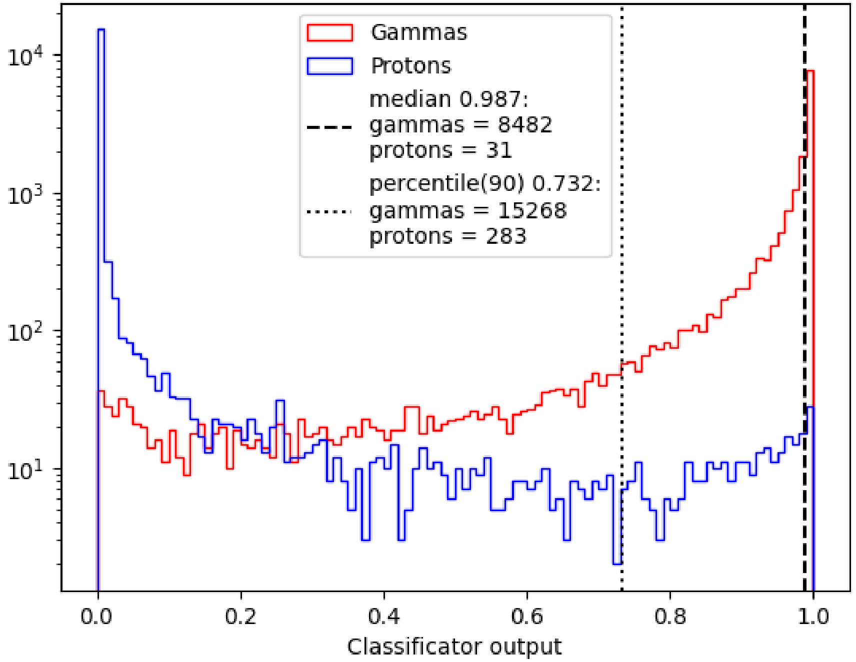




CNN TRAINING AND VALIDATION

- Training:
  - Gamma: 50.4 thousands. * 6 rotations
  - Hadrons: 52.4 thousands. * 6 rotations.

- Validation:
  - Gamma: 16.9 thousands
  - Hadrons: 17.3 thousands



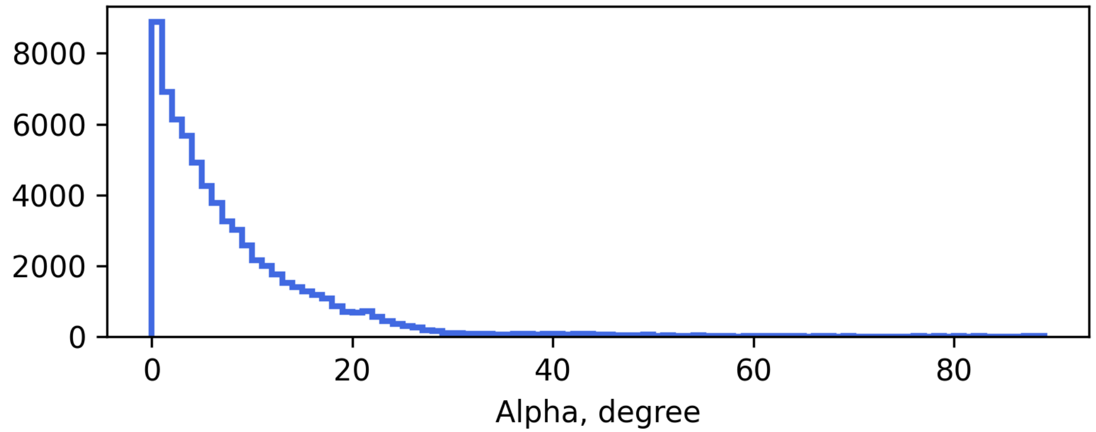
- Thus, the background suppression at the 50% gamma loss threshold is **no more than 1/1000.**

CNN TRAINING AND VALIDATION

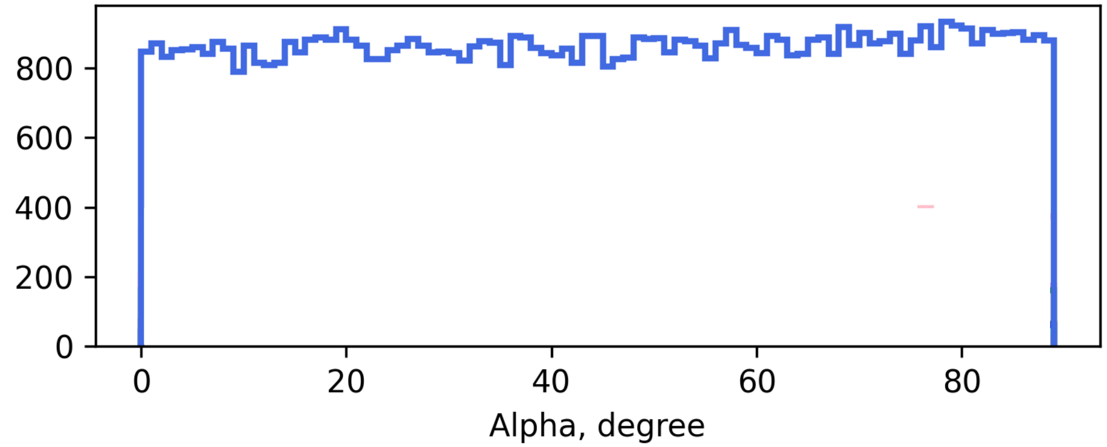
- **Conditions for events**: lg(Size) >2.08, Alpha<20° or Alpha<6°
  - We are not interested in images severely distorted by cleaning;
  - Remains:
    - Alpha<20°: gamma ~90%, hadrons ~20%
    - Alpha<6° : gamma ~50%, hadrons >10%
  - Significantly reduces the number of events and speeds up data processing;
  - Let CNN to consider the specific point on the camera more closely.

- Hence, training with Alpha<20° and lg(Size) >2.08:
  - MC гамма: 38400 (1-200 TeV, 30-40° zenith)
  - Hadrons: 40000

Gamma events

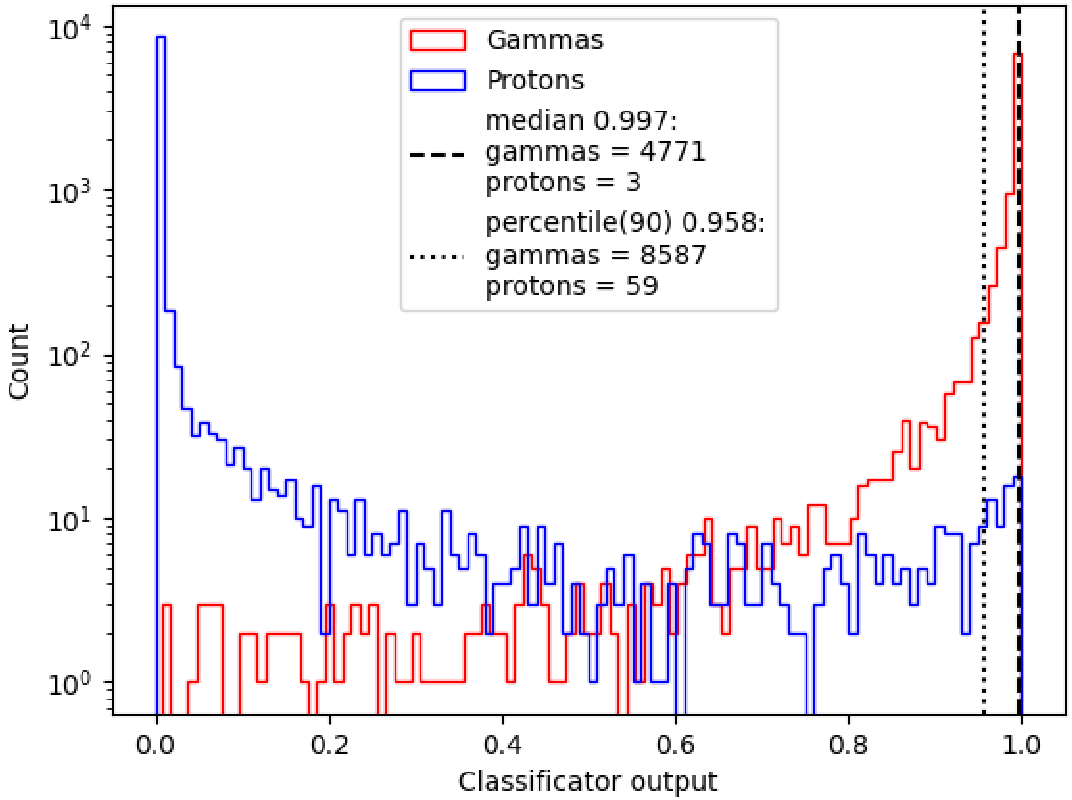


Proton events



CNN TRAINING AND VALIDATION

Validation: 9500 MC gamma и 9600 hadron events



At class threshold 0.9965:

- 4771 gamma events
- 3 hadrons

«Gamma-hadron» ratio is more 1/100, leading the general background suppression from 1/2000 to 1/10000

SEPARATION OF THE GAMMA SIGNAL FROM THE CRAB NEBULA AND COMPARISON WITH STANDARD METHODS

- After estimation of background suppression, 6 dates of Crab Nebula were selected in November of 2019-2020 observation season and analyzed using Hillas Parameters (HP) cuts applied in the TAIGA collaboration and CNN classifier.

- HP cuts:

Size > 120 phe; 0.36° < Dist< 1.44°;
0.024° < Width< 0.068°·lg(Size)-0.047°;
Lendth > 0.145°·lg(Size); Θ^2 <0.05°

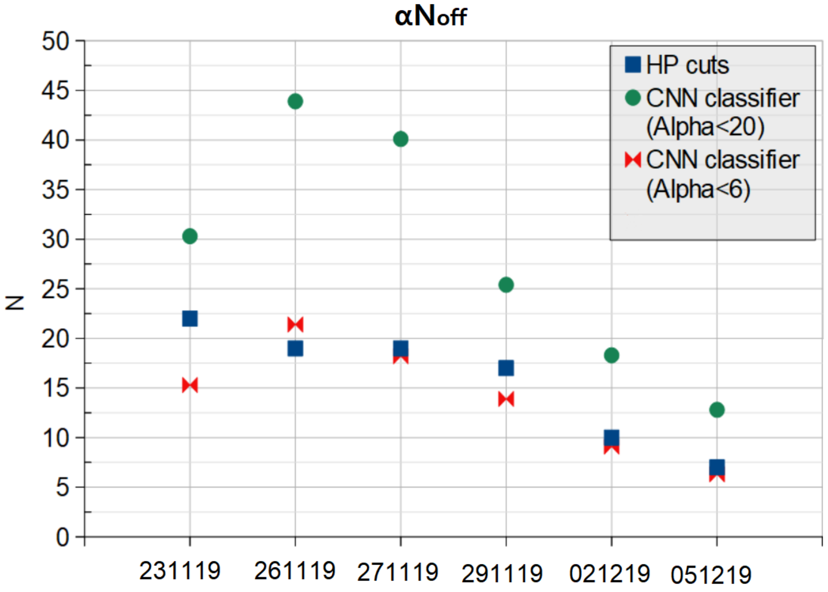
- Significance S_{LiMa} (*α – time ratio of ON- and OFF-observations*):

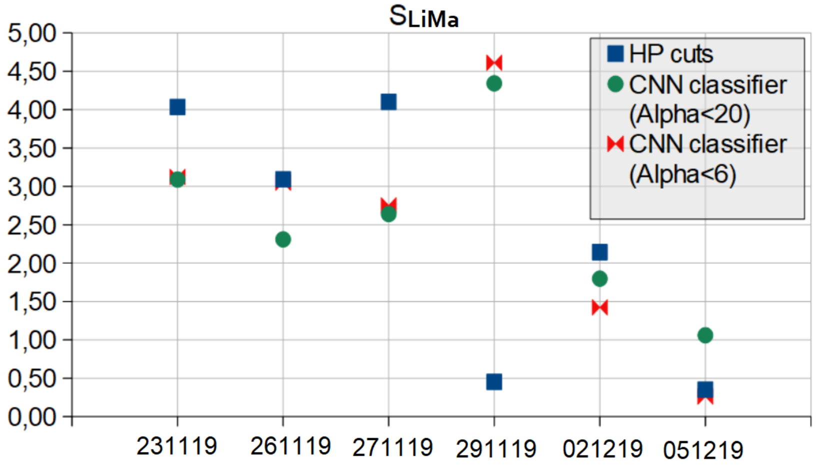
$$S_{LiMa} = \sqrt{2}\left\{N_{on}\ln\left[\frac{1+\alpha}{\alpha}\left(\frac{N_{on}}{N_{on}+N_{off}}\right)\right] + N_{off}\ln\left[(1+\alpha)\left(\frac{N_{off}}{N_{on}+N_{off}}\right)\right]\right\}^{1/2}$$

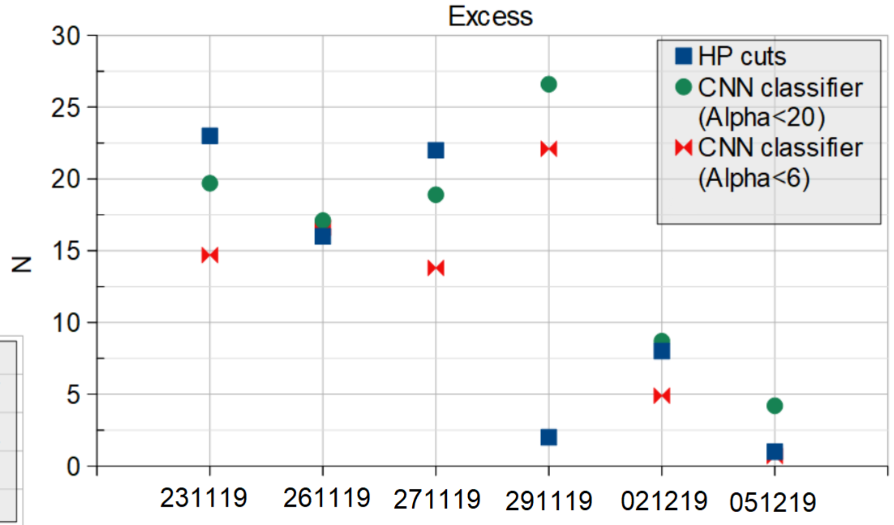
Observation time, hours	Processing method	αN_{OFF}	N_{ON}	Excess	α	S_{LiMa}
21,75	HP cuts	94	166	72	0.01	**6.32**
	CNN (Alpha<20)	170.8	266	95.2	0.01	**6.37**
	CNN (Alpha<6)	84.2	157	72.8	0.01	**6.67**

- Coincidence – 46 events in ON-point (CNN (Alpha<20)).

SEPARATION OF THE GAMMA SIGNAL FROM THE CRAB NEBULA AND COMPARISON WITH STANDARD METHODS

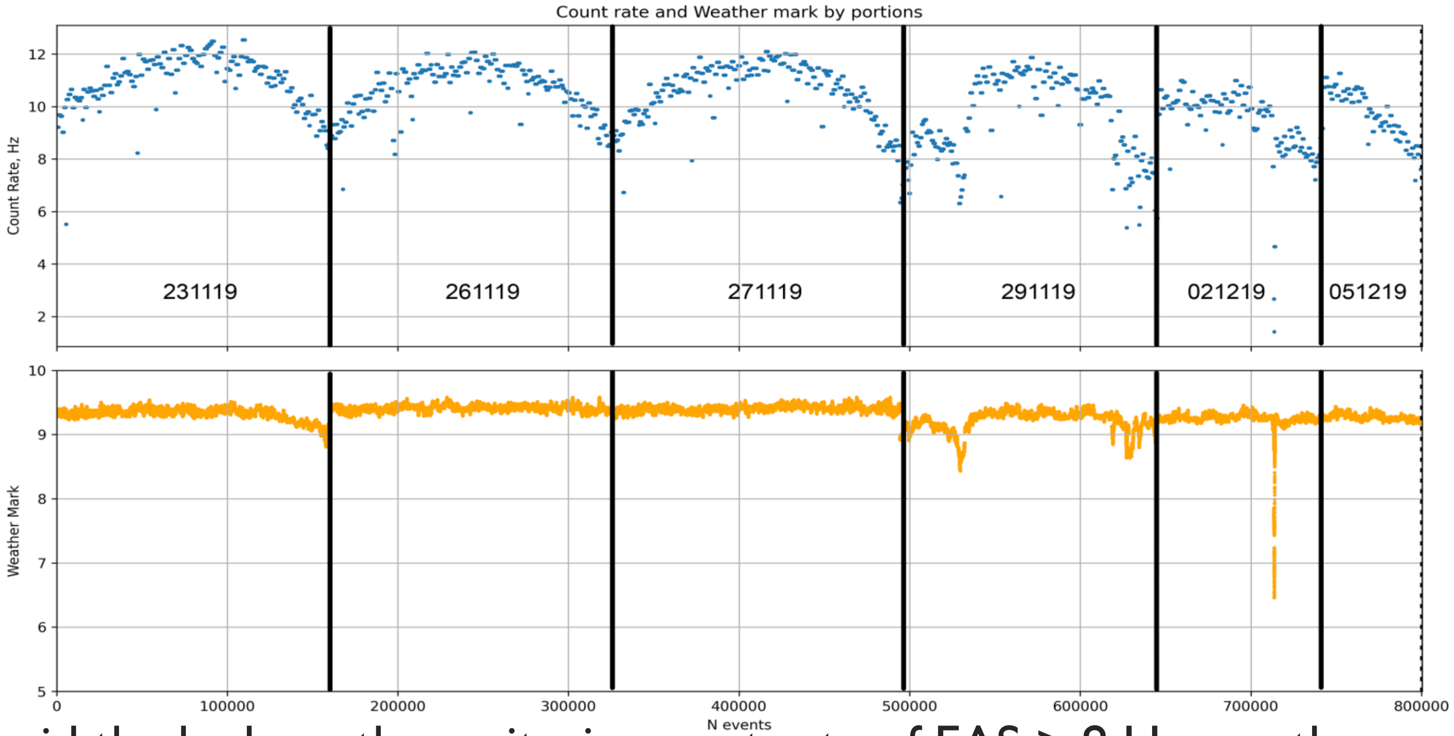






- Background suppression is good for Alpha<6, yet, according simulation, we loose 50% before CNN classification;
- S_{LiMa} doesn't differ more then 1σ;
- the difference in results is quite large for 291119 run – atmosphere transparency fluctuation?

SEPARATION OF THE GAMMA SIGNAL FROM THE CRAB NEBULA AND COMPARISON WITH STANDARD METHODS

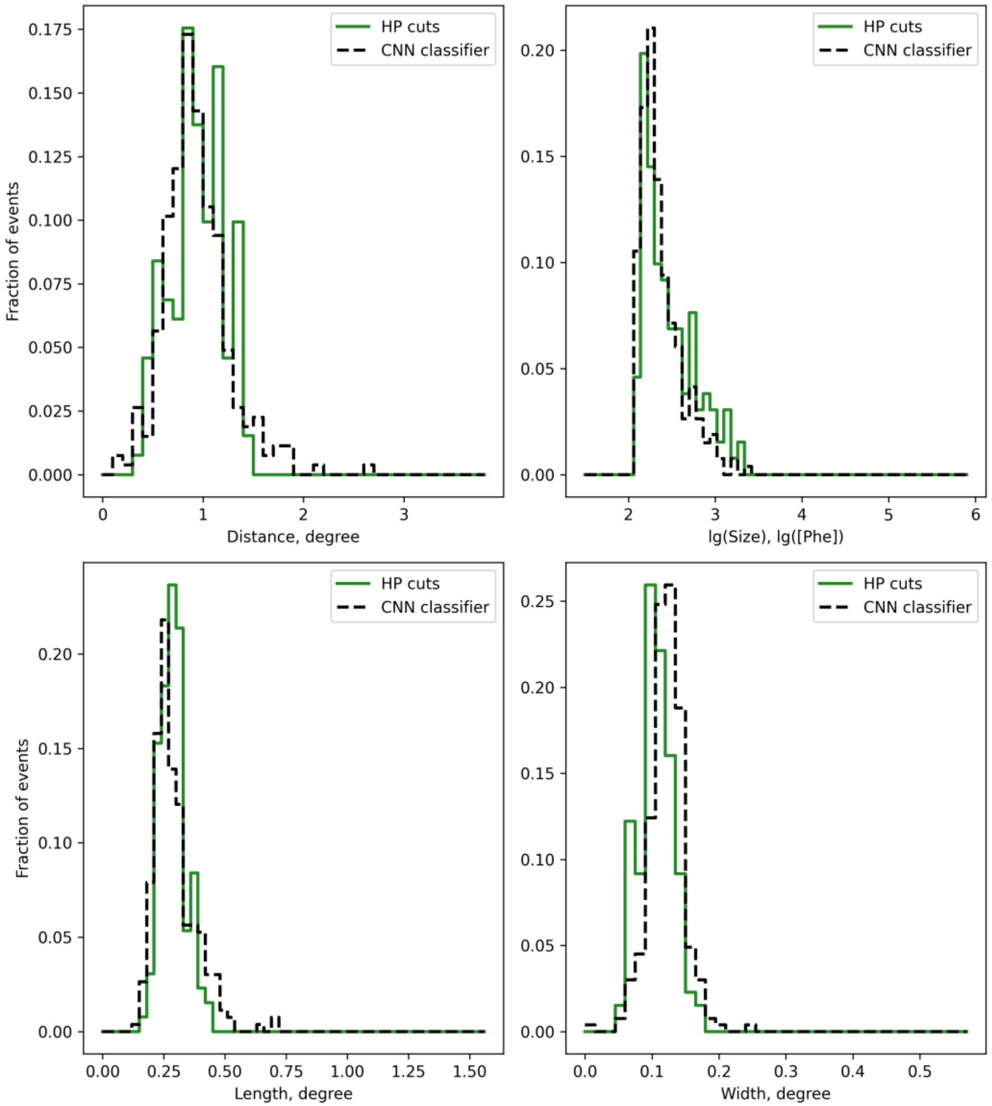


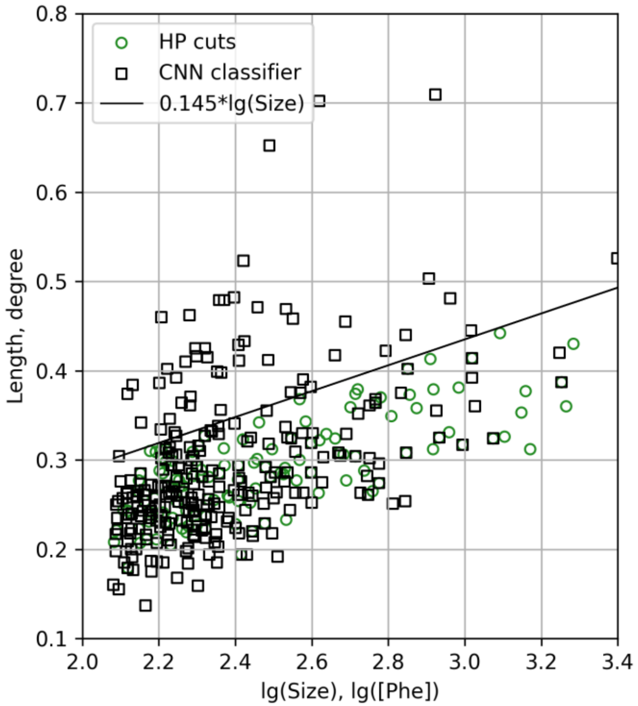
- To avoid the bad weather, criteria: count rate of EAS > 8 Hz; weather mark > 9

Observation time, hours	Processing method	αN_{OFF}	N_{ON}	Excess	α	S
21	HP cuts	94	160	66	1/10	**5.84**
	CNN (Alpha<20)	161.8	253	91.2	1/10	**6.26**
	CNN (Alpha<6)	79.3	148	68.7	1/10	**6.48**

SEPARATION OF THE GAMMA SIGNAL FROM THE CRAB NEBULA AND COMPARISON WITH STANDARD METHODS

Hillas parameters distributions for ON point





- The shape of the Hillas parameters distributions approximately coincide.
- Length and Width dependance on lg(Size) also roughly correspond.

CONCLUSION

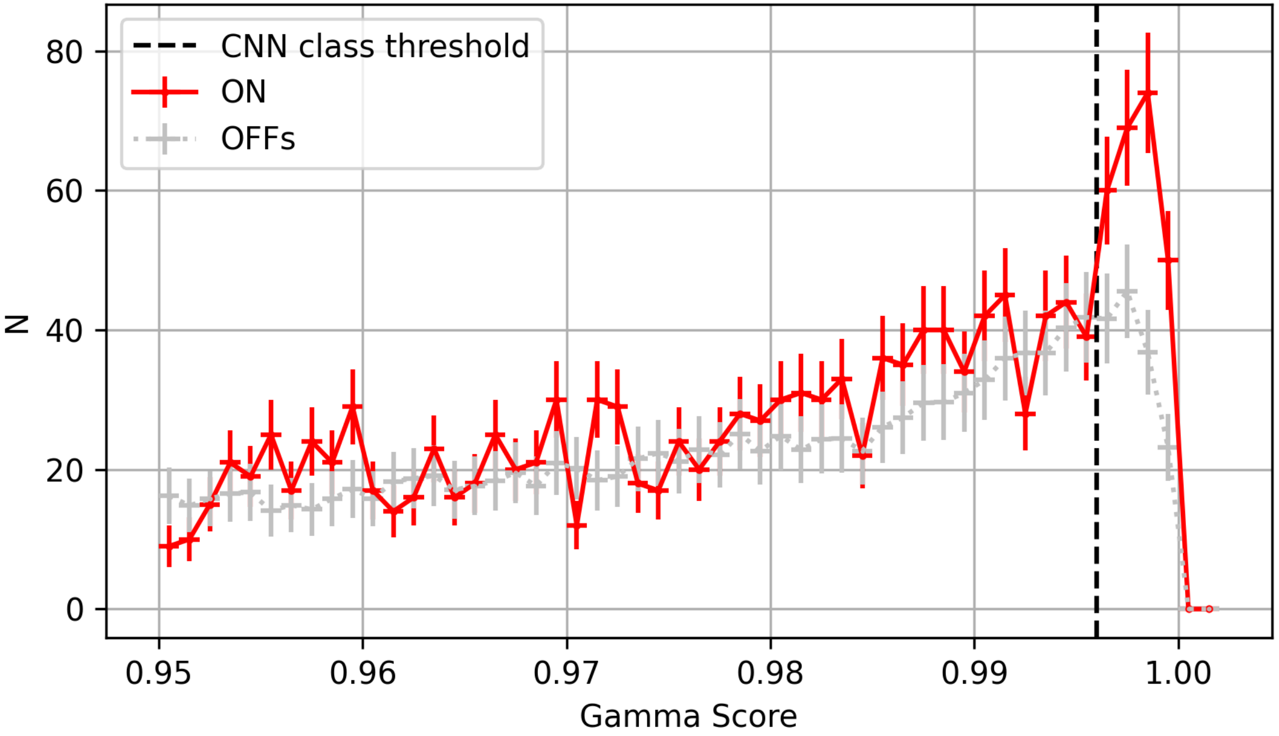
- During this work conditions were obtained and applied for classifying data from Cherenkov telescopes when strong class imbalance occurred.
- Even though the complete separation was not achieved, CNN allows to separate gamma-like events on the same level of significance as Hillas Parameters cuts;
- Thus, the signal from Crab Nebula were obtained on 6.6σ level of significance for 21 hours of observation time. 73 gamma events were obtained during this period of time.

- **Perspectives**:
  - To process the more experimental data (the whole 2019-2020 season and others);
  - To develop and test neural network methods for energy reconstruction and energy spectrum of observed by TAIGA-IACT sources;
  - To process experimental data of different observe gamma sources with neural network methods;

THANK YOU FOR ATTENTION!

BACKGROUND SLIDES

- CNN with Alpha<20°



- CNN with Alpha<6°