



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



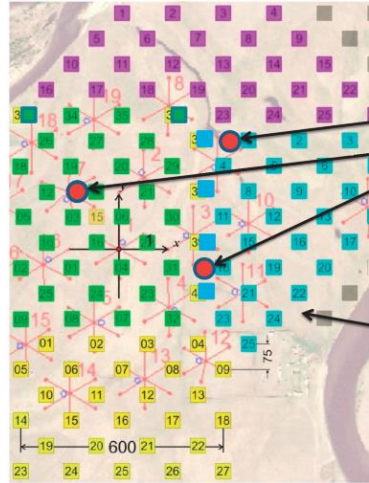
GAMMA/HADRON SEPARATION IN THE TAIGA EXPERIMENT WITH NEURAL NETWORK METHODS

Elizaveta Gres¹, A. Kryukov², P. Volchugov², A. Demichev², D. Zhurov¹,
Ju. Dubenskaya², S. Polyakov², S. Polyakov², A. Vlaskina²

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.



$51^{\circ} 48' 35''$ N
 $103^{\circ} 04' 02''$ E
675 m a.s.l.



PROBLEM AND AIM OF THE RESEARCH

- TAIGA-IACT detect gamma events 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 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

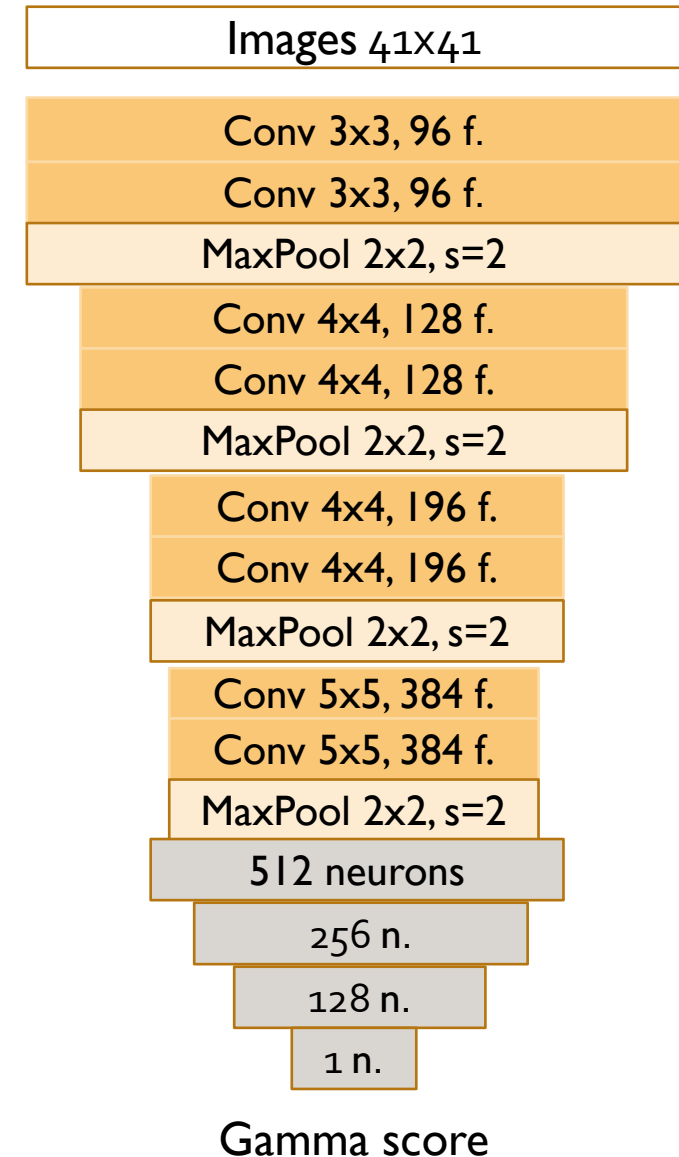


IMAGE PREPROCESSING

- 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 (ON) to another (opposite) offset position (OFF)

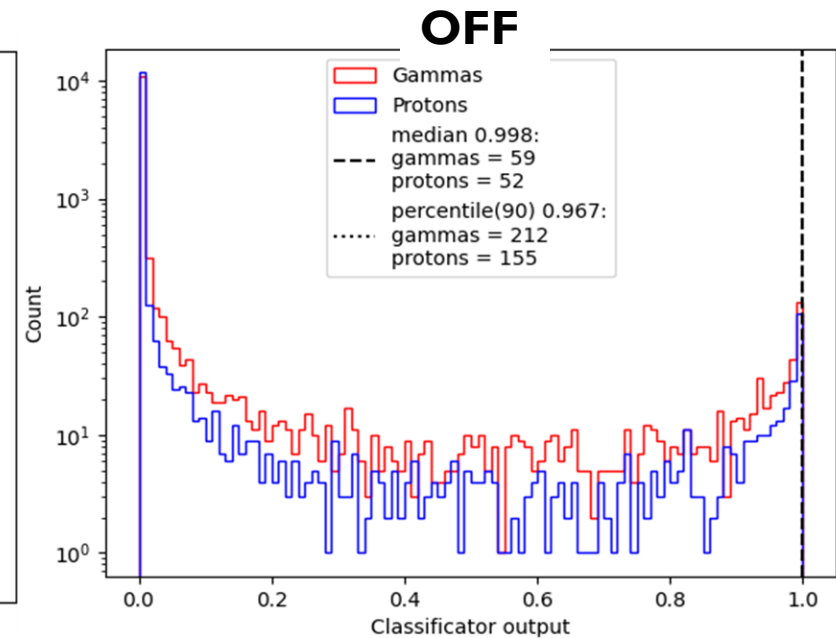
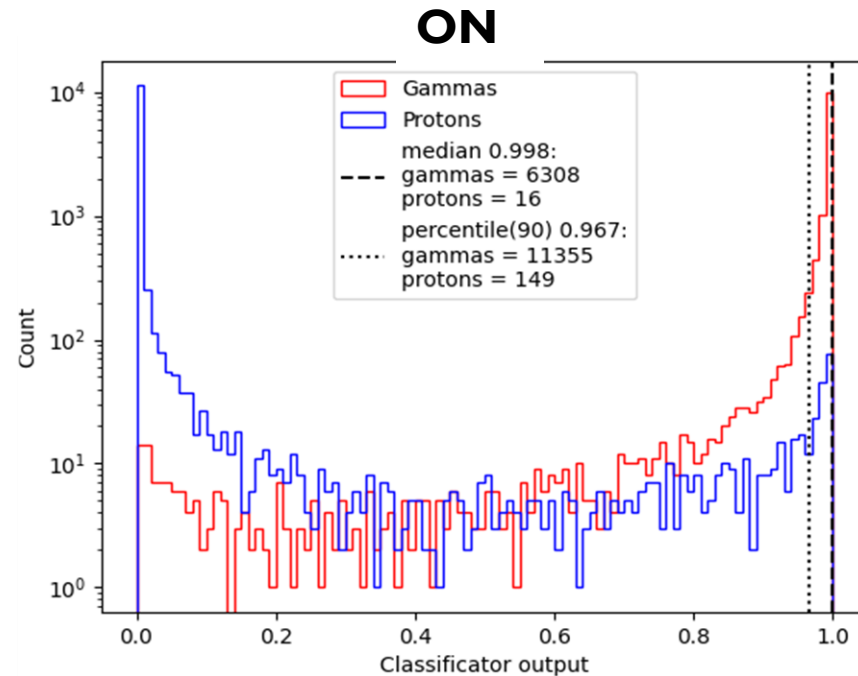
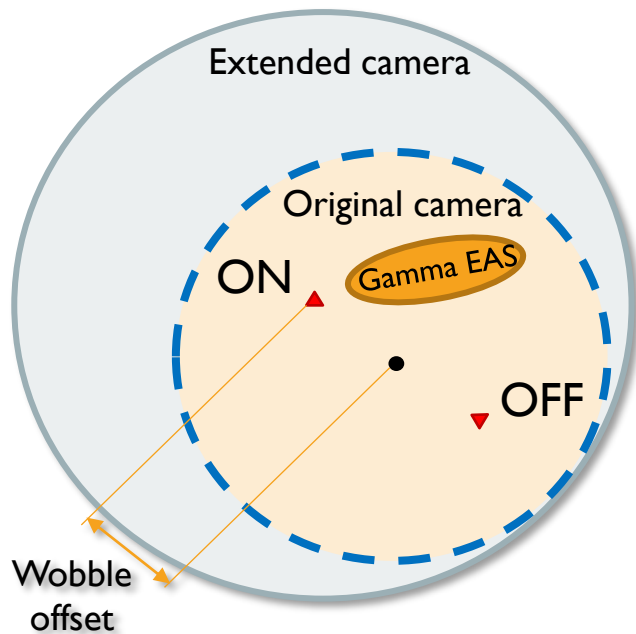
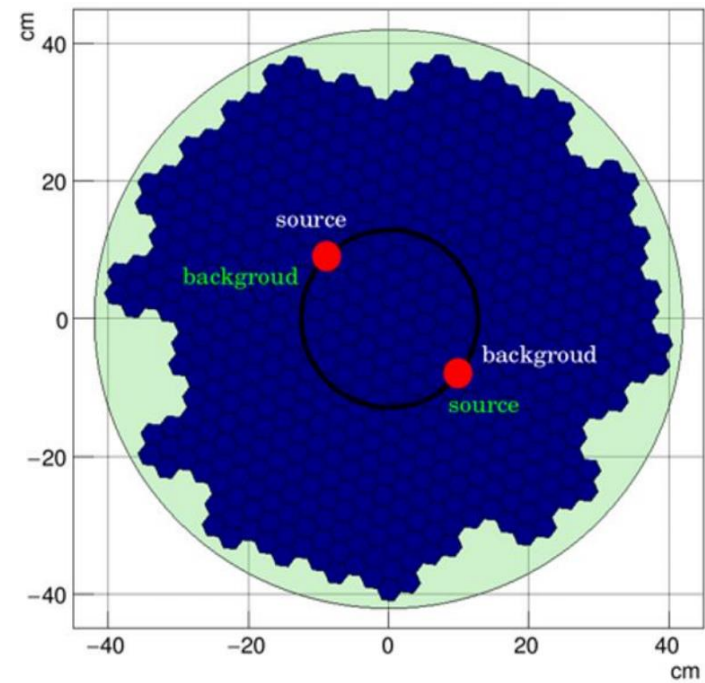
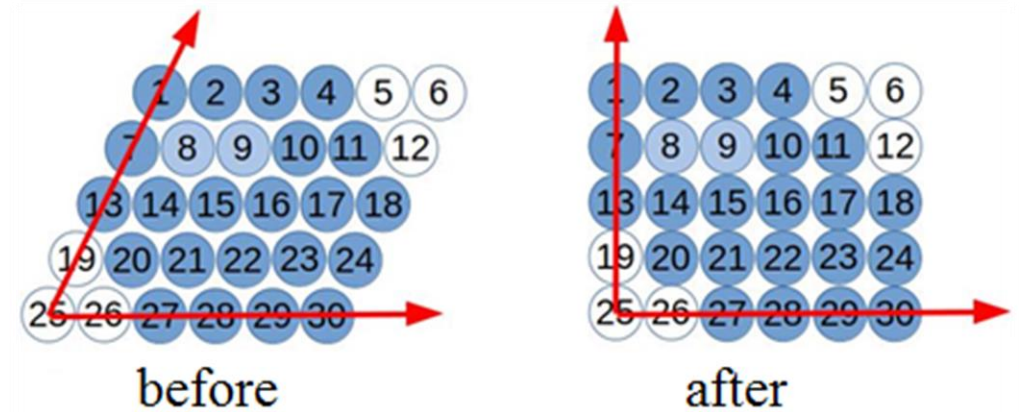
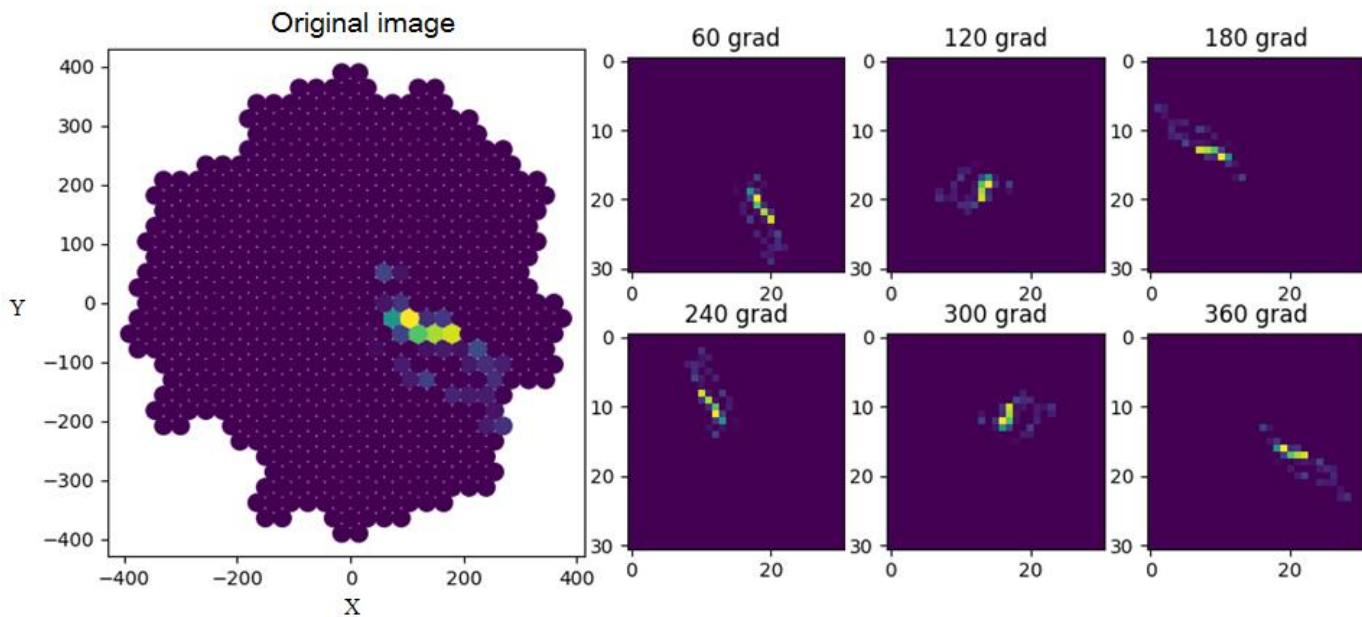


IMAGE PREPROCESSING

- **III.** Train set expansion by 60° rotation;
- **IV.** Transformation to a square image with the axial method;
- **V.** Logarithmic amplitude normalization to [0;1]:

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



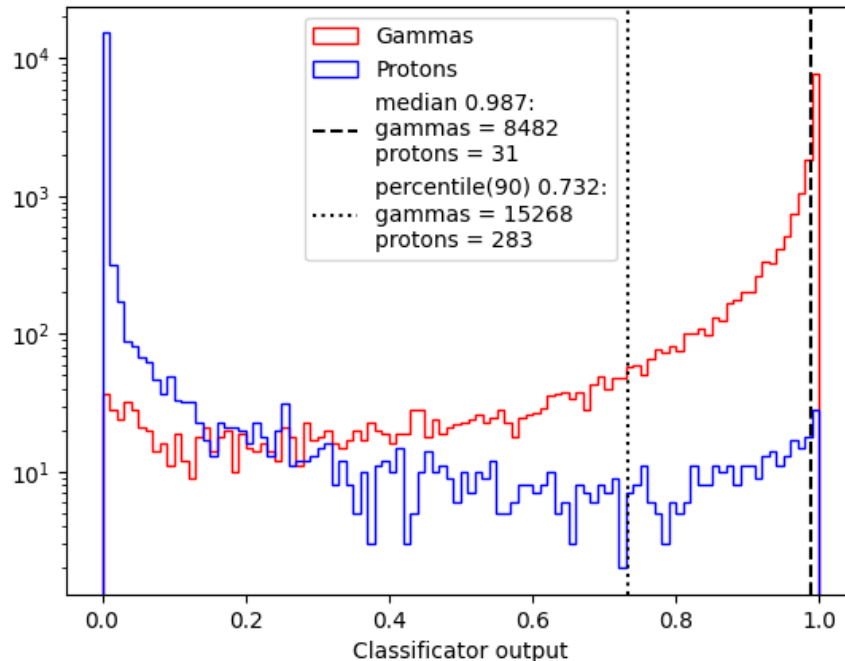
CNN TRAINING AND VALIDATION

- Training:

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

- Validation:

- Gamma: 16.9 thousands
- Hadrons: 17.3 thousands



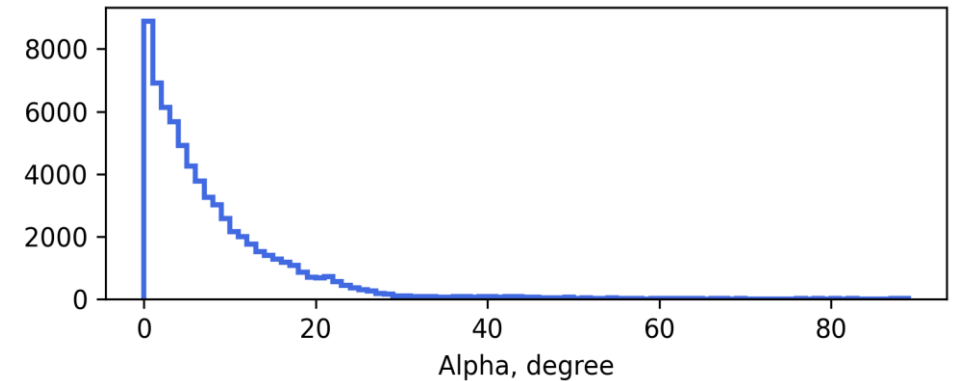
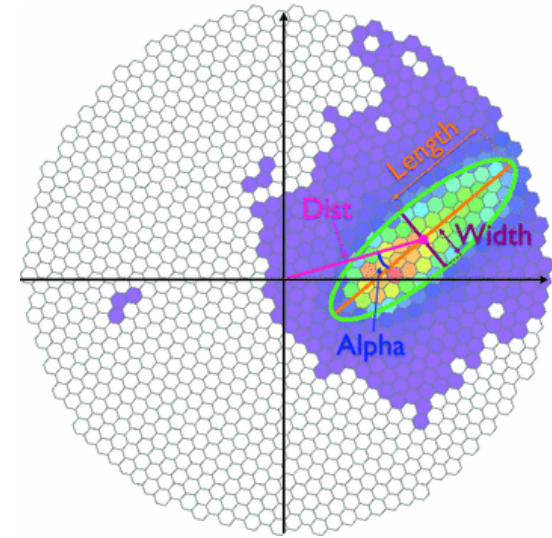
At class threshold 0.987:

- 8482 gamma events
- 31 hadrons

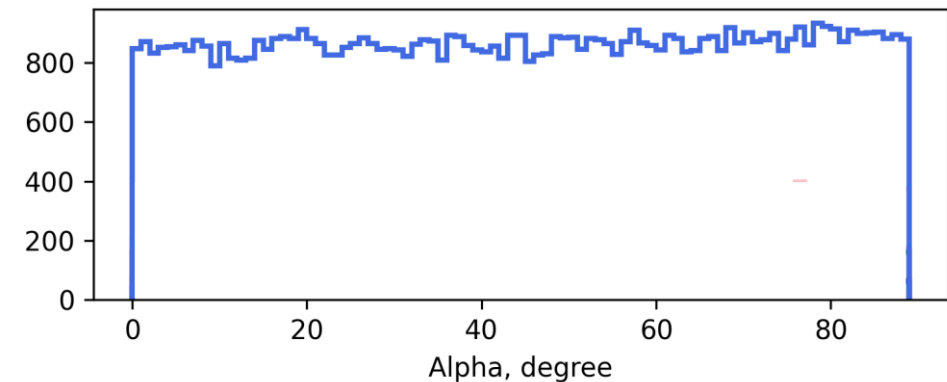
- Thus, the background suppression at the 50% gamma loss threshold is **no more than 1000.**

CNN TRAINING AND VALIDATION

- **Conditions for events:** $\lg(\text{Size}) > 2.08$, $\text{Alpha} < 20^\circ$ or $\text{Alpha} < 6^\circ$
 - We are not interested in images severely distorted by cleaning;
 - Remains:
 - $\text{Alpha} < 20^\circ$: gamma $\sim 90\%$, hadrons $\sim 20\%$
 - $\text{Alpha} < 6^\circ$: gamma $\sim 50\%$, hadrons $< 10\%$
 - Significantly reduces the number of events and speeds up data processing.
- Hence, training with $\text{Alpha} < 20^\circ$ and $\lg(\text{Size}) > 2.08$:
 - MC gamma: 38400 (1-200 TeV, 30-40° zenith)
 - Hadrons: 40000



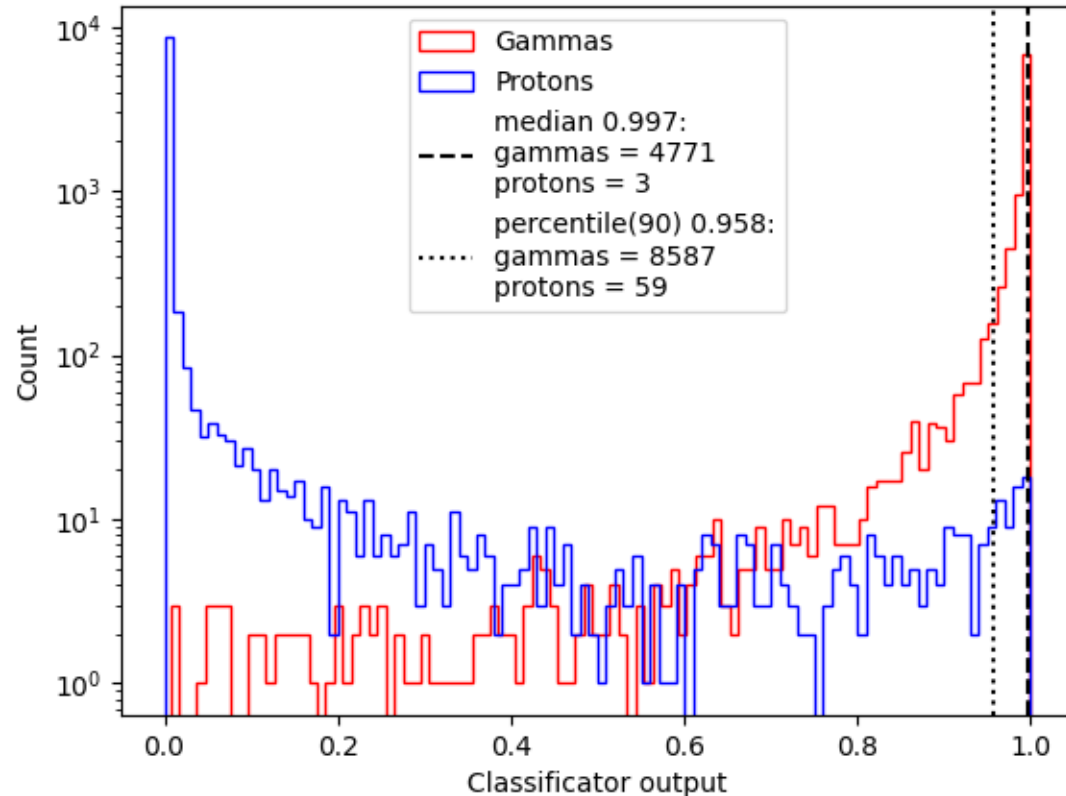
Gamma events



Proton events

CNN TRAINING AND VALIDATION

Validation: 9500 MC gamma и 9600 hadron events



At class threshold 0.9965:

- 4771 gamma events
- 3 hadrons

Background suppression for validation set is up to 1000, leading the general hadron suppression to **10000**.

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

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

- HP cuts:

$$\begin{aligned} & \text{Size} > 120 \text{ phe}; 0.36^\circ < \text{Dist} < 1.44^\circ; \\ & 0.024^\circ < \text{Width} < 0.068^\circ \cdot \lg(\text{Size}) - 0.047^\circ; \\ & \text{Length} > 0.145^\circ \cdot \lg(\text{Size}); \Theta^2 < 0.05^\circ \end{aligned}$$

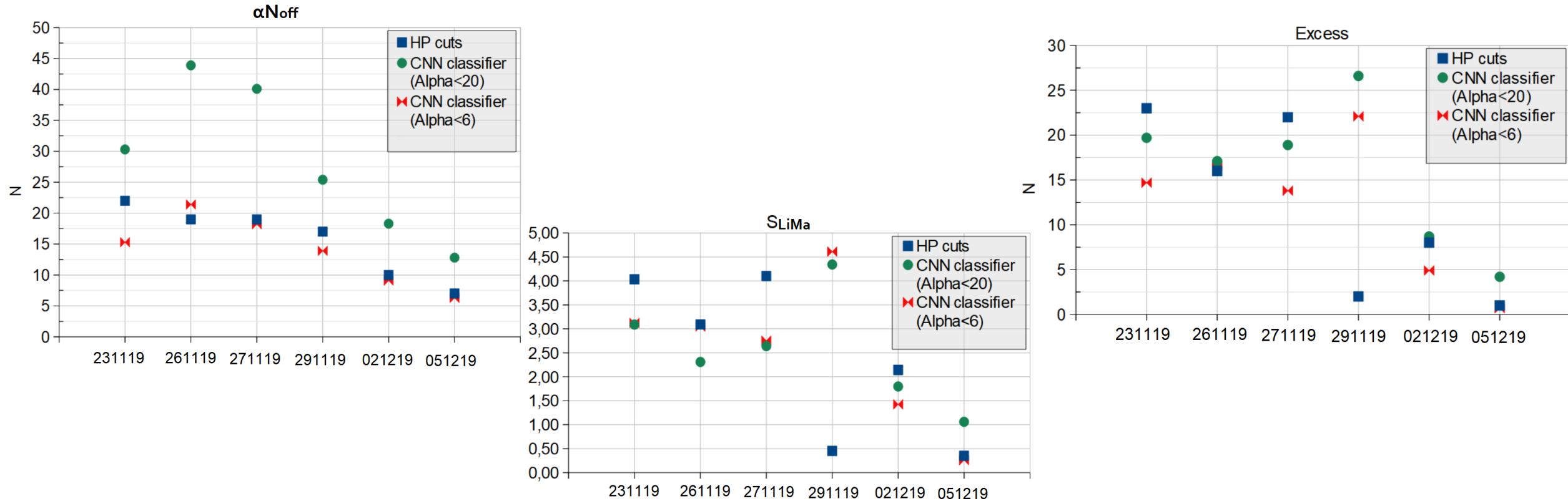
- Significance S_{LiMa} (α – time ratio of ON- and OFF-observations):

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

Observation time, hours	α	Processing method	αN_{OFF}	N_{ON}	Excess	S_{LiMa}
21,75	0.1	HP cuts	94	166	72	6.32
		CNN (Alpha<20)	170.8	266	95.2	6.37
		CNN (Alpha<6)	84.2	157	72.8	6.67

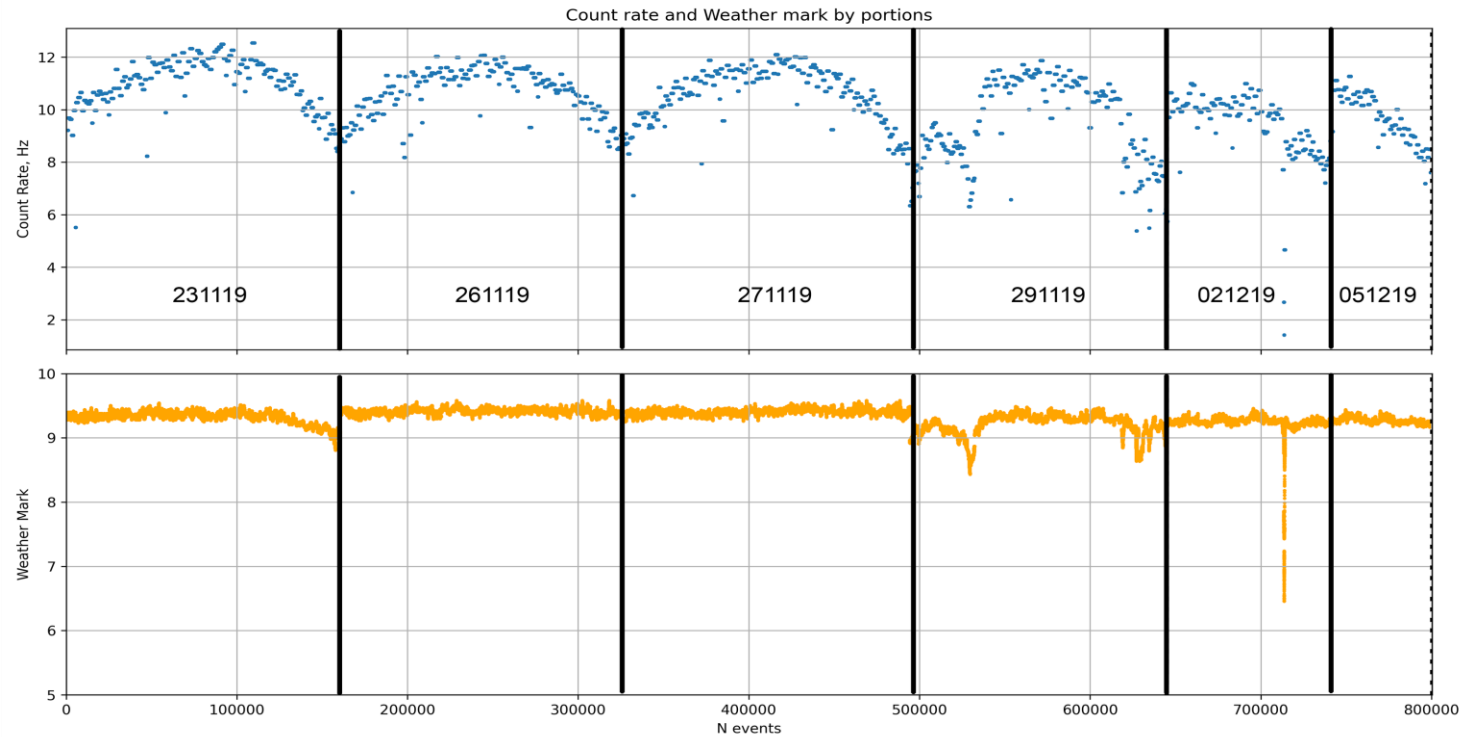
- Coincidence – 46 events in ON-point (CNN (Alpha<20)).

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



- 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 METHOD

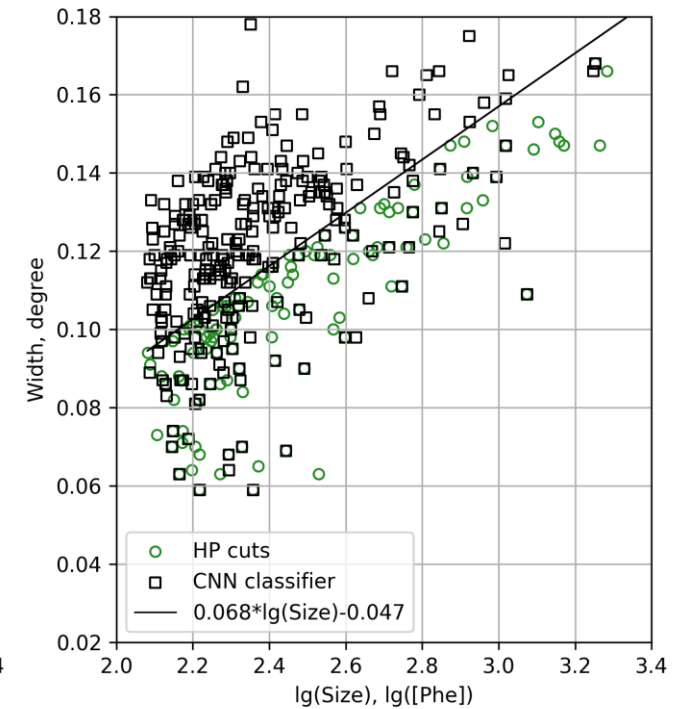
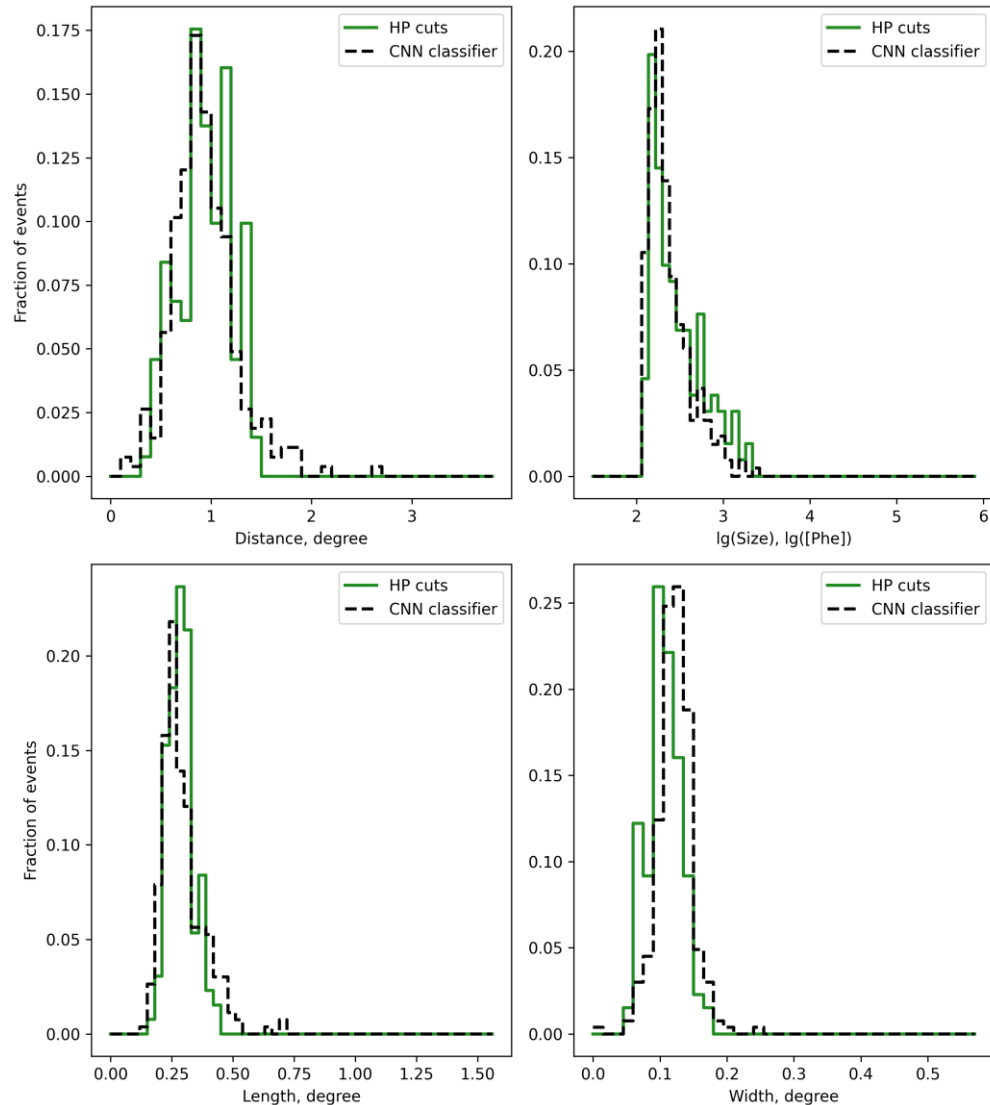


- To avoid the bad weather, criteria: count rate of EAS > 8 Hz; weather mark > 9

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 METHOD

Hillas parameters distributions for ON point



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

CONCLUSION

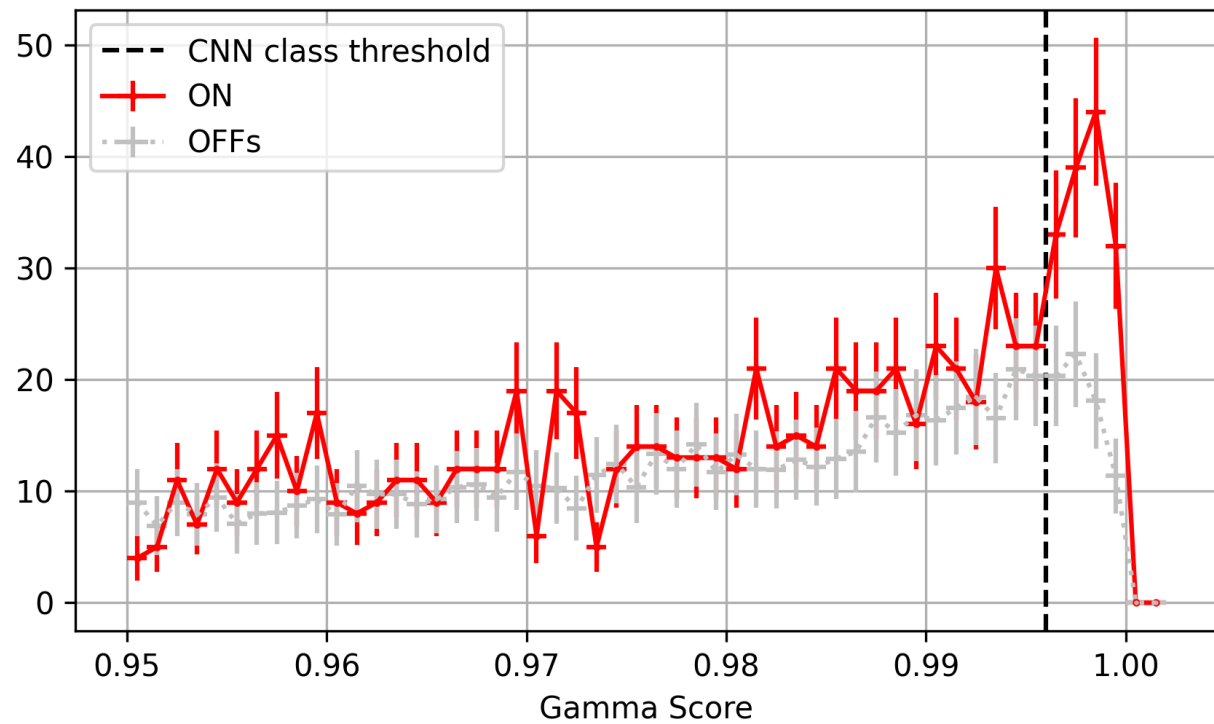
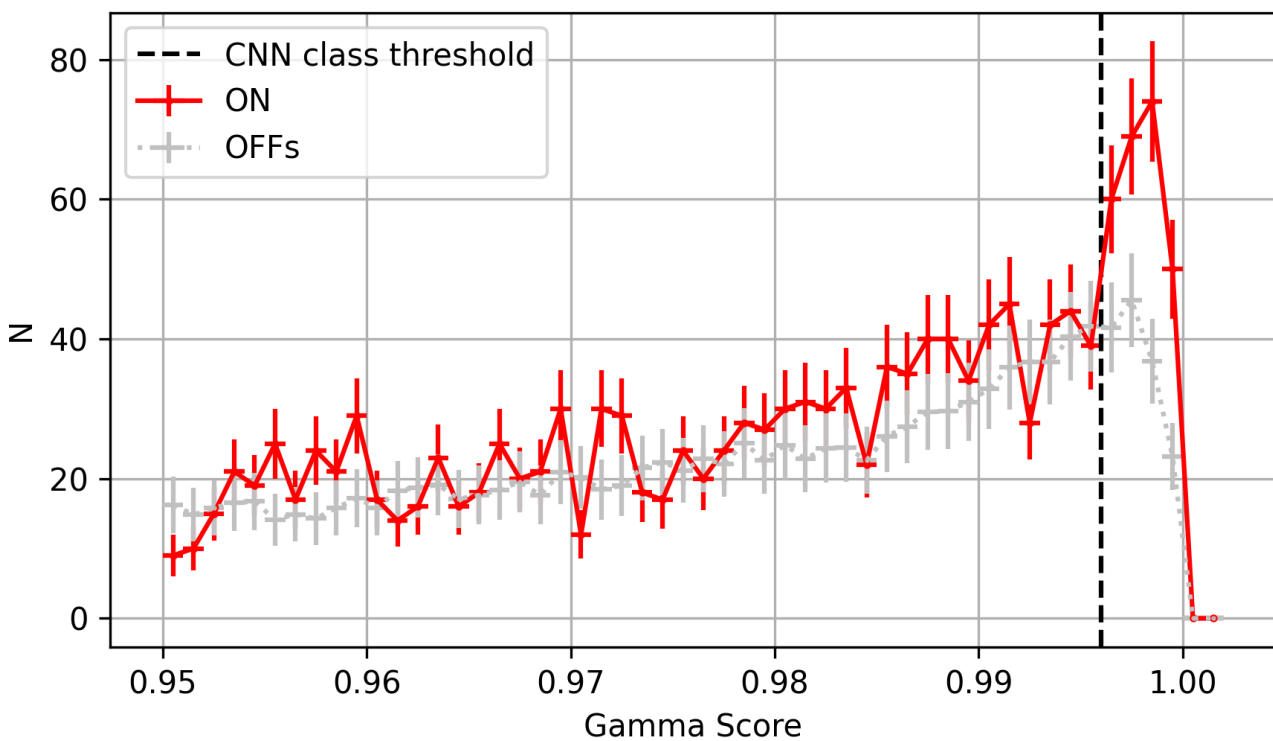
- During this work selection conditions were obtained and applied for classifying data from Cherenkov telescopes when strong class imbalance occurred.
- Even though the complete separation was not achieved, CNN allows to separate gamma-like events on the same level of significance as Hillas Parameters cuts;
 - Thus, the signal from Crab Nebula were obtained on 6.2σ level of significance for 21 hours of observation time. 68 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 Alpha20°

• CNN with Alpha6°



Отбор	Выборка 1-50 ТэВ. Число оставшихся событий
После $\lg(\text{size}) > 2.08$	5111
После $\lg(\text{size}) > 2.08$ & $\alpha < 20$	4884 (95%)
После $\lg(\text{size}) > 2.08$ & $\alpha < 6$	3212 (62%)
HP cuts	1290 (25%)
CNN_alpha < 20	2010 (39%)
CNN_alpha < 6	1310 (25%)
Пересечение HP & CNN_alpha < 20	984