

The use of convolutional neural networks for processing images from multiple IACTs in the TAIGA experiment

Stanislav Polyakov,* Andrey Demichev, Alexander Kryukov and Evgeny Postnikov

*Lomonosov Moscow State University, Skobeltsyn Institute of Nuclear Physics (SINP MSU),
1(2) Leninskie gory, Moscow 119991, Russian Federation
E-mail: s.p.polyakov@gmail.com, demichev@theory.sinp.msu.ru,
kryukov@theory.sinp.msu.ru, evgeny.post@gmail.com*

TAIGA experiment uses hybrid detection system for cosmic and gamma rays that currently includes three imaging atmospheric Cherenkov telescopes (IACTs). Previously we used convolutional neural networks to identify gamma ray events and estimate the energy of the gamma rays based on an image from a single telescope. Subsequently we adapted these techniques to use data from multiple telescopes, increasing the quality of selection and the accuracy of estimates. All the results have been obtained with the simulated data of TAIGA Monte Carlo software.

*37th International Cosmic Ray Conference (ICRC 2021)
July 12th – 23rd, 2021
Online – Berlin, Germany*

*Presenter

1. Introduction

Imaging atmospheric Cherenkov telescopes (IACTs) can detect Cherenkov radiation produced by cascades of particles originating from a collision of a high-energy particle with upper atmosphere, known as extensive air showers. Individual IACTs and IACT arrays are used in numerous gamma-ray observatories such as MAGIC, H.E.S.S., VERITAS. They are also used along with other types of detectors in the TAIGA (Tunka Advanced Instrument for cosmic ray physics and Gamma-ray Astronomy) experiment where a second Cherenkov telescope was commissioned and a third one installed in 2020 [1].

One of the main problems faced by scientists using IACT images is separating the images of events caused by gamma rays from the more common images of events caused by other types of primary particles such as protons. A conventional method based on image parameters (so-called Hillas parameters) was originally proposed in [2]. Following the breakthroughs in applying machine learning to image recognition machine learning methods also became commonly used in gamma astronomy. In particular, the very successful convolutional neural networks (CNNs) were applied to the analysis of IACT images in multiple experiments including VERITAS [3], CTA [4], and H.E.S.S. [5]. In our previous work we used CNNs to identify the primary particle types and estimate the energy of the gamma rays using images from a single TAIGA telescope [6, 7].

In this paper we extend the use of convolutional neural networks for the analysis of TAIGA IACT images to stereoscopic mode. We compare the performance of the neural networks trained on images from one and two telescopes in identifying images corresponding to gamma events and estimating the energy of the gamma rays.

2. Image classification

The neural networks were trained and evaluated using images generated by a Monte Carlo simulation software CORSIKA [8]. The data set for the networks identifying the primary particle type consisted of stereoscopic images of 3400 gamma events and 9306 proton events with the energy from 1 to 45 TeV, obtained by two telescopes. The distance between the telescopes was 300–350 m. Figures 1 and 2 show the examples of a gamma and a proton event images, respectively.

We used the following neural network architecture:

Conv2D 5×5 , W
 AvgPool 2×2
 Conv2D 5×5 , W
 AvgPool 2×2
 Conv2D 3×3 , W
 AvgPool 2×2
 Flatten $3 \times 3 \times W \rightarrow 9W$
 Fully connected layer, $3W$
 Fully connected layer, W
 Output layer, 2

with $W = 10, 15, 25, 50, 100, 150, 250$ and either one or two images as input. PyTorch library was used to implement the networks.

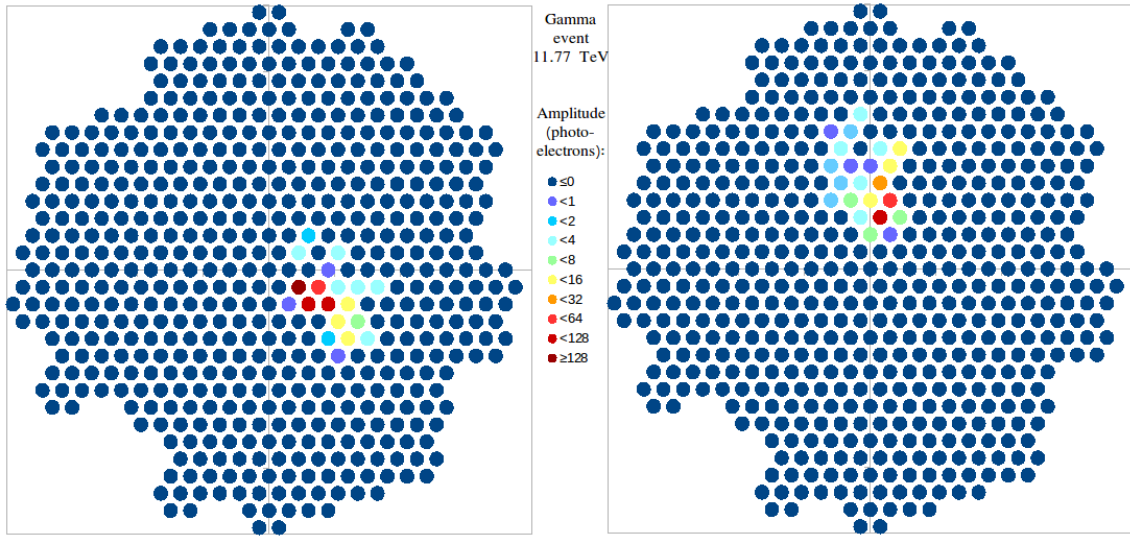


Figure 1: Stereoscopic TAIGA IACT image of a gamma event, $E = 11.77$ TeV.

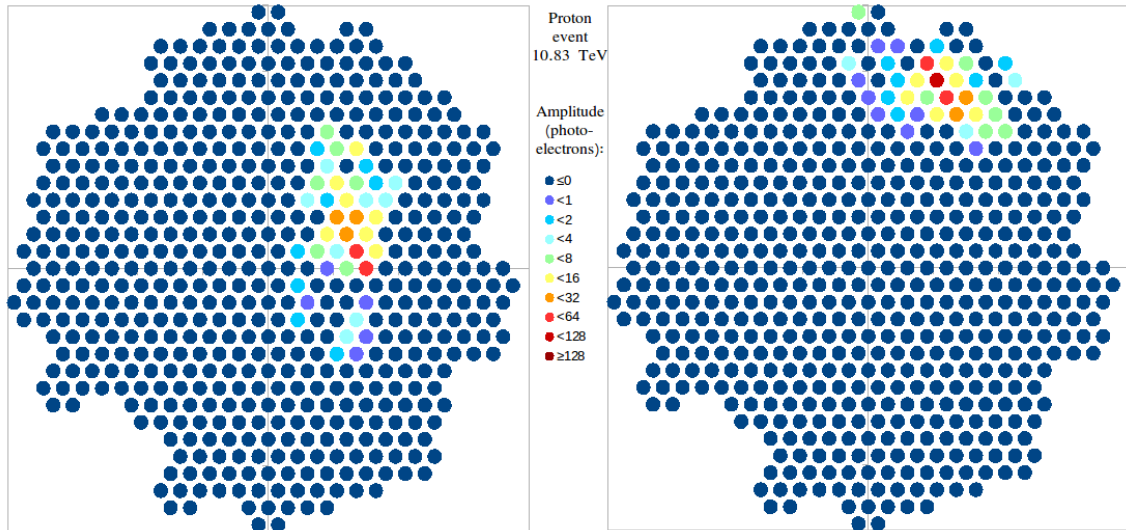


Figure 2: Stereoscopic TAIGA IACT image of a proton event, $E = 10.83$ TeV.

All neural networks were trained on 80% of the data set and evaluated using the remaining 20%, which was repeated 10 times with different training subsets selected at random. The networks were trained for 500 epochs with dropout ($p = 0.5$) applied before each fully connected layer, including the output layer. The learning rate for gradient descent was initially set at 0.1 and decreased by a factor of 10 if the average loss did not reach a new minimum for 20 consecutive epochs.

The networks trained on stereoscopic images showed significantly improved performance. One possible measure of performance is the hadron suppression rate (the true positive rate to false positive rate ratio) with a condition that the fraction of correctly identified gamma events (true positive rate) is kept above 1/2. For the neural networks trained on monoscopic images the average hadron suppression rate was 74–98, and for the networks trained on stereoscopic images it was

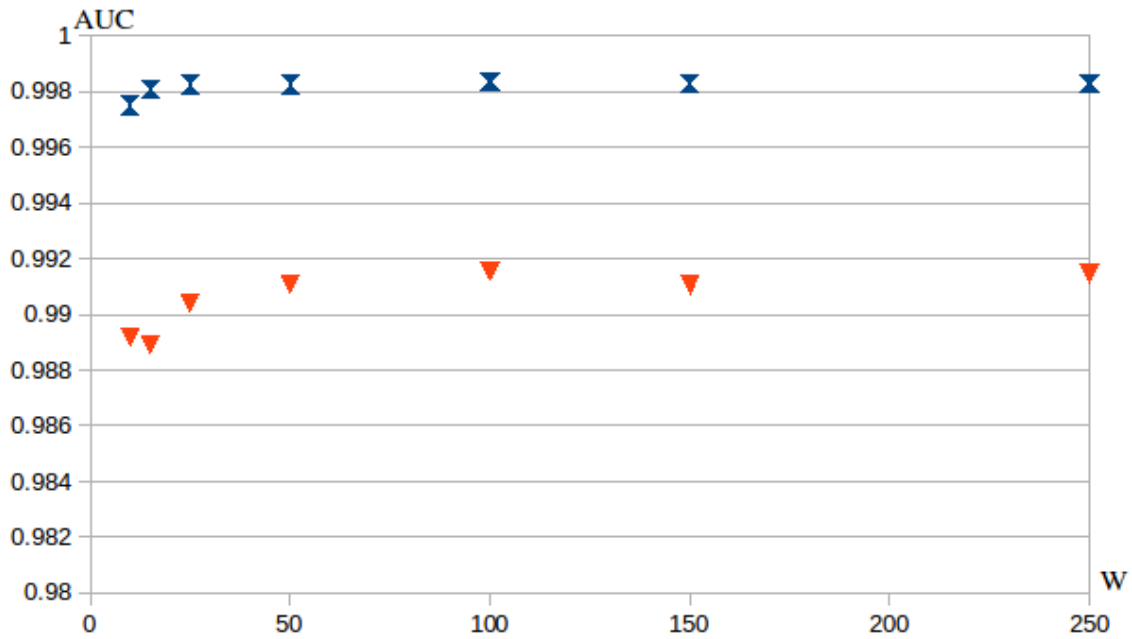


Figure 3: Average ROC AUC for the CNNs trained on monoscopic (triangles) and stereoscopic (double triangles) images.

340–454. For the similar networks adding a second input image increased the hadron suppression rate by a factor of 4.0–5.4.

We also compared the average area under a receiver operating characteristic curve (ROC AUC) for these networks. For the networks trained on monoscopic images the average AUC was between 0.9889 and 0.9916, and for the networks trained on stereoscopic images it was between 0.9975 and 0.9983 (fig. 3).

3. Estimation of gamma ray energy

The data set for the networks estimating the gamma ray energy was generated using CORSIKA and consisted of stereoscopic images of 18359 gamma events with the energy from 1 to 50 TeV. The images were obtained by two telescopes at a distance of 324 m from each other.

We used the following neural network architecture:

Conv2D 5×5 , $\lfloor W/4 \rfloor$
 AvgPool 2×2
 Conv2D 5×5 , $\lfloor W/2 \rfloor$
 AvgPool 2×2
 Conv2D 3×3 , $\lfloor W/2 \rfloor$
 AvgPool 2×2
 Flatten $3 \times 3 \times \lfloor W/2 \rfloor \rightarrow 9 \lfloor W/2 \rfloor$
 Fully connected layer, W
 Fully connected layer, W
 Fully connected layer, W
 Output layer, 1

with $W = 12, 25, 50, 100, 200$ and either one or two images as input.

All neural networks were trained on 80% of the data set and evaluated using the remaining 20%, which was repeated 10 times with different training subsets selected at random. The dropout with a rate p was applied during training after each layer except for the output, with $p = \frac{k}{40}$, $k = 1, \dots, 20$. We chose the value of p that minimized the average relative error. The learning rate for gradient descent was initially set at 0.1 and decreased by a factor of 10 if the average loss did not reach a new minimum for 20 consecutive epochs. The training was stopped after the learning rate decreased below 0.001 (this always required less than 2000 epochs).

The average relative error of the energy estimates was 24.0% to 26.0% for the networks trained on monoscopic images and 12.5% to 14.0% for the networks trained on stereoscopic images. For the similar networks adding a second input image decreased the average relative error by a factor of 1.7–2.0 (fig. 4).

We also trained the neural networks for monoscopic and stereoscopic images with $W = 100$ using the gamma events from the first data set (described in section 2). The average relative error of the energy estimates was 22.0% for monoscopic images and 16.0% for stereoscopic images.

4. Conclusion

We compared the performance of convolutional neural networks trained on simulated monoscopic and stereoscopic images from TAIGA telescopes. We found that adding a second telescope image significantly improved both the accuracy of the primary particle type identification (hadron suppression rate increased from 74–98 to 340–454, ROC AUC increased from 0.9889–0.9916 to 0.9975–0.9983) and the accuracy of the gamma ray estimation (average relative error decreased from 24.0%–26.0% to 12.5%–14.0% and 22.0% to 16.0% on the two data sets we used).

Acknowledgments

This work was carried out in the framework of R&D State Assignment No.115041410196.

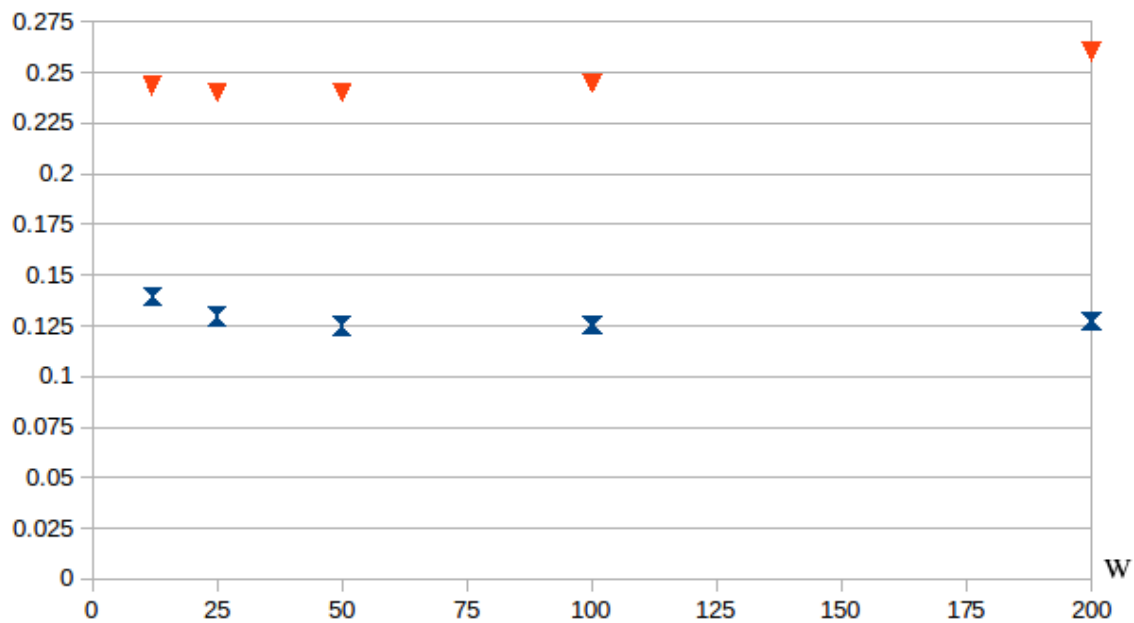


Figure 4: Average relative error of energy estimation by neural networks trained on monoscopic (triangles) and stereoscopic (double triangles) images.

References

- [1] E.B. Postnikov, I.I. Astapov, P.A. Bezyazeev, M. Blank and A.N. Borodin, *First detection of gamma-ray sources at tev energies with the first imaging air cherenkov telescope of the taiga installation*, *J. Phys. Conf. Ser.* **1690** (2020) 012023.
- [2] A.M. Hillas, *Cerenkov Light Images of EAS Produced by Primary Gamma Rays and by Nuclei*, in *19th International Cosmic Ray Conference (ICRC19)*, vol. 3 of *International Cosmic Ray Conference*, pp. 445–448, Aug., 1985.
- [3] Q. Feng, J. Jarvis, VERITAS Collaboration et al., *A citizen-science approach to muon events in imaging atmospheric cherenkov telescope data: the muon hunter*, in *35th International Cosmic Ray Conference (ICRC2017)*, vol. 301, 2017.
- [4] S. Mangano, C. Delgado, M.I. Bernardos, M. Lallena, J.J.R. Vázquez, CTA Consortium et al., *Extracting gamma-ray information from images with convolutional neural network methods on simulated cherenkov telescope array data*, in *IAPR Workshop on Artificial Neural Networks in Pattern Recognition*, pp. 243–254, Springer, 2018.
- [5] I. Shilon, M. Kraus, M. Büchele, K. Egberts, T. Fischer, T. Holch et al., *Application of deep learning methods to analysis of imaging atmospheric cherenkov telescopes data*, *Astroparticle Physics* **105** (2019) 44.
- [6] E. Postnikov, A. Kryukov, S. Polyakov, D. Shipilov and D. Zhurov, *Gamma/hadron separation in imaging air cherenkov telescopes using deep learning libraries tensorflow and pytorch*, *J. Phys. Conf. Ser.* **1181** (2019) 012048.

- [7] E. Postnikov, A. Kryukov, S. Polyakov and D. Zhurov, *Deep learning for energy estimation and particle identification in gamma-ray astronomy*, vol. 2406, pp. 90–99, 2019.
- [8] D. Heck, J. Knapp, J.N. Capdevielle, G. Schatz and T. Thouw, *Corsika: A monte carlo code to simulate extensive air showers*, Tech. Rep. FZKA-6019, Karlsruhe (2, 1998).