THE USE OF CNN FOR IMAGE ANALYSIS FROM CHERENKOV TELESCOPES IN THE TAIGA EXPERIMENT

A. Kryukov[1], D. Zhutov[2], E. Postnikov[1], S. Polyakov[1]

[1] *M.V.Lomonosov Moscow State University, Skobeltsyn Institute of Nuclear Physics. Leninskie gory, 1, bld.2, Moscow, 119992, Russia*

[2] *Irkusk State University, Karl Marks, 1, Irkuts, 664003, Russia*

E-mail: kryukov@theory.sinp.msu.ru

Artificial neural networks is a modern powerful tool for solving various problems in many areas. In particular, they are excellent for various aspects of image analysis because of their ability to find patterns which are too complex or numerous for a human researcher to extract and teach the machine how to recognize them. This paper describes the use of convolutional neural networks (CNN) for the problems of classifying the type of primary particles and estimating their energy using images obtained from the Cherenkov telescope (IACT) in the TAIGA experiment. For the problem of classifying primary particles, it was shown that the use of CNN significantly improved the quality criterion for the correct classification of gammas compared to traditional methods using the Hillas parameters. For the problem of estimating the energy of primary gammas, the use of CNN allowed us to obtain good results for extensive air showers, whose centers are located far enough from the telescope. In particular, it is important for the Cherenkov telescope in the TAIGA experiment, which uses a wide-angle camera when traditional methods do not work. Our neural network was implemented using the PyTorch and TensorFlow libraries. Monte Carlo event sets obtained using the CORSIKA program were used to train the CNN. CNN training was performed on both ordinary servers and servers equipped with Tesla P100 GPUs.

Keywords: Machine Learning, Convutional neural network, TensorFlow, PyTorch, GPU, Astroparticle Physics, Image Air Cherenkov Telescope, TAIGA, Gamma Astronomy

1. Introduction

Artificial neural networks is a modern powerful tool for solving various problems in many areas. In particular, they are excellent for various aspects of image analysis because of their ability to find patterns which are too complex or numerous for a human researcher to extract and teach the machine how to recognize them. Traditionally, artificial neural networks are used to solve various tasks of image analysis, such as classification and regression problem. In the last decade, this method has been increasingly used for the analysis of scientific data. The difference between scientific and traditional data is that scientists often consider an "image" in the multidimensional space of abstract parameters. For example, in gamma astronomy the physicists analyze the "images" of air shower. The main tasks for the analysis are to identify the type of primary particle and determine the parameters of the primary particle, such as its energy.

Very high energy gamma rays produce only one millionth of the cosmic ray flux [1]. Thus, the separation of gamma rays from other cosmic rays, which are mainly protons, is a very important problem. A common way to solve this problem is to use empirical variables, such as Hillas parameters [2]. Successive reductions in the Hillas parameters can be used to remove background events. The optimal cutting values are determined by the Monte Carlo simulation of the telescope images. But now deep learning methods are increasingly used to identify cosmic rays in the IACT images [3,4,5].

This paper describes the use of convolutional neural networks (CNN) for the problems of classifying the type of primary particles and estimating their energy using images obtained from the Cherenkov telescope (IACT) in the TAIGA experiment. For the problem of classifying primary particles, it was shown that the use of CNN significantly improved the quality criterion for the correct classification of gammas compared to traditional methods using the Hillas parameters. For the problem of estimating the energy of primary gammas, the use of CNN allowed us to obtain good results for extensive air showers (EAS), whose centers are located far enough from the telescope. In particular, it is important for the Cherenkov telescope in the TAIGA experiment, which uses a wide-angle camera when traditional methods do not work. Our neural network was implemented using the PyTorch and TensorFlow libraries. Monte Carlo event sets obtained using the CORSIKA program were used to train the CNN. CNN training was performed on both ordinary servers and servers equipped with Tesla P100 GPUs.

The work is a part of the Karlsruhe-Russian astroparticle data life cycle initiative [7]. This initiative aims to develop an open science system for collecting, storing, and analyzing astroparticle physics data. Currently it includes data of TAIGA and KASCADE [8] experiments.

The structure of the article is as follows. In sections 2 and 3 we describe the identification of a gamma event and the determination of the energy of a primary gamma by a CNN. In conclusion, the obtained results and future works are discussed.

2. Particle identification

A CNN is a very good deep learning method for the classification problem. We use a CNN for recognition of the IACT images [9,10]. The advantage of CNN is a fully automatic algorithm, including automatic extraction of image features, in contrast to Hillas parameters which require some preliminary processing to extract them. To build the CNN the free software tools PyTorch [11] and TensorFlow [12] were selected. Since both tools are implemented for square grids, we have to transform the hexagonal shape of the TAIGA-IACT pixels into a regular square grid. For this, we use oblique coordinates with an angle of 60 degrees. Of course, this is only one of the possible ways of such a transformation [9]. As an example, fig. 1 shows the structure of the CNN, which was built using TensorFlow.

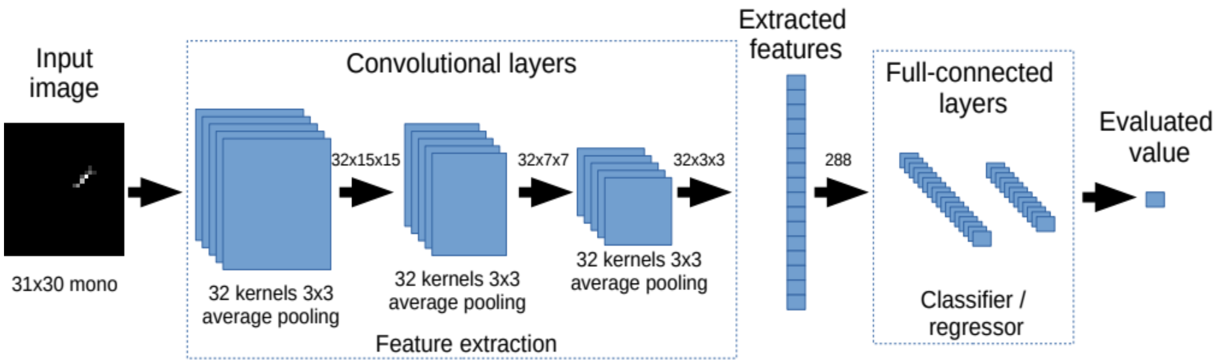

Figure 1. CNN for classification. The output of convolutional layers (extracted features) is fed to a fully connected network, which estimates the output value

The training datasets contained gamma-ray and proton images (Monte Carlo of TAIGA-IACT, energy distributions in the range of 2−60 and 3−100 TeV respectively with the spectral slope of -2.6). The test datasets (different from the training ones) of gamma-ray and proton images in random proportion (blind analysis) were classified by each of the tools: TensorFlow and PyTorch. The simulation dataset is a collection of the TAIGA-IACT images. Each of them consists of 560 pixels arranged in the form of a hexagonal lattice. The simulating dataset consists of 30,273 images generated from protons and 25,492 images generated from high-energy gamma-quanta (55,765 EAS images in total). The initial sample of 55,765 images was randomly divided into three subsamples: 60% of the original for training, 15% for verification (validation) and 25% for the final test. It was also found out that expanding the training sample by rotating the images at angles that are multiple of 60 degrees (the angle of symmetry of the camera) can improve the quality of classification on the test and verification samples. Thus, using image rotations, he sample for training was expanded from 32614 to 195684 events.

The accuracy of identification of the type of primary particles for the best network configuration in the training and test sample after training is 91.29% and 90.02%, respectively. The ROC AUC score was 0.9712 for the training and 0.9647 for the test sample. The quality factor (Q-factor) was also used to evaluate the quality of gamma events detection in the test sample. This factor

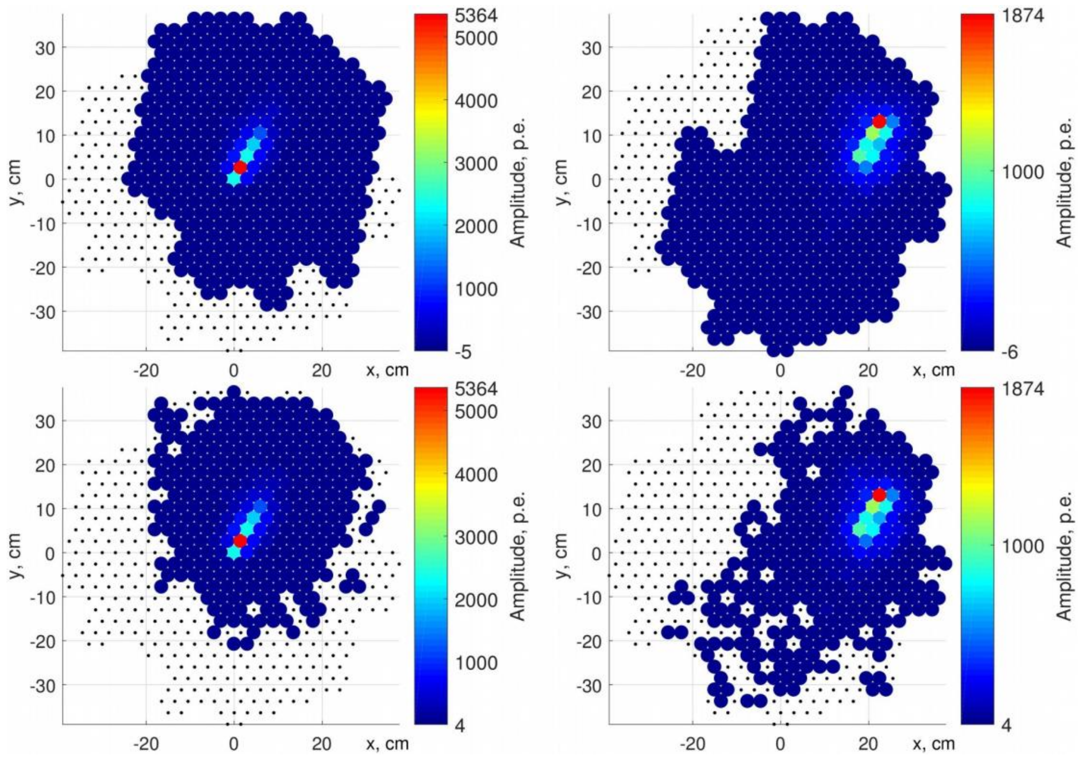

Figure 2. Gamma-ray (left panel) and proton (right panel) images before image cleaning (top panel) and after image cleaning with a low threshold (bottom panel)

indicates an improvement in the significance of the statistical hypothesis that events do not belong to the background compared with the value before the selection of events. The value of the Q-factor obtained using the best CNN configuration among all trained networks at the optimal threshold was 2.99. The classical technique using the two main Hillas parameters that determine the linear dimensions of the image and its orientation to the source, under the same conditions, provides a Q-factor of 1.70. It should be kept in mind that these results were obtained with poor image cleaning and without any other selection criteria, for example, without selection criteria for the total number of photoelectrons in the image (the so-called image size) (see fig.2). After applying image size sampling over 60 photoelectrons, the Q-factor reached a value of 4.10 for the convolution network, while for the Hillas parameters it was only 2.76.

When training CNN to identify the type of primary particle, the Adagrad optimizer was used

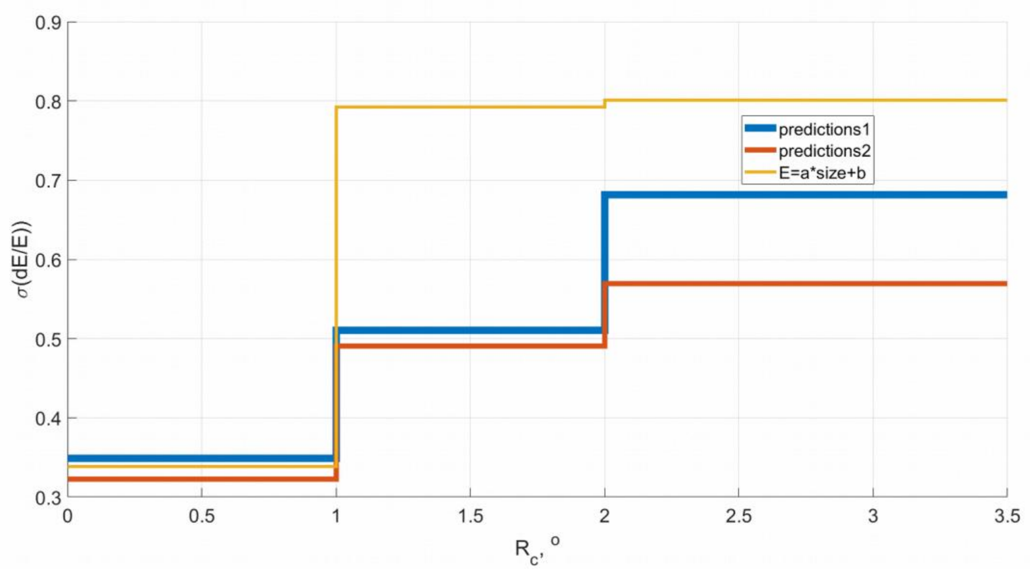

Figure 3. Precision of energy estimation

with a set learning speed of 0.05 and binary cross-entropy as a loss function. The training was carried out using the early stopping criterion to interrupt the training procedure, when the loss function for the test sample did not decrease over 30 epochs. The training lasts ~ 150 eras (duration ~ 9 minutes). For calculations, we used the NVIDIA Tesla P100 graphics processor, which allowed us to increase the performance in calculations by ~ 11 times compared to using an ordinary CPU.

3. Estimation of energy of primary gamma's

Conventional method of energy prediction is based on linear correlation between the energy and the image size (the sum of photo-electrons over the image), which works only for gamma-rays incident very close to the telescope (up to 100-150 m on the ground which equivalently up to ~1 degree on the camera). While our preliminary results did not improve the accuracy of estimating the energy of showers, the use of CNN gives good results for showers whose axis is rather far from the IACT. The fig. 3 show that the energy error obtained by CNN (blue and red curves) is much smaller than the method of image size (orange line) at a distance of 1 to 2 degrees. This region is very important for TAIGA IACT that has wide angle camera and registers many showers more than 1 degree.

4. Conclusion

Convolutional neural network**s** have great potential for data analysis in astroparticle physics and, in particular, for IACT images analysis. The main advantage of the deep learning method over the conventional one is that it does not use difficult motivated heuristics to identify the primary particle. The identification quality factor of this method is twice as high as the quality factor for the method based on Hillas parameters. Another advantage of CNN is the ability to parallelize the calculations performed for the network, which makes these tasks well suited for GPUs, the use of which accelerates calculations by ten or more times. Modern software products such as PyTorch and TensorFlow provide convenient high-level tools for neural network configuration. In particular, to switch from CPU to GPU, one just needs to change several compiler options.

In the future, we plan to explore the capabilities of the generative adversarial networks (GAN) for fast generation of realistic IACT images instead of rather slow MC generator CORSIKA.

The work is supported by RSF 18-41-06003. The authors also express deep gratitude to Yu.Dubenskaya for her help in preparing the articles.

References

[1] Lorenz E, Wagner R, Very-high energy gamma-ray astronomy // The European Physical Journal H, **37**(2012), 459–513. DOI:10.1140/epjh/e2012-30016-x

[2] Hillas A M. Cerenkov Light Images of EAS Produced by Primary Gamma Rays and by Nuclei //In Proc. 19th Int. Cosmic Ray Conf. (La Jolla), 3(1985), 445–448

[3] Feng Q, Lin T T Y for the VERITAS Collaboration. The analysis of VERITAS muon images using convolutional neural networks. Proc. the International Astronomical Union Symposium S325: Sorrento, Italy, October 19-25, 12(2016), 173–179. ArXiv:astro-ph.IM/1611.09832.

[4] Nieto, D.; Brill, A.; Kim, B.; Humensky, T. Exploring deep learning as an event classification method for the Cherenkov Telescope Array // Proceedings of Science, 301(2017), 809. ArXiv:astro-ph.IM/1709.03483.

[5] Shilon I, et al. Application of Deep Learning methods to analysis of Imaging Atmospheric Cherenkov Telescopes data // ArXiv:astro-ph.IM/1803.10698.

[6] Postnikov E et al. Commissioning the joint operation of the wide angle timing HiSCORE Cherenkov array with the first IACT of the TAIGA experiment // Proc. 35th Int. Cosmic Ray Conf., PoS(ICRC2017)756

[7] Igor Bychkov et al. Russian–German Astroparticle Data Life Cycle Initiative // Dat, 3(2018), 56. DOI:10.3390/data3040056

[8] Apel, W.D. et al. The KASCADE-Grande experiment // Nucl.Instrum.Meth. A620 (2010) 202-216

[9] Nieto D. et al. for the CTA Consortium. Exploring deep learning as an event classification method for the Cherenkov Telescope Array // Proceedings of Science 2017. PoS(ICRC2017), 809

[10] Shilon I., et al. Application of Deep Learning methods to analysis of Imaging Atmospheric Cherenkov Telescopes data // arXiv:1803.10698. 2018.

[11] PyTorch software // URL: http://pytorch.org

[12] TensorFlow software // URL: http://www.tensorflow.org