Preliminary results of convolutional neural network models in HiSCORE experiment

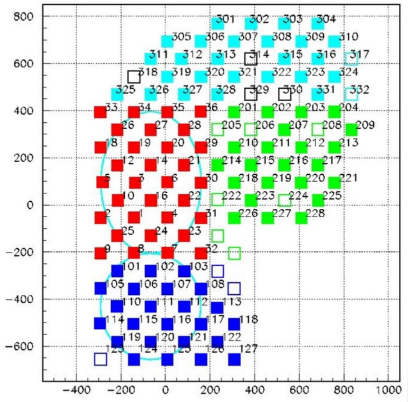
Vlaskina Anna[1], Alexander Kryukov[1] , Andrey Demichev[1], Julia Dubenskaya[1], Stanislav Polyakov[1], Dmitry Zhurov[2], Elizaveta Gres[2]

[1]Skobeltsyn Institute of Nuclear Physics, Moscow State University

[2]Research Institute of Applied Physics, Irkutsk State University

The work was supported by RSF, grant no.22-21-00442.
The work was done using the data of UNU "Astrophysical Complex of MSU-ISU» (agreement EB-075-15-2021-675)

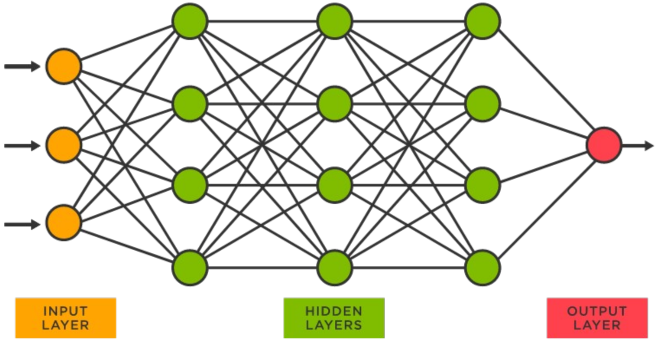
HiSCORE experiment



- Timing array for cosmic ray and high-energy gamma-ray physics
- Part of the TAIGA experiment
- Captures the amplitudes of signal and its time of arrival
- Array is split into 4 clusters
- Energetic threshold: 80 TeV for cosmic rays and 40 TeV for gamma rays

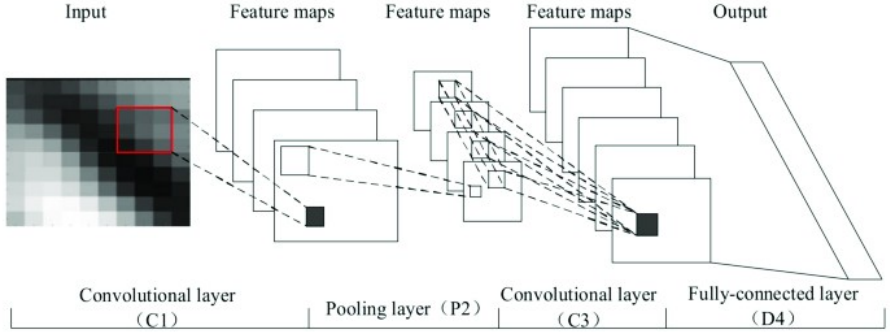
Artificial neural networks

- consists of the simplest interconnected processors - neurons
- each of the neurons of the layer summarizes the information from the previous layer and after the nonlinear transformation passes the result to the next layer.
- network parameters are corrected using backpropagation method



Convolutional neural networks (CNN)

CNNs - type of neural network for data processing with grid topology. The result of the convolution operation is called a feature map. Due to sparse connectivity, a convolutional neural network extracts only significant features.



Proposed methods

Convolutional network

- Array of stations with timing /amplitude data is presented as an image.
- φ and θ of shower axis are extracted from the image as a feature

Fully-connected network

- Time/amplitude data is presented as a function of multiple arguments
- Expected result of the function: φ and θ of shower axis

Convolutional network

Kernel size: 2x2

Number of parameters: 1,665,818

Optimizer: ADAM

Learning rate: 0,001

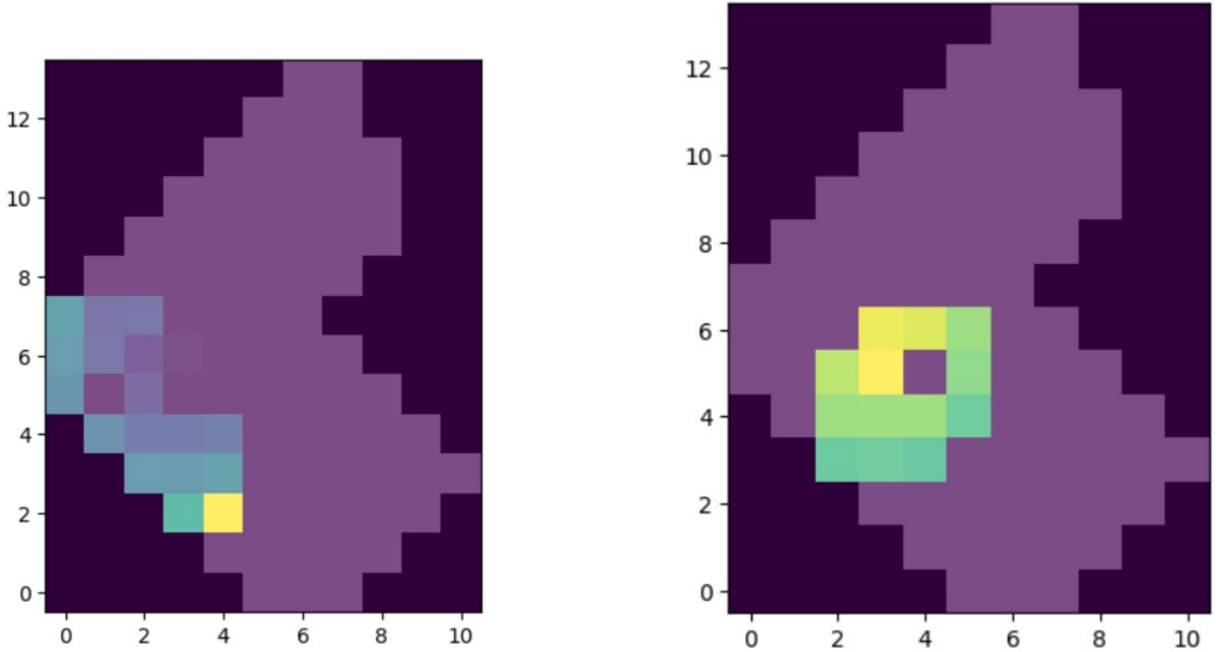
Error: MSE $=\sum 1/n$ (y_true - y_pred)^2

Epochs: 20

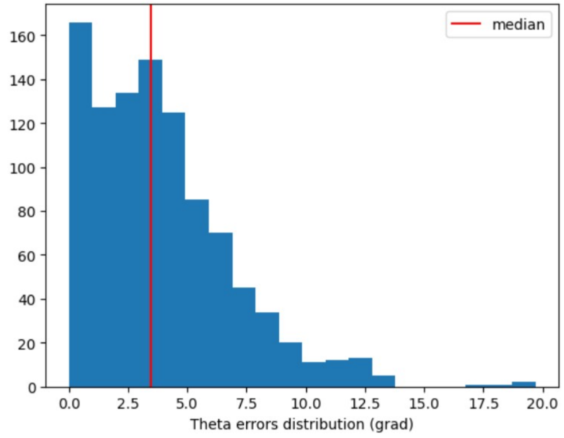
Padding: ‘same’

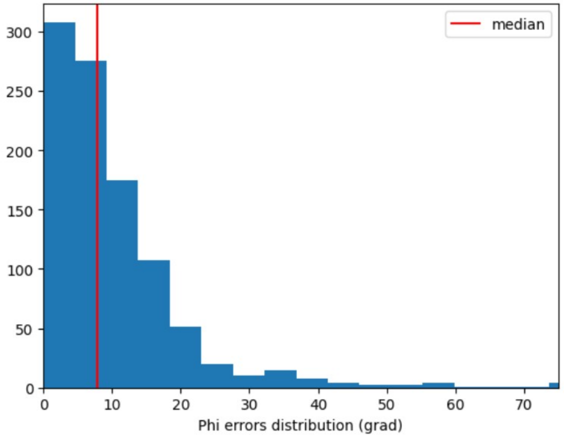
Layer type	Output shape	Activation function
Convolutional 2D, filters = 128	14,14,128	LeakyReLU
Max pooling 2D, pool_size = (2,2)	7,7,128	-
Convolutional 2D, filters = 256	7,7,256	LeakyReLU
Max pooling 2D, pool_size = (2,2)	3,3,256	-
Flatten	2304	-
Dense	576	LeakyReLU
Dense	72	LeakyReLU
Dense	2	Linear

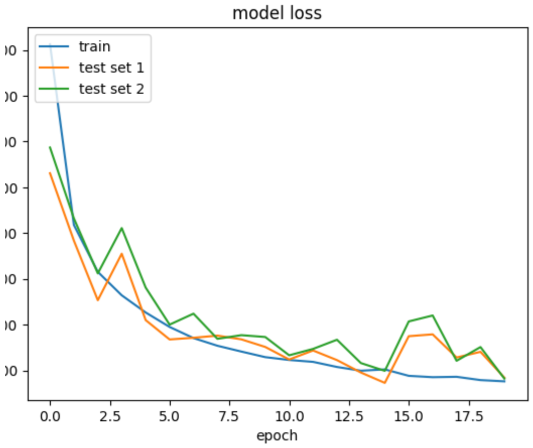
HiSCORE events



CNN model results



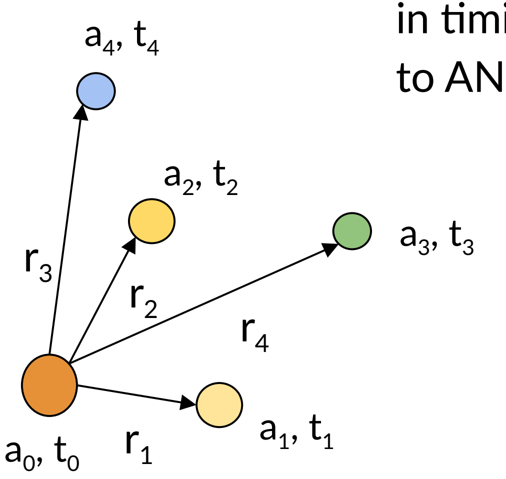




Распределение абсолютных ошибок для тестовой выборки

Reasoning behind ‘dense’ method

- Stations with no signal are taken as zero-value pixels in timing array - confusing to ANN



$$\delta_i = t_0 - t_i$$

Dense-ANN method:

- a fully connected network plays the role of an approximator of the function $F(X) \to \{\theta, \varphi\}$, $Xm = \{a_0, r_1, \alpha_1, \delta_1, a_1, r_2, \alpha_2, \delta_2, a_2, r_3, \alpha_3, \delta_3, a_3\}$,
- no-signal stations are excluded from the F(X)

Dense network

Тип слоя	Output shape	Activation function
Flatten	(None, 25)	-
Dense	(None, 100)	LeakyReLU
Dense	(None, 500)	LeakyReLU
Dense	(None, 1000)	LeakyReLU
Dropout (10%)	(None, 1000)	-
Dense	(None, 200)	LeakyReLU
Dense	(None, 10)	LeakyReLU
Dense	(None, 2)	Linear

Number of parameters: 756,332

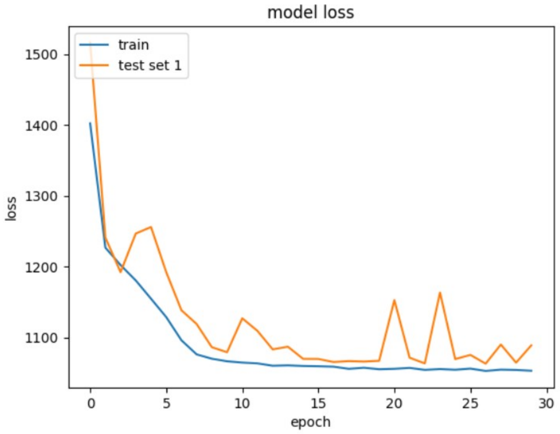
Optimizer: ADAM

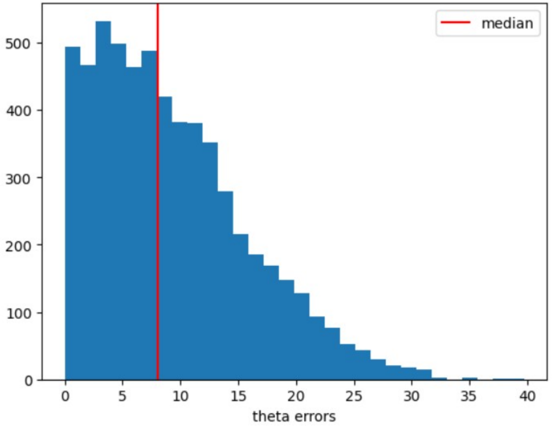
Learning rate: 0,001

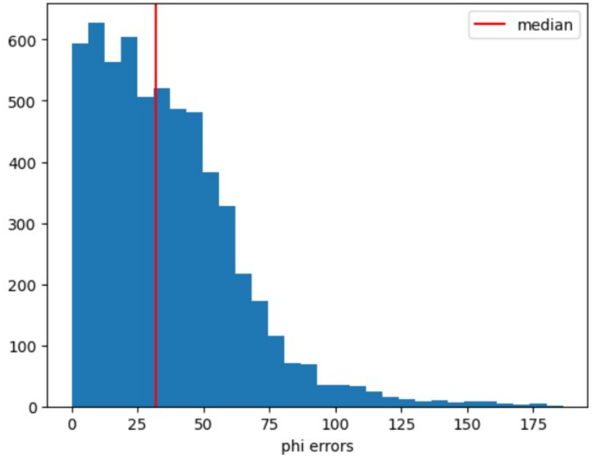
Error: MSE $=\sum 1/n\ (y\_true - y\_pred)^2$

Epochs: 30

Dense model results







Conclusions

Conclusions

MAE for Convolutional model:

4° for theta 32° for phi

MAE for Dense model:

9° for theta 39° for phi

Further investigation is needed. At present, the accuracy of the method is insufficient for the purposes experiment. We associate this with highly unbalanced angle data set.

We hope that hyperparameter optimization of networks, as well as improving the learning process will dramatically improve the accuracy of determining the angle phi, and the accuracy of determining the angle theta make it less than 1 degree.

Thank you for your attention!