THE PRELIMINARY RESULTS ON ANALYSIS OF TAIGA-IACT IMAGES USING CONVOLUTIONAL NEURAL NETWORKS

Elizaveta Gres (ISU, Irkutsk)

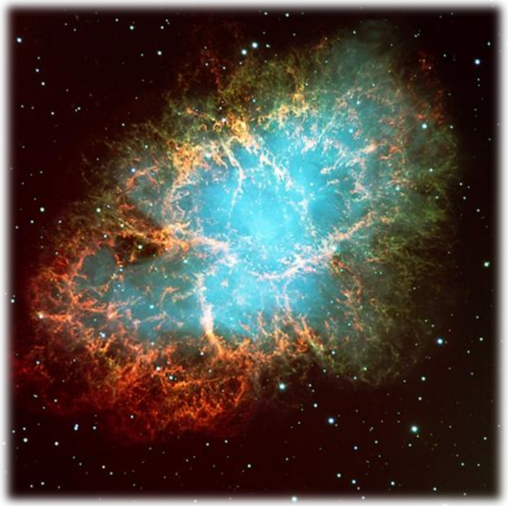
A. Kryukov (MSU, Moscow)

5th International Workshop on Deep Learning in Computational Physics, June 28-29, 2021

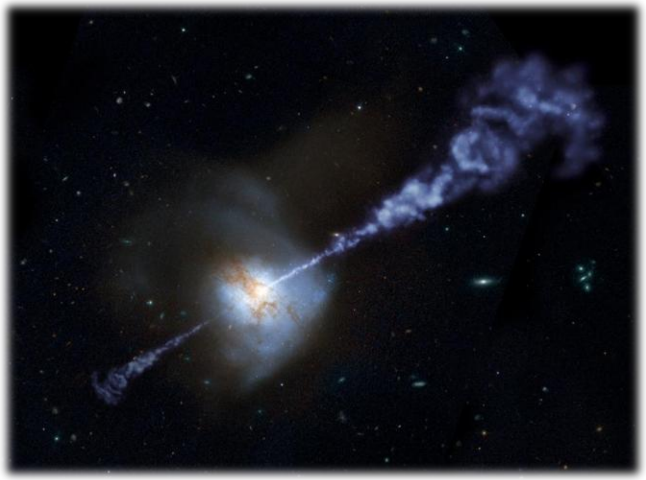
Scientific studies in gamma-astronomy

- Its are the identification and research of high-energy gamma radiation sources.

- The flux, energy spectrum, direction of arrival helps to understand the generation mechanism of high energy gamma radiation and the morphology of sources.



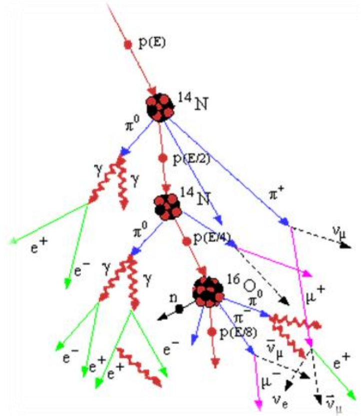
Supernova remnant (Crab Nebula)



Active galactic nucleus

TAIGA-IACT

- Telescopes register Cherenkov radiation created during Extensive Air Showers (EASs).



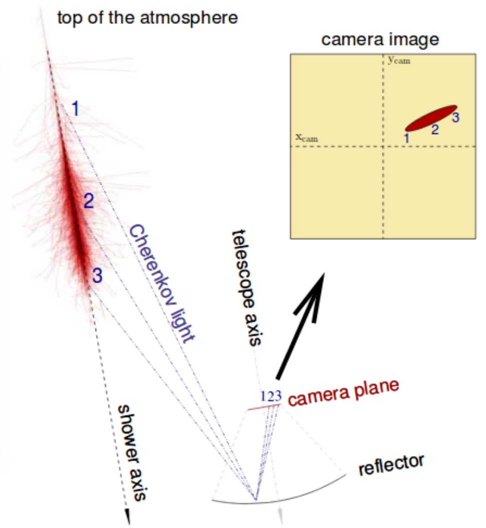
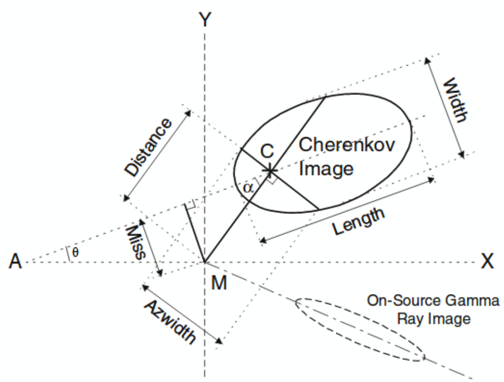
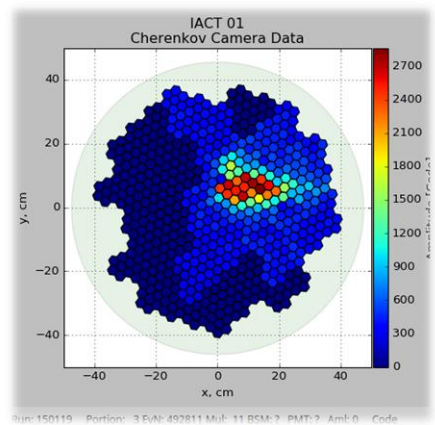


Image analysis techniques:

- Hillas parameter – the image description by an ellipsoid with certain parameters;

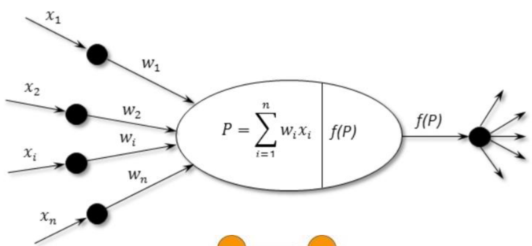


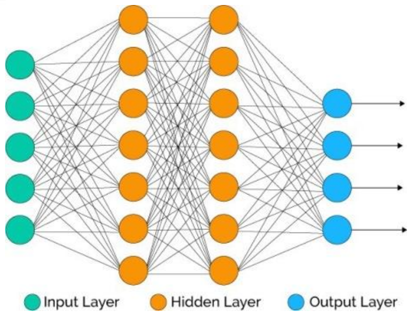


- The Machine Learning.

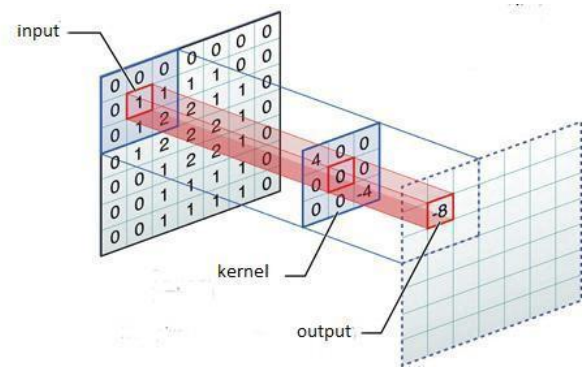
THE MACHINE LEARNING: CONVOLUTIONAL NEURAL NETWORKS (CNN)

A mathematical model of a neuron (*perceptron*) and a multilayer structure of an neural network:





CNN: the presence of convolutional layers, where *filters* (or *kernels*) are applied that highlight more general structures and features in the image:



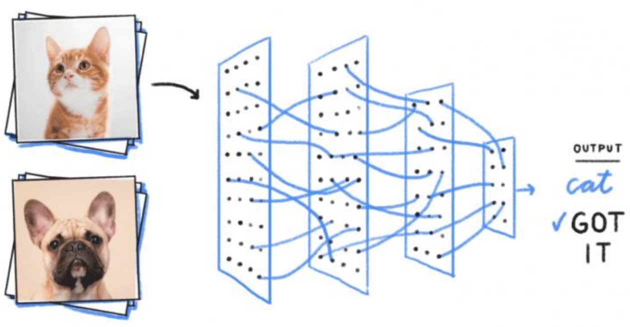
- CNNs are one of the best ways to analyze images.
- At the moment, the use of CNN for processing IACT images has not been implemented.

Thus, this work will allow us to study perspectives of this method in relation to the processing of EAS images.

The main tasks solved in this work:

Classification:

- The separation of images from gamma photons and proton images.



Regression:

- The recovery of the energy of the primary particle from the IACT image;

- The comparison of the energy recovery quality in the case of observations with one and two telescopes.

Model data

The data description:

- Monte-Carlo events simulated with CORSIKA (provided by SINP MSU)

Set	Task	Total events (gamma/proton)	Train / validation separation	Energies
#1 (small)	Classification	30 000 (17 500 / 12 500)	20 000 / 10 000	Hadron: 2-100 TeV γ: 1-60 TeV
#2 (big)	Classification+ regression	200 000 (100 000 / 100 000)	160 000 / 40 000	Hadron: 5-100 TeV γ: 2-50 TeV
#3	Regression	18 000 (only gamma)	12 000 / 6 000	1-50 TeV

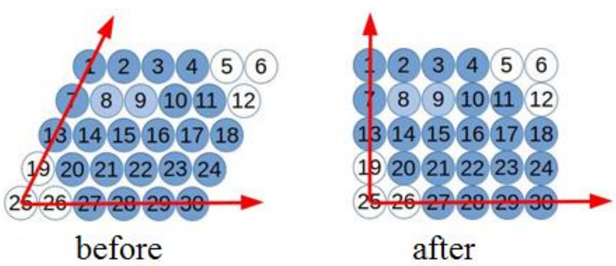
The training samples of small set (#1) were artificially expanded (6 times) by rotating the image each 60 degrees (as camera's symmetry angle)

Model data

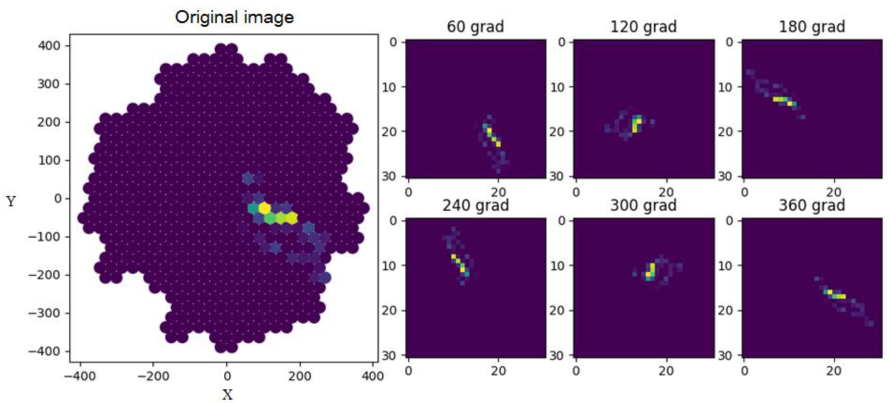
<u>*Image preprocessing*:</u>

- *Cleaning*: the zeroing out negative pixel values, after single non-zero pixels;
- *Pixelization:* the conversion of a hexagonal image to a square form;
- *Scaling:* Logarithmic scaling of pixel amplitudes (x_i) by the formula:

$$\tilde{x}_i = \frac{1}{9} ln(1 + x_i)$$

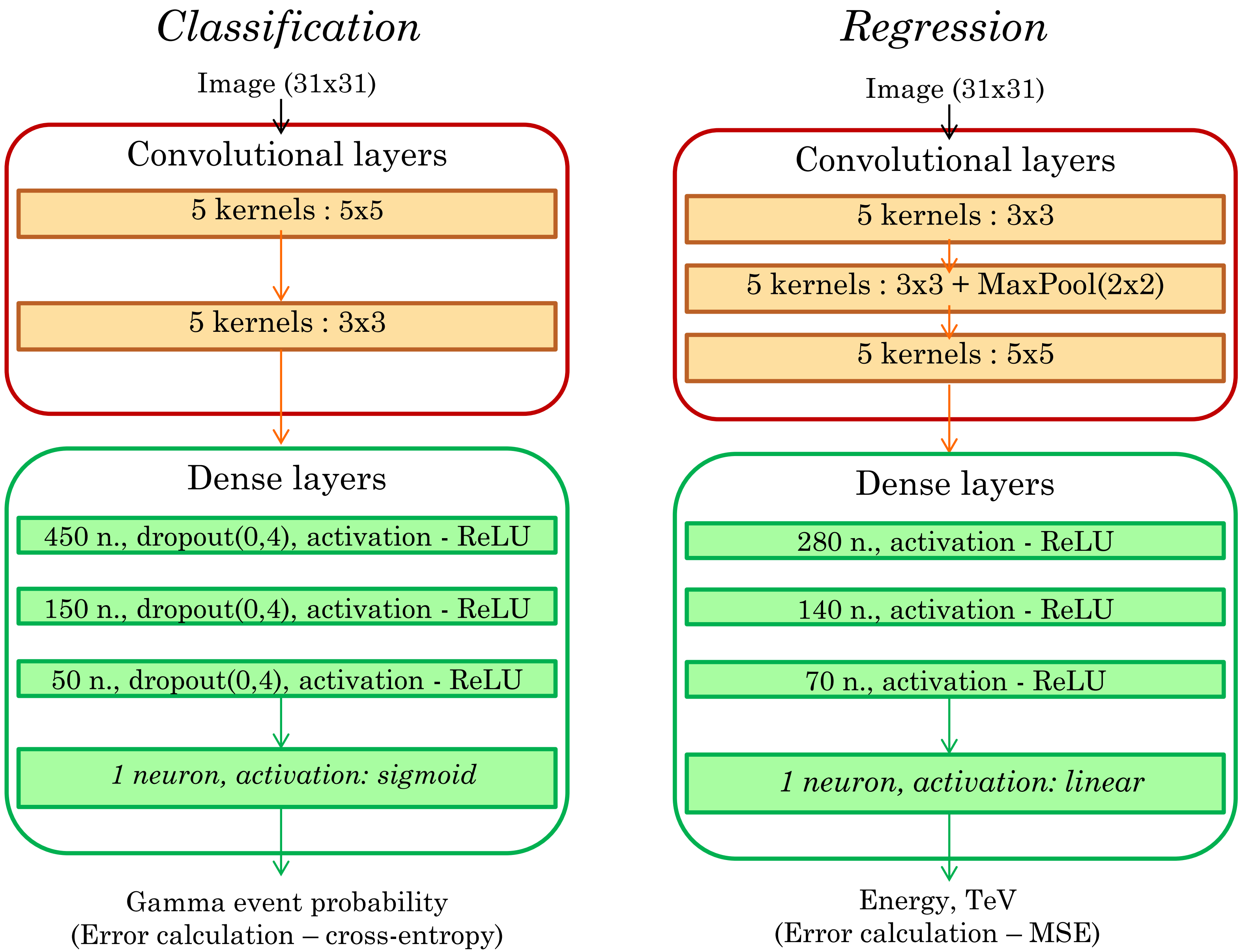


The principle of pixelization



Example of pixelization and rotation the image each 60 degrees

The architecture of CNNs: user model



The architecture of CNNs: ResNet and GoogLeNet

ResNet

Image (31x31)

Conv. 64 kernels : 5x5 + MaxPool(3x3)

2 ResNet's blocks (64 k., 3x3)

3 ResNet's blocks (128 k., 3x3)

1 ResNet's block (256 k., 3x3)

AvrPooling(7x7)

1 neuron

GoogLeNet

Image (31x31)

Conv. 64 kernels : 5x5 + MaxPool(3x3)

Conv. 64 kernels : 1x1

Conv. 192 kernels : 5x5 + MaxPool(3x3)

Inception block

Inception block

Inception block

Conv. 192 kernels : 5x5 + MaxPool(3x3)

1024 n. + dropout(0.7)

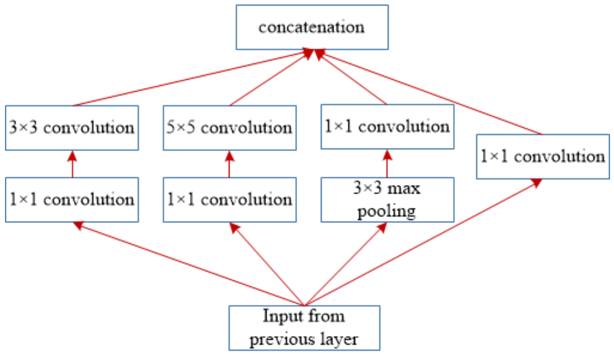
1 neuron

ResNet's block (N k., mxm)

Conv. N kernels : mxm + BatchNorm

Conv. N kernels : mxm + BatchNorm

Inception block



CLASSIFICATION: RESULTS

- Achieved validation accuracy: 96%
- Classification estimation for analysis: $Q = \dfrac{S_{after}}{S_{before}} = \dfrac{N_{g\_g}}{\sqrt{N_{g\_g} + N_{h\_g}}} \Big/ \dfrac{N_g}{\sqrt{N_{total}}}$

  N_{g_g} – the number of true gamma events identified by the CNN as gammas;
  N_{h_g} – the number of proton events identified by the CNN as gammas events;
  N_g – the total number of gamma events in the set;
  N_{total} – the number of all events in the set.

- The class separation threshold: the 50% of true gamma-quanta should remain.

Data classification with balanced classes

Training	N_{total}	N_g	N_h	N_{g_g}	N_{h_g}	S_{after}	Q
#1	10 000	5876	4124	2971	16	54,36	0,93
#2	40 000	20 000	20 000	11677	180	107,24	1,07

There is no improvement for classification with balanced classes

Classification: results

- There is a strong imbalance in the fluxes of gamma radiation and cosmic rays (1:100 000) in real experiment. Therefore, a classification quality check was carried out for unbalanced classes.

Data classification with unbalanced classes

Training (#N_set)	N_{total}	N_g	N_h	N_{g_g}	N_{h_g}	S_{after}	Q
User model (#1)	4 182	58	4 124	25	21	3,69	4,11
User model (#2)	36 783	35	36 748	13	187	0,92	5,04
ResNet (#2)	36 783	35	36 748	18	279	1,04	5,72
GoogLeNet (#2)	36 783	35	36 748	19	262	1,13	6,21

There is a good suppression of proton events, but significance is still low.

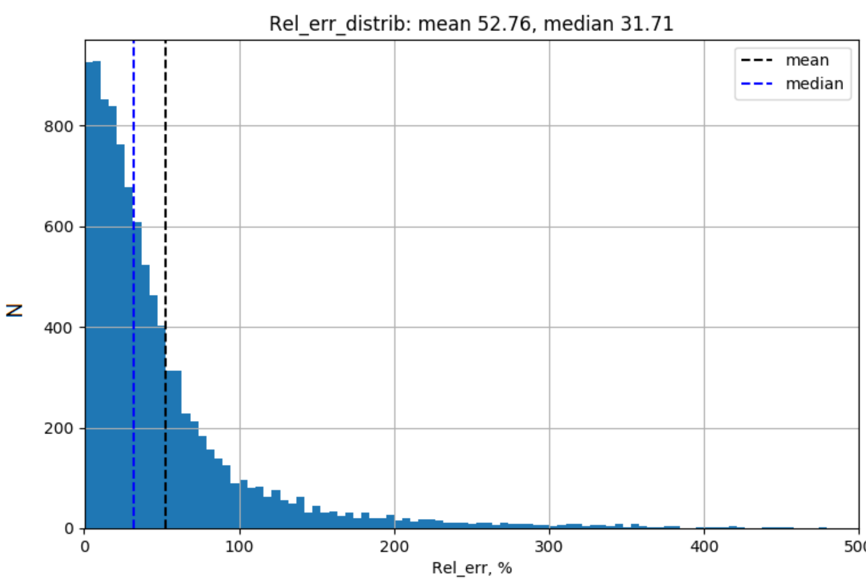
Regression: results

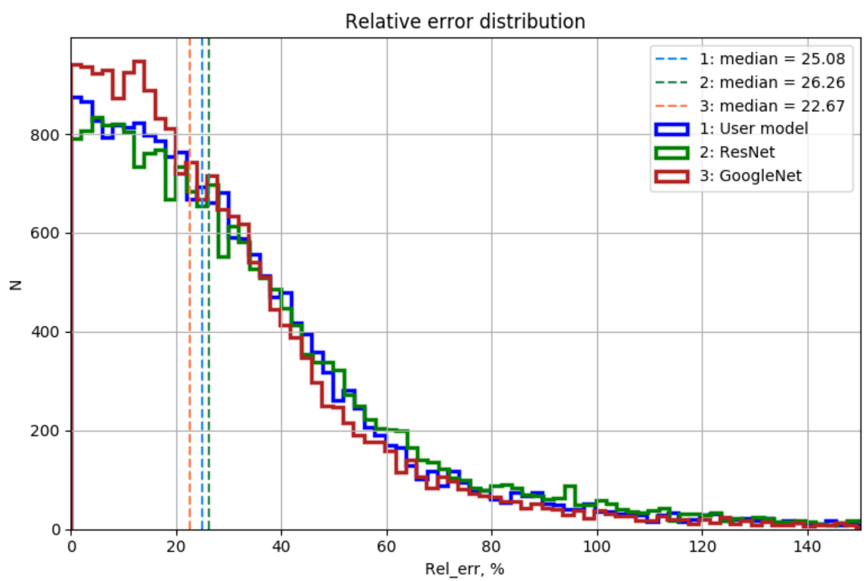
- Regression estimation: $Rel\_err = \frac{|E_{pred} - E_{true}|}{E_{true}}$

  E_{pred} – the energy predicted by CNN;

  E_{true} – the true energy value.

- 32% relative error in case of mixed (gammas and protons) events; 23-25% – in case of only gamma-photons.

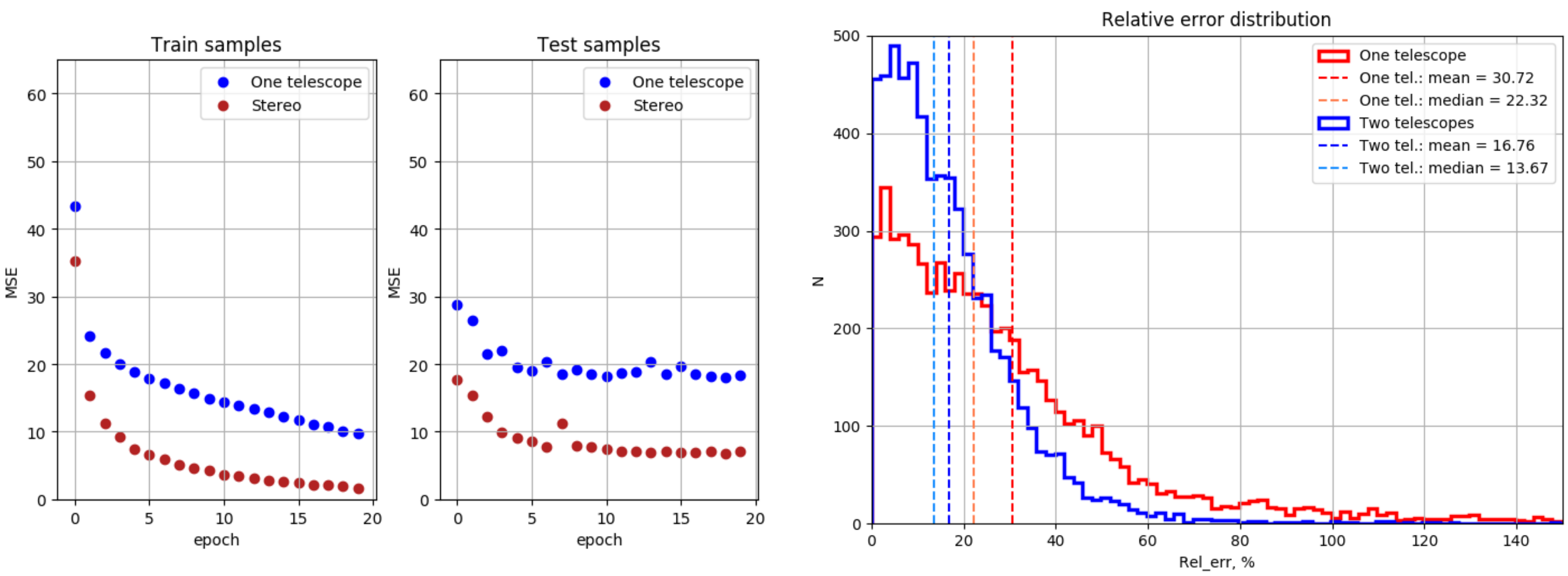




- Different CNN structure do not improve results significantly.

Regression: mono- and stereo-modes

- The second telescope was included in the CNN structure by adding a second channel with convolutional layers.
- The MSE and relative error on validation samples decreased by half in stereo mode.



Conclusion

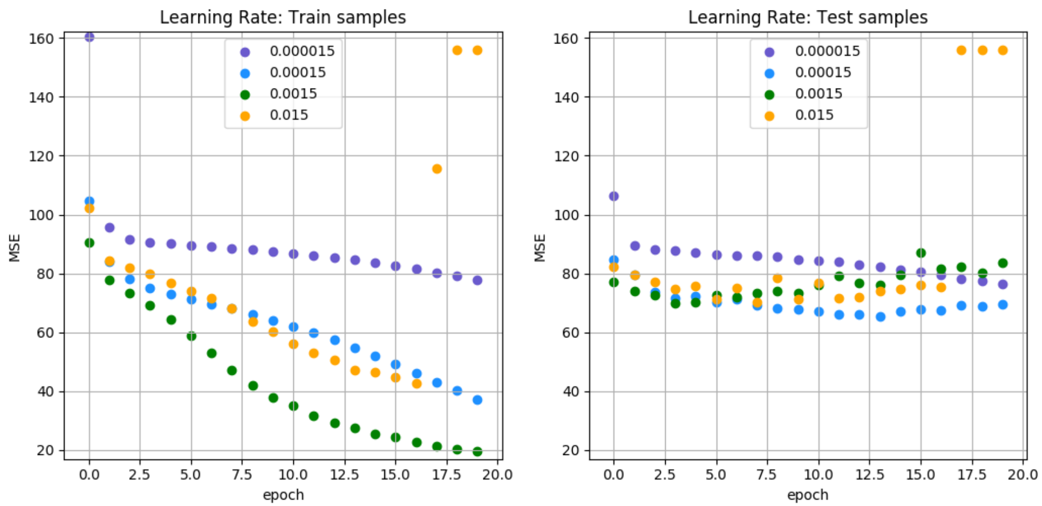
- CNN classification suppresses the proton background greatly (around 100 times), but significance is low (around 1 sigma).
- Regression showed the 24% for one telescope, and 14% – for two telescopes.
- ResNet and GoogLeNet demonstrated a slight results improvement.

- Thus, this method can be used for the energy restoration, as it gives good results. Also CNN for good background suppression can be considered as additional selection threshold.

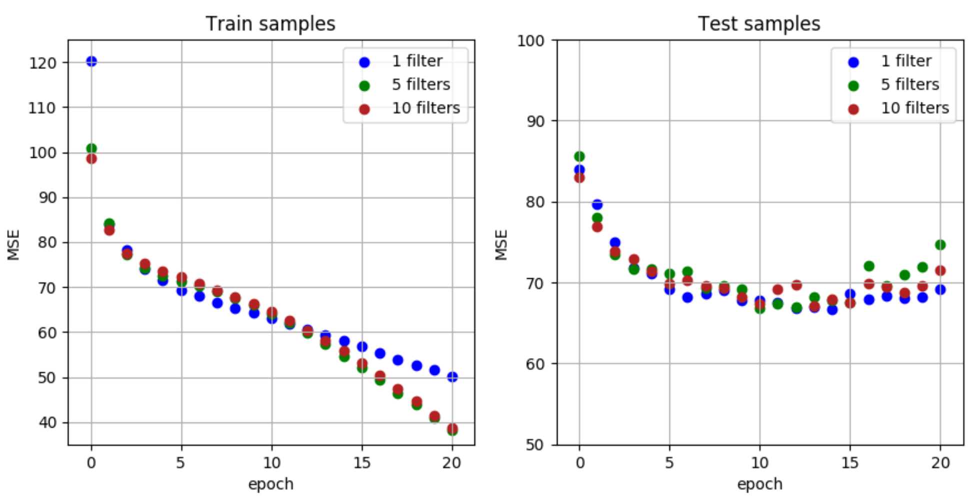
- Future plans:
  - Detailed comparison with traditional method of event selection (by Hillas parameters);
  - Study and application of this method for energy recovery in the case of observations by several (more than 2) telescopes;
  - Search and study of various CNN structures for solving the classification problem.


Thank you for your attention !

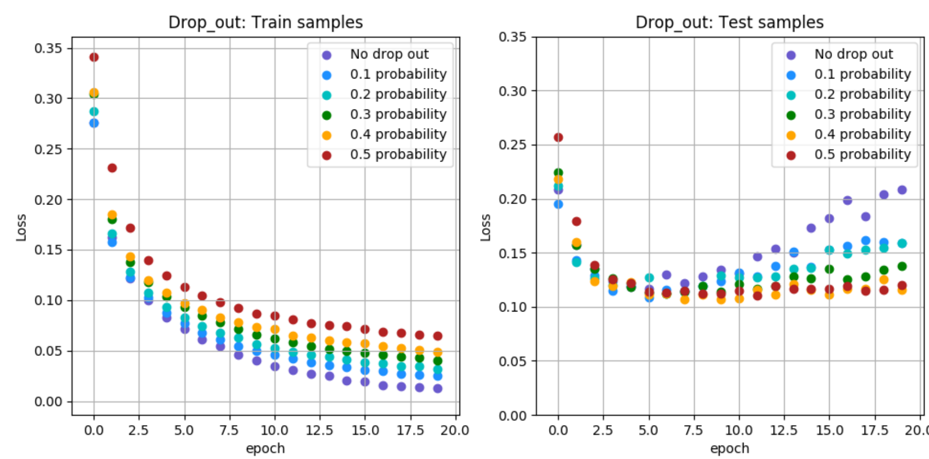
The investigation of CNN architectures (#1)



Training and validation for different learning rate

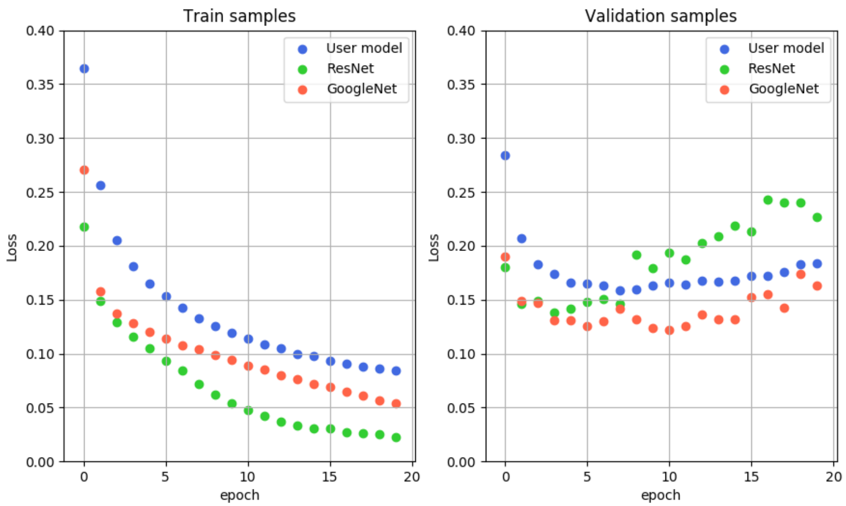


Training and validation for different number of filters

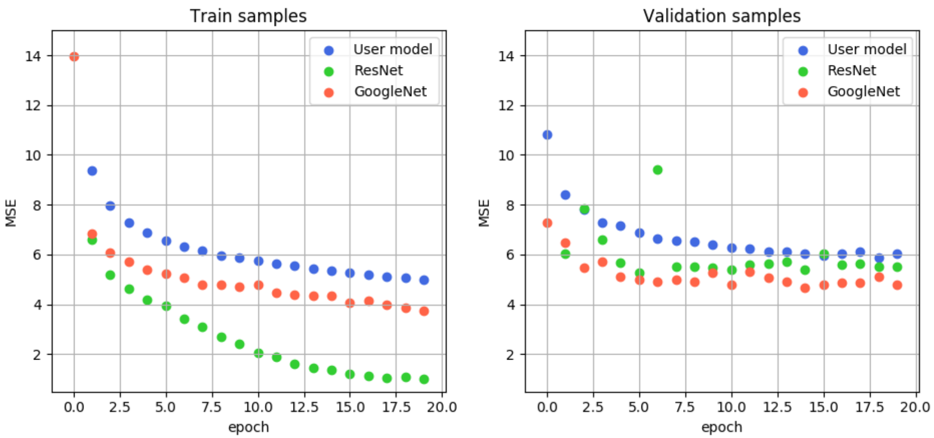


Training and validation for addition dropout-regularization

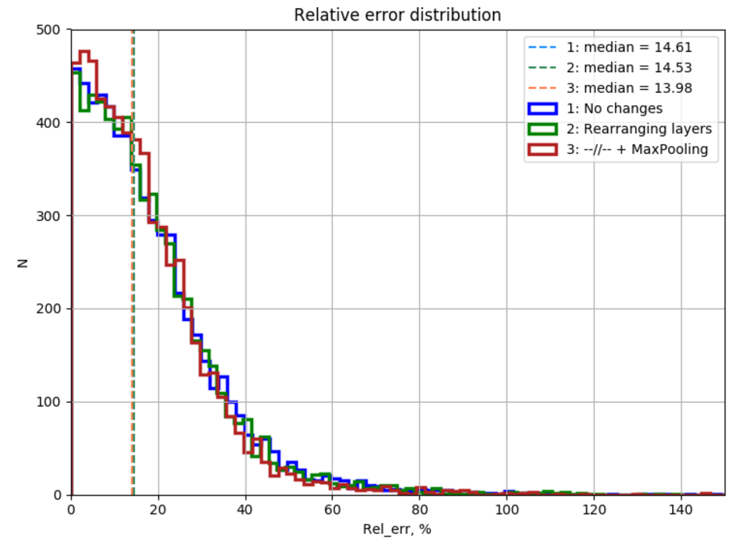
The investigation of CNN architectures (#2)



Training and validation different CNN's (classification)



Training and validation different CNN's (regression)



Relative error for different layer order and for addition of MaxPool-regularization

Hillas parameters

log(Size)	>	1,9
Concentration	>	0,55
Ellipticity	>	0,5
Distance	∈	(0,3; 2,0)

N_{total}	N_g	N_h	N_{g_g}	N_{h_g}	S_{after}	Q
30000	17513	12487	2631	69	50,63	0,50
12662	175	12487	29	69	2,93	1,88
36783	35	36748	2	185	0,15	0,80

Training	N_{total}	N_g	N_h	N_{g_g}	N_{h_g}	S_{after}	Q
#1	10 000	5876	4124	2971	16	54,36	0,93
#2	40 000	20 000	20 000	11677	180	107,24	1,07

Training	N_{total}	N_g	N_h	N_{g_g}	N_{h_g}	S_{after}	Q
User model (#1)	4 182	58	4 124	25	21	3,69	4,11
User model (#2)	36 783	35	36 748	13	187	0,92	5,04
ResNet (#2)	36 783	35	36 748	18	279	1,04	5,72
Googlenet (#2)	36 783	35	36 748	19	262	1,13	6,21