Application of convolutional neural networks for data analysis in TAIGA-HiSCORE experiment

Vlaskina Anna[1,2], Alexander Kryukov[1]

[1]M.V. Lomonosov Moscow State University, Leninskie Gory, Moscow, 119991, Russian Federation

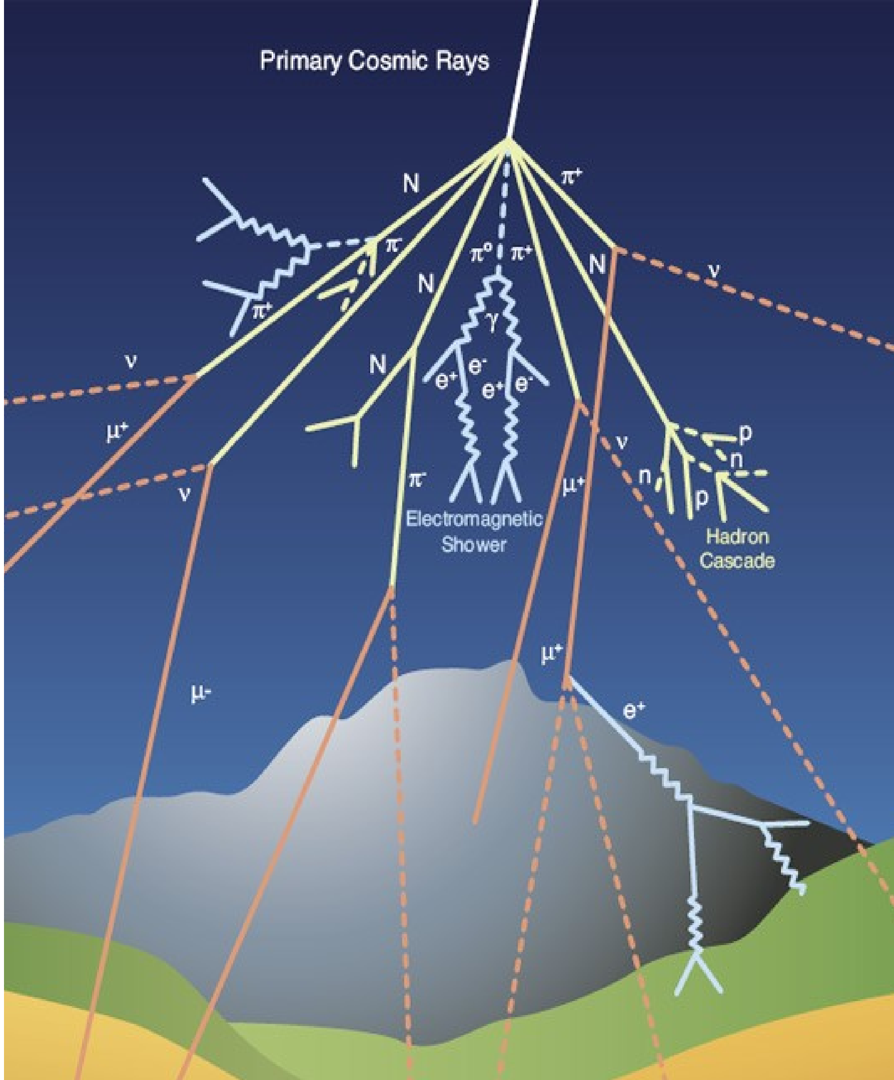
[2]Skobeltsyn Institute of Nuclear Physics, 1(2) Leninskie gory, Moscow 119991, Russian Federation

Extensive Air Showers

- Primary high-energetic particle reaches the atmosphere, generating a massive cascade of secondary particles.

- Secondary particles emit Cherenkov light

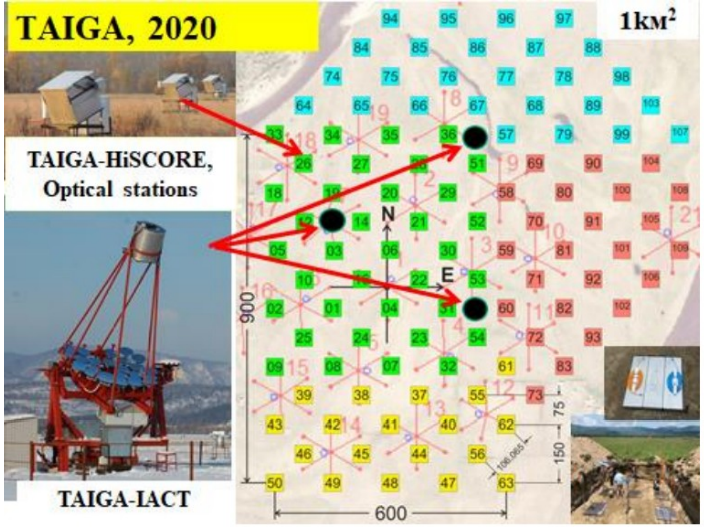
- Air shower parameters: primary particle type, its arrival direction, position of central axis, particle energy



TAIGA experiment
(Tunka Advanced Instrument for cosmic rays and Gamma Astronomy)

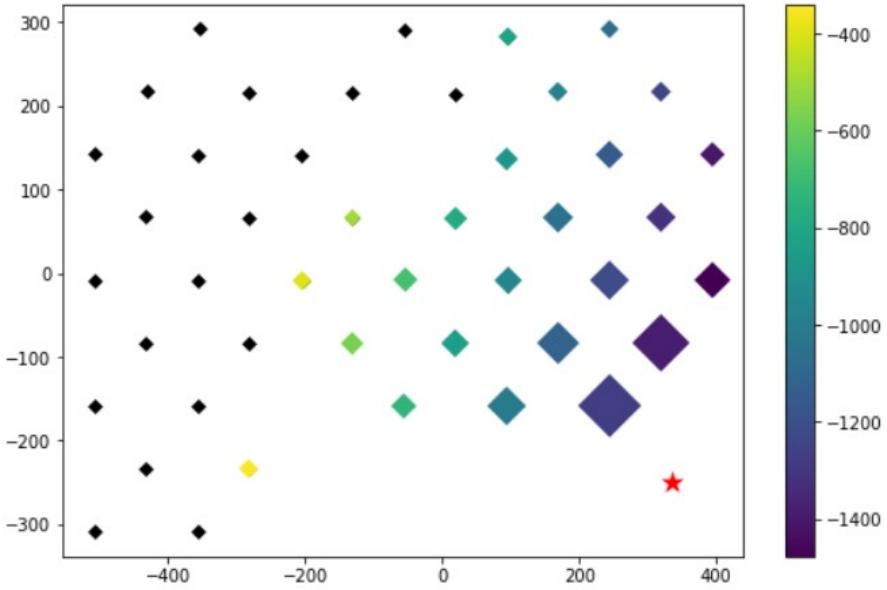
- Instrument for ground-based astronomy from a few PeV to several TeV
- Consists of Cherenkov Telescopes (TAIGA-IACT), Cherenkov stations (TAIGA-HiSCORE), scintillation counters and other instruments.

TAIGA-IACT and TAIGA-HiSCORE



Proposed method

- While working with signal registration times, we look for a time isoline and use it to reconstruct the direction of the shower axis.
- To determine the energy of the primary particle, we use the HiSCORE signal amplitudes.
- We consider an array of signal registration times and an array of signal amplitudes as images.
- For EAS parameter reconstruction, it is proposed to use deep learning technologies, in particular, convolutional neural networks.



An example of an EAS event modeled for the HiSCORE array. The color indicates the time of registration of the particle, the value of the dot corresponds to the value of the signal amplitude.

Artificial Neural Networks (ANN)

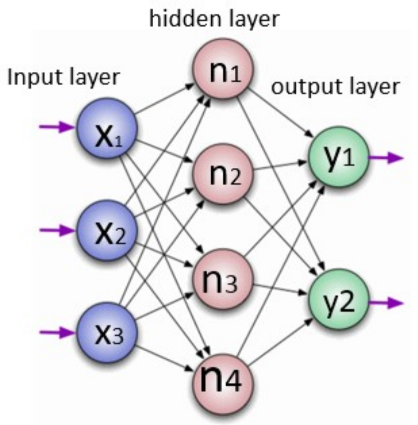
- The input of the neuron is given the expression:

$$y = \sum (w_i * x_i + b_i)$$

- The value of the neuron at the output:

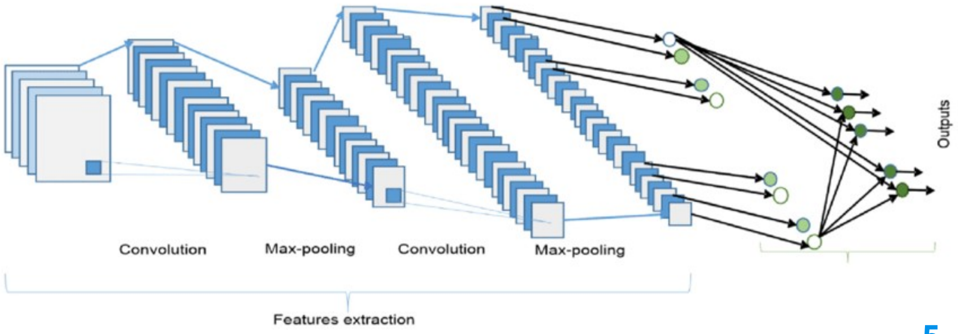
$$y_{out} = f(y)$$

- The goal of training is to find the local minimum of the error function of the weights. To find the local minimum, the gradient descent method is used.
- The network parameters are corrected using the backpropagation method.



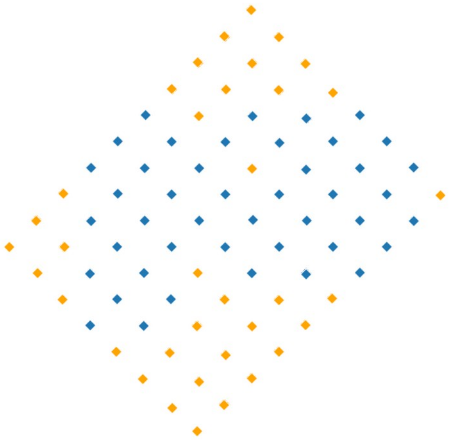
CNNs - type of neural network for data processing with grid topology. The result of the convolution operation is called a feature map.

Due to sparse connectivity, a convolutional neural network extracts only significant features. CNNs also are equivariant to translation.

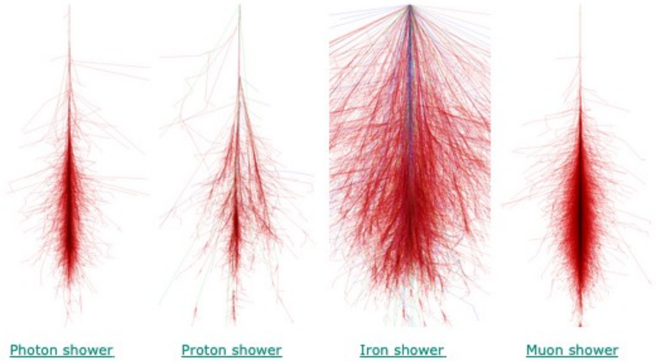


Training and validation sets

- For network training and validation, model data sets obtained using the CORSIKA program were used.
- The sample consists of 12216 events, the simulation was performed for 44 HiSCORE stations.
- The data on EAS registration times are normalized to the latest EAS registration time in the event.
- The array of stations is extended to a rectangular shape: points with zero values are added, the coordinate axes are rotated by 45°



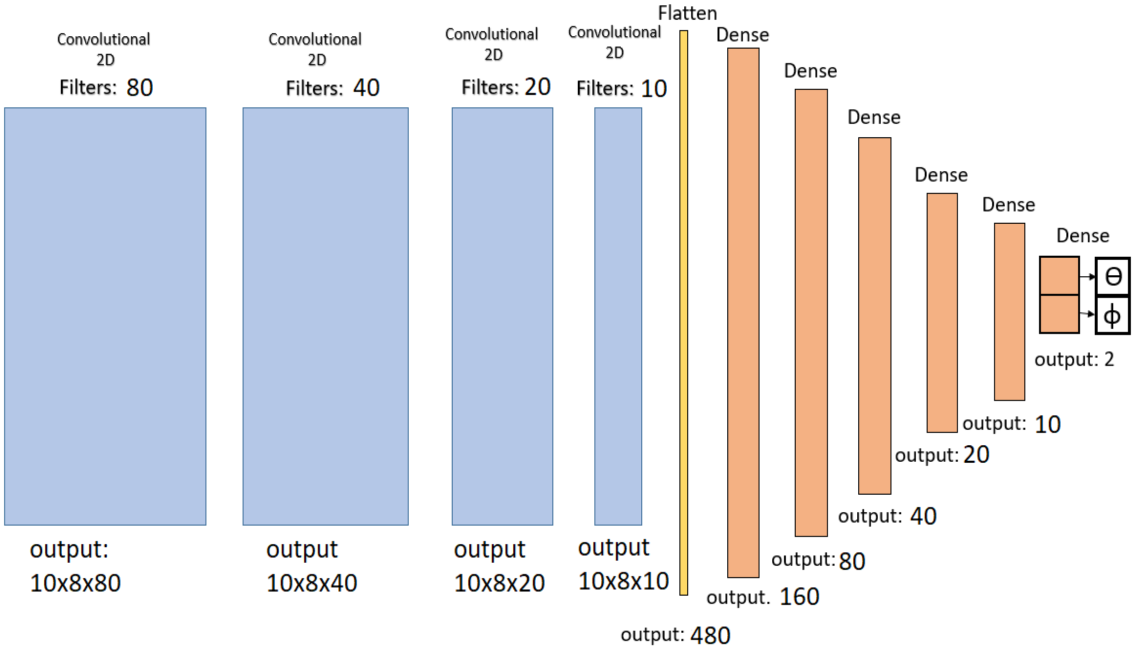
An array of HiSCORE stations reduced to a rectangular shape



CORSIKA simulated images of EAS

Structure of model

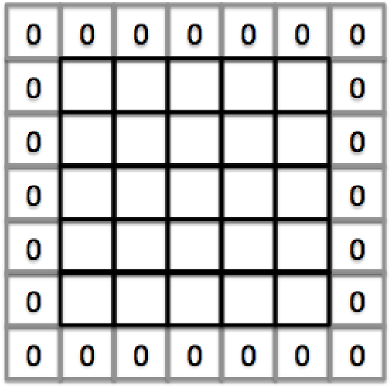
- The network receives as input data on the EAS signal registration time for each triggered station in the event.
- At the output, we expect to get the angles of the EAS axis θ and ϕ determined by the neural network.
- The model will be adapted to determine the energies: instead of data on the time of signal registration, signal amplitude data are used.



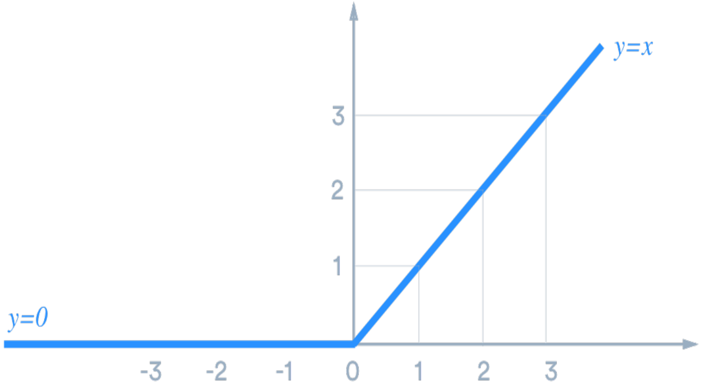
Our model is a convolutional neural network with a serial connection of neural layers. To extract meaningful features from two-dimensional datasets, convolutional layers come first in the neural network.

Structure of model

- Kernel size: 2x2
- Number of model parameters: 135,402
- Optimizer: ADAM (Adaptive Moment Estimation)
- Learning rate: 0,001
- Loss: MSE
- Number of epochs: 50
- Batch size: 15 events
- Padding: ‘same’



Padding - adding pixels around the edges of the image



Activation function ReLU

Results

- MSE = 2.85 (total root mean square error)
- Estimating the average absolute error for the parameters:

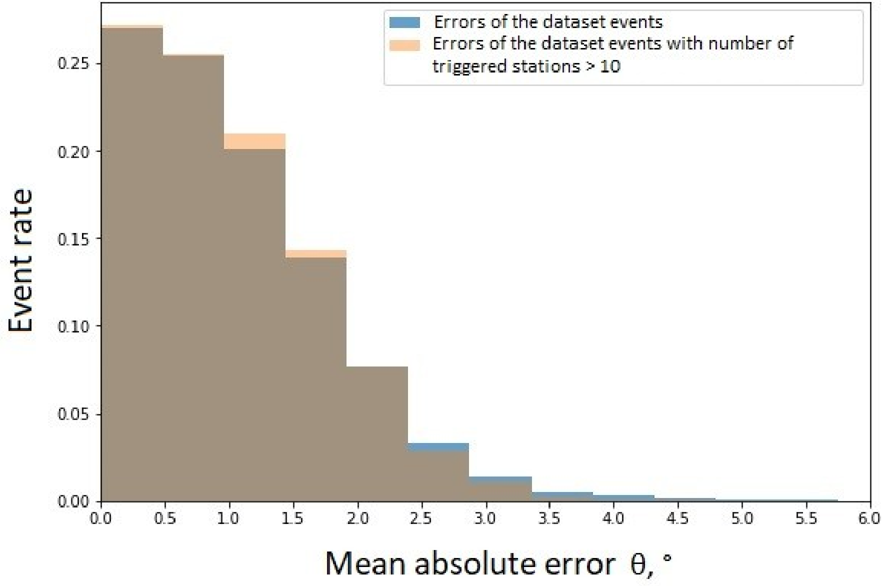
Mean Absolute Error θ: 0.97°

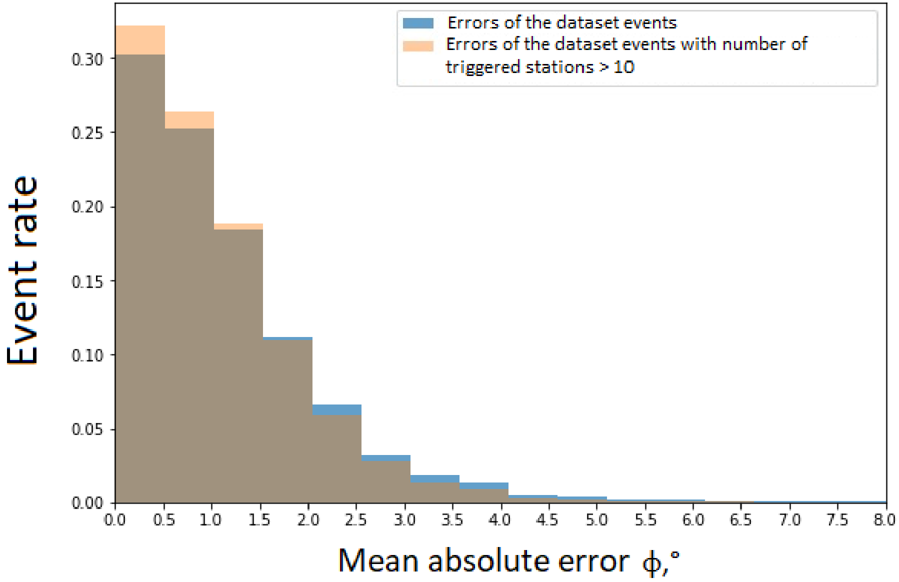
Mean Absolute Error ϕ: 1.38°

Number of triggered stations	$\Delta\theta$, °	$\Delta\varphi$, °	Number of events
≤10	**1.34**	**1.68**	**7**
≥10	**0.85**	**1.12**	**10**

Results

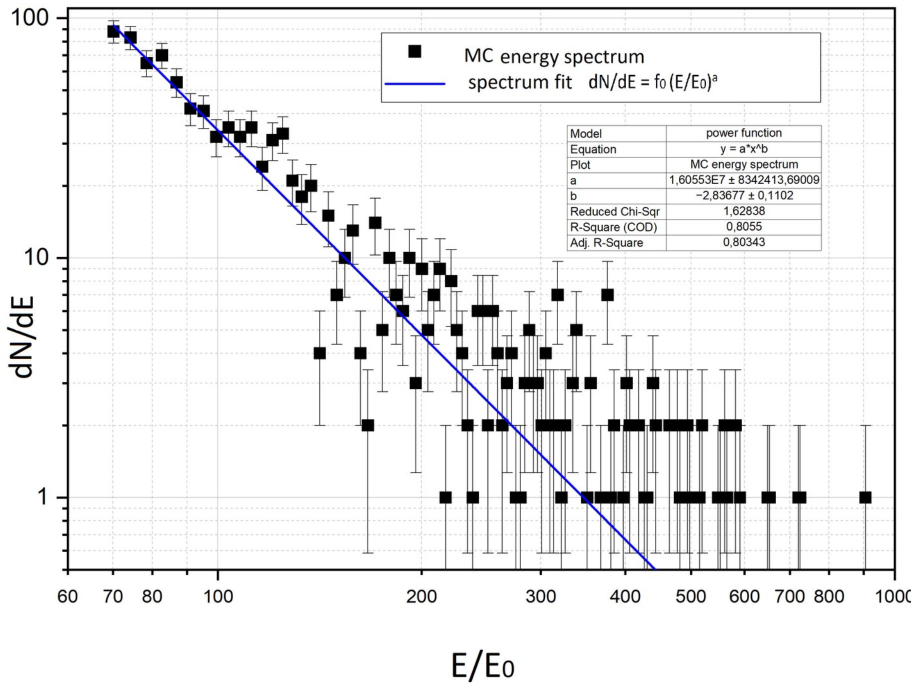
θ and φ error distributions:

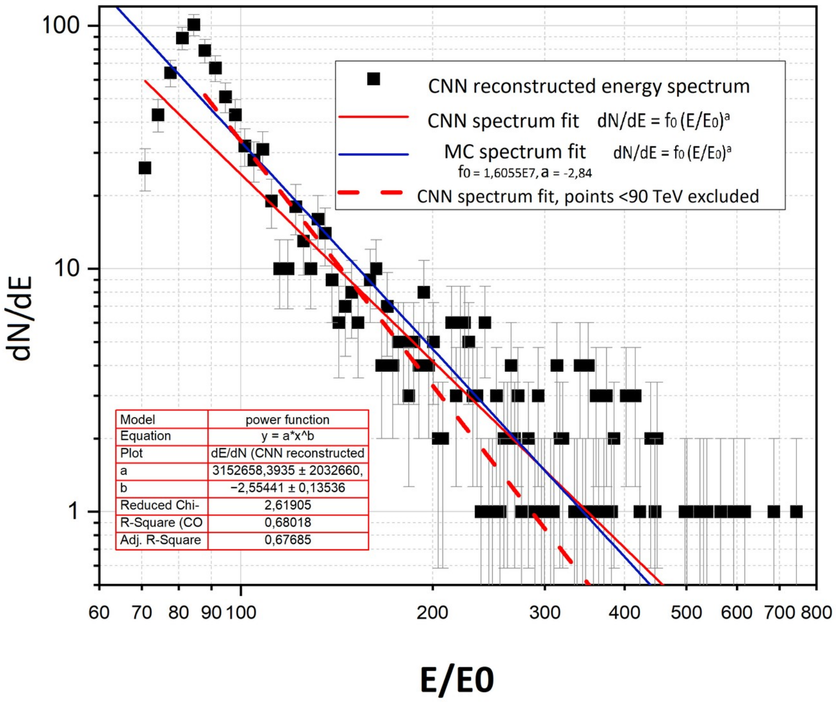




Results: energy spectra

- Fit function: , f_0= 1 TeV





Results: energy spectra

- Fit function: , = 1 TeV
- Comparing to fit in MAGIC experiment (J. Aleksić, S. Ansoldi et al., 2015)

dN/dE	a	Red. Chi-Sq
MC spectrum	-2.83 +/- 0.11	1.628
CNN spectrum	-2.55 +/- 0.14	2.619
MAGIC spectrum	-2.47 +/- 0.01	1.818

Conclusions

- Achieved accuracy of angle determination: ~1°
- The slope of the spectrum reconstructed using the neural network is in good agreement with the MAGIC experiment.

Perspectives:

- Using CNNs for real TAIGA-HiSCORE data
- Considering different network architectures for pattern recognition

Thank you for your attention!

contact e-mail: nina.vankalas@gmail.com

Additional slides

№	Layer type	Output shape	Number of output filters	Activation function
1	Convolutional 2D	10x8x80	80	ReLU
2	Convolutional 2D	10x8x40	40	ReLU
3	Convolutional 2D	10x8x20	20	ReLU
4	Convolutional 2D	10x8x10	10	ReLU
5	Flatten	480	-	-
6	Dense	160	-	ReLU
7	Dense	80	-	ReLU
8	Dense	40	-	ReLU
9	Dense	20	-	ReLU
10	Dense	10	-	ReLU
11	Dense	2	-	ReLU

Adaptive Moment Estimation (Adam) Optimizer

$$m_t = \beta_1 m_{t-1} + (1 - \beta_1)\nabla_\theta J(\theta)$$

$$v_t = \beta_2 v_{t-1} + (1 - \beta_2)\nabla_\theta^2 J(\theta)$$

$$\alpha = \eta \frac{\sqrt{(1 - \beta_2^t)}}{(1 - \beta_1^t)}$$

$$\theta = \theta - \alpha \frac{m_t}{\sqrt{v_t} + \epsilon}$$