Search for optimal deep neural network architecture for gamma detection at KASCADE

Speaker:
Margarita Tsobenko
Co-authors:
P. Bezyazeekov, S. Golovachev, D. Kostunin, V. Lenok,
I. Plokhikh, N. Petrov, D. Reutsky, V. Sotnikov,
V. Tokareva, O. Shchegolev

Supported by the Russian Science Foundation (RSF)

Motivation

- Gamma astronomy enters PeV era
- LHAASO detected PeV gammas[1]. Tibet detected diffuse sub-PeV gammas[2]
- KASCADE exposure comparable with LHAASO one at PeV energies => should contain gamma-rays in dataset
- Using DL with modern architectures[3]

1. Zhen Cao et al.“ Ultrahigh-energy photons up to 1.4 petaelectronvolts from 12 γ-ray Galactic sources”. In: 2021
2. M. Amenomori et al. “First Detection of sub-PeV Diffuse Gamma Rays from the Galactic Disk: Evidence for Ubiquitous Galactic Cosmic Rays beyond PeV Energies”. In: 2021
3. Chao Jin et al. “Classifying cosmic-ray proton and light groups in LHAASO-KM2A experiment with graph neural network”. In: 2019

Data description

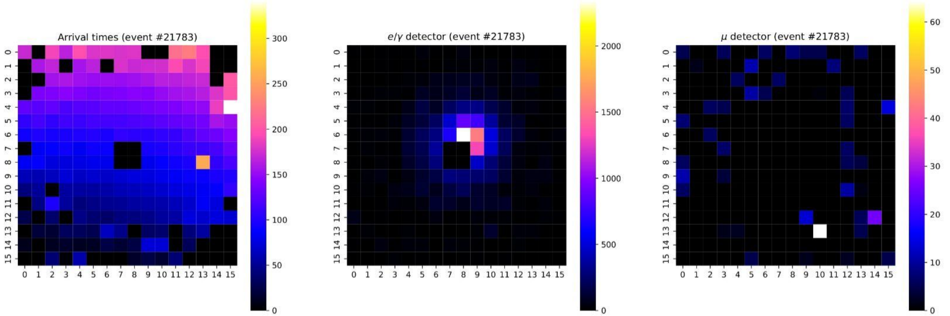
Each event consists of the following components:

- 3 arrays (16x16 shape):
  - arrival times per station (ns)
  - e/γ energy deposit per station (MeV)
  - muon energy deposit per station (MeV)

- Reconstructed features:
  - energy of the primary particle
  - arrival direction
  - number of electrons and muons
  - shower age

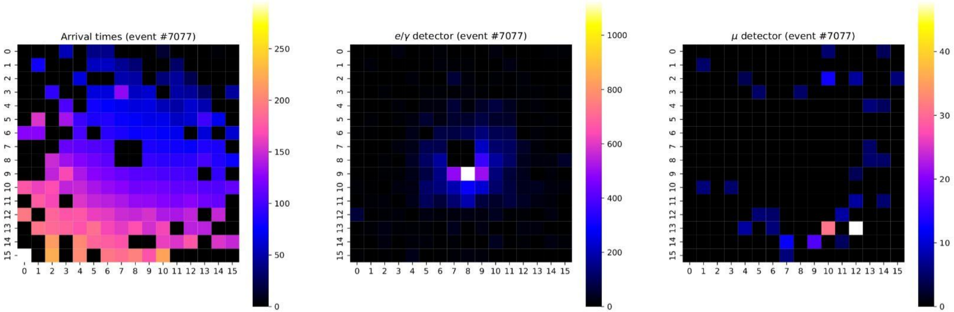
Data and Simulations are taken from KCDC
A.Haungs et al; Eur. Phys. J. C (2018) 78:741;
“The KASCADE Cosmic ray Data Centre KCDC: granting open access to astroparticle physics research data”;
(doi: 10.1140/epjc/s10052-018-6221-2)

Simulated event



lg E [eV] = 15.25, θ = 15.96°, Core distance = 17.91 m

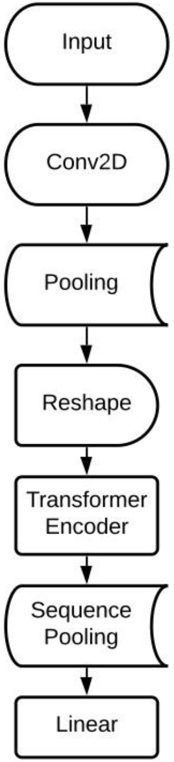
Real event



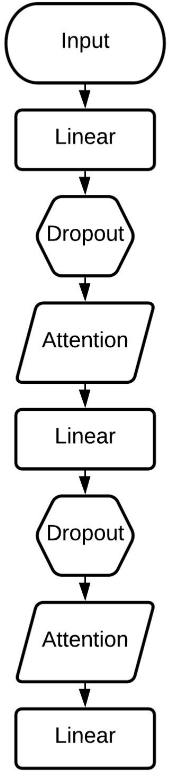
lg E [eV] = 15.10, θ = 14.98°, Core distance = 15.04 m

Models Overview

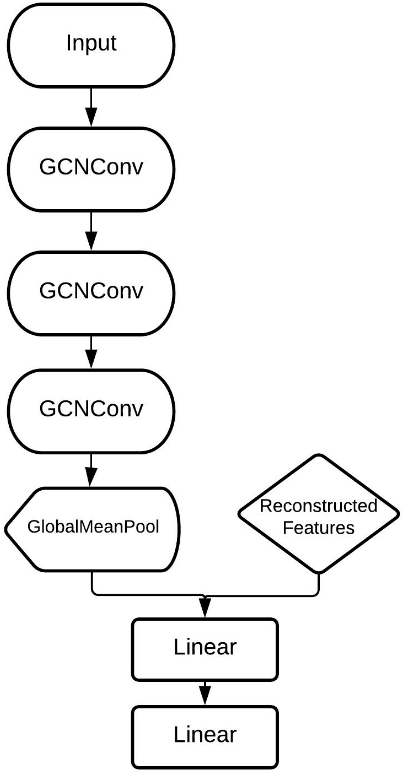
Compact Convolutional Transformer



Self-Attention Network



Graph Convolutional Network



Basic differences

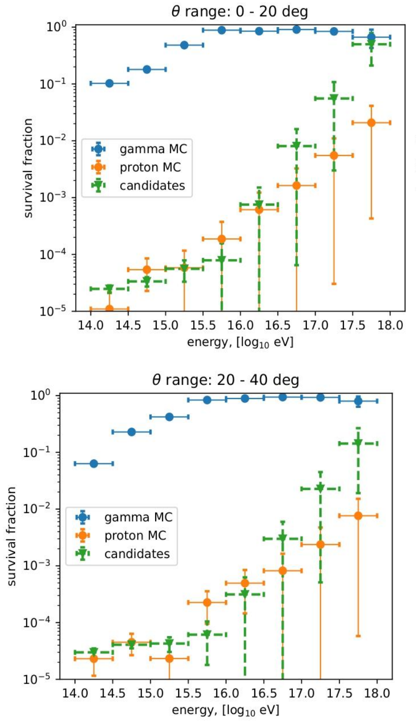
Compact Convolutional Transformer	**Self-Attention Network**	**Graph Convolutional Network**
Input: two-channel image	Input: one-dimensional array	Input: graph with real distances between detector stations
2D Convolutions, Attention, Pooling	Attention	Graph Convolutions
Partially spatially invariant	Spatially non-invariant	Spatially non-invariant

Implementation details

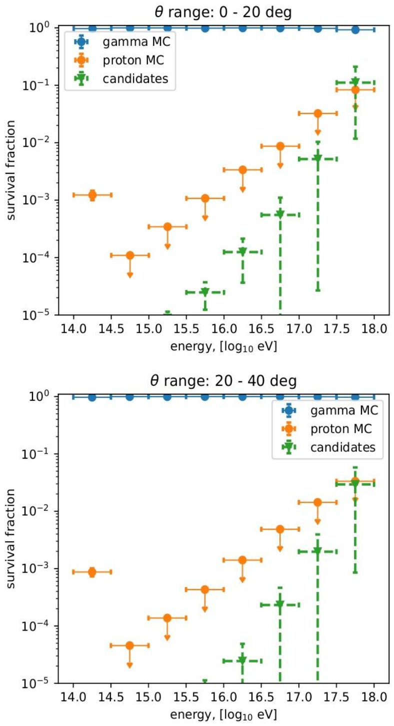
Compact Convolutional Transformer	Self-Attention Network	Graph Convolutional Network
Reconstructed features are not used	Reconstructed features are not used	Reconstructed features are used
Arrival time is not used	Arrival time is not used	Arrival time is used
Number of parameters: 30531	Number of parameters: 30183	Number of parameters: 2952

Intermediate Results

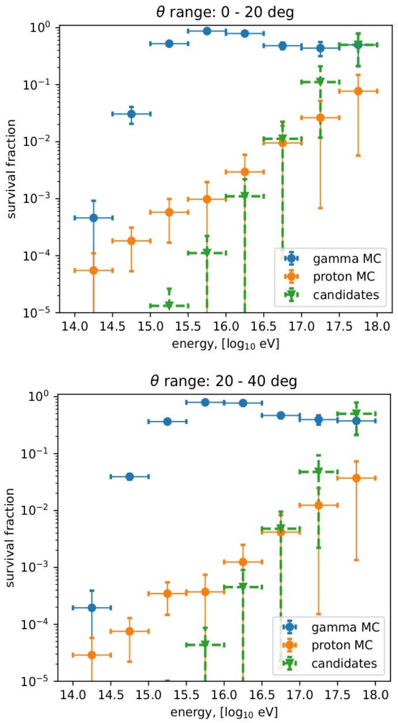
Compact Convolutional Transformer



Self-Attention Network



Graph Convolutional Network



All networks are trained with QGSJet-II.04 protons, photons are mixed from different models

Conclusion

- We are testing different architectures for gamma search at KASCADE
- We use semi-blind approach (80% of data are hidden until classifier is finalized)
- Compact Convolutional transformer satisfactorily describes data
- We need more Monte-Carlo events for network evaluation
  (especially for Self-Attention network)
- GNN has order of magnitude less trainable parameters

Future plans

- Improve the quality of models
  (try new optimization techniques, improve architectures)
- Conduct an ablation study
  (how number of high-energy events in training set affects performance)
- Use more simulated data