Deep learning methods for noise filtering in the NA61/SHINE experiment

Katarzyna Schmidt[1]

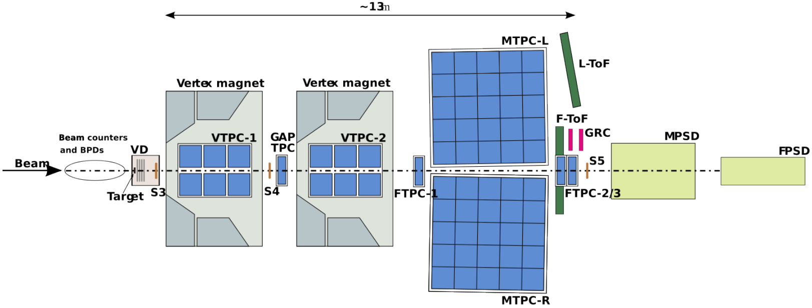
on behalf of:

Yuliia Balkova[1], Wojciech Bryliński[2], Kordian Makulski[2], Marcin Słodkowski[2], Oskar Wyszyński[3]

University of Silesia, Katowice, Poland[1]
Warsaw University of Technology, Warsaw, Poland[2]
Jan Kochanowski University, Kielce, Poland[3]

ACAT2024 11-15 March 2024
Charles B. Wang Center, Stony Brook University, USA

The NA61/SHINE is a fixed target experiment using a large acceptance hadron spectrometer located in the North Area H2 beam line of the CERN SPS



The detection system of the NA61/SHINE experiment consists mainly of 4 large TPC chambers
(2 Vertex-TPC chambers placed in a magnetic field and 2 Main-TPC chambers) and 4 smaller chambers (Gap-TPC and 3 Forward-TPC chambers).

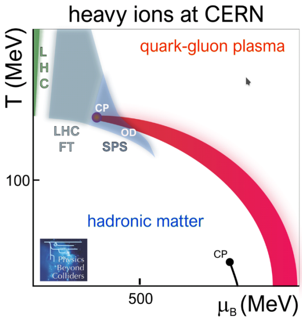
NA61/SHINE physics program

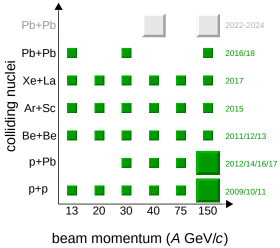
Strong interactions physics:

- Study of the properties of onset of deconfinement and onset of fireball.
- Properties of QCD matter (EoS).
- Exploration region of the critical point of strongly interacting matter.
- Direct measurement of heavy quarks (open charm) at SPS energies.

Cosmic ray and neutrino physics:

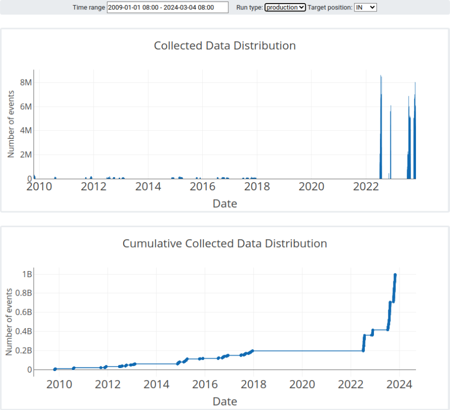
- Measurements of nuclear fragmentation cross section for cosmic ray physics.
- Measurement for neutrino programs at J-PARC and FERMILAB.



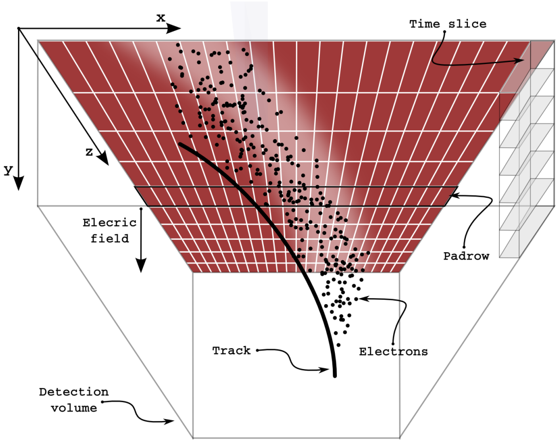


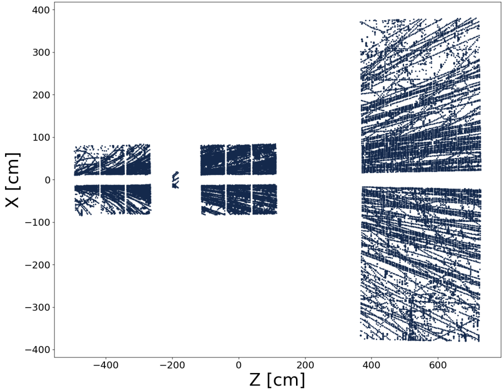
NA61/SHINE detector upgrade 2022

- Time-Projection-Chambers (TPCs) read-out electronics upgrade
- new or upgraded sub-detectors
- new DAQ system =>**2022: 58 mln, 2023: 309 mln events of Pb + Pb at 150 GeV/c**



TPC and real experimental event



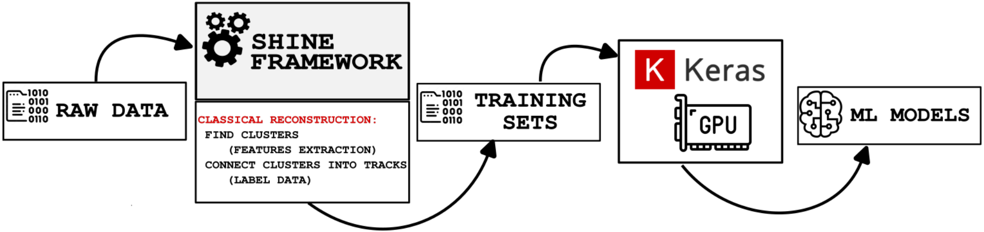


Filtering methods

1. Classical reconstruction - reconstruct local tracks using physics-based algorithms and remove the clusters which do not belong to any track
2. ML models
   - Dense NN (DNN) - clusters properties as the input (charge, X, Y, Z positions, MaxADC...)
   - Convolutional NN (CNN) - 2D images as the input (for each cluster find pad ID and timestamp of max charge deposit then include signal from ± 5 neighbouring pads and ± 9 timestamps)

UNIVERSITY OF SILESIA IN KATOWICE

Implementation part 1



NN architectures

Layer (type)	Output Shape	Parameters #
normalization (Normalization)	(None, 8)	17
dense (Dense)	(None, 256)	2304
dropout (Dropout)	(None, 256)	0
dense_1 (Dense)	(None, 128)	32896
dropout_1 (Dropout)	(None, 128)	0
dense_2 (Dense)	(None, 128)	16512
dense_3 (Dense)	(None, 64)	8256
dense_4 (Dense)	(None, 1)	65

Total parameters: 60,050
Trainable parameters: 60,033
Non-trainable parameters: 17

Layer (type)	Output Shape	Parameters #
conv2d (Conv2D)	(None, 9, 18, 8)	80
max_pooling2d (MaxPooling2D)	(None, 4, 9, 8)	0
conv2d_1 (Conv2D)	(None, 2, 7, 8)	584
max_pooling2d_1 (MaxPooling2D)	(None, 1, 3, 8)	0
flatten (Flatten)	(None, 24)	0
dense (Dense)	(None, 64)	1600
dropout (Dropout)	(None, 64)	0
dense_1 (Dense)	(None, 32)	2080
dropout_1 (Dropout)	(None, 32)	0
dense_2 (Dense)	(None, 16)	528
dropout_2 (Dropout)	(None, 16)	0
dense_3 (Dense)	(None, 1)	17

Total parameters: 4,889
Trainable parameters: 4,889
Non-trainable parameters: 0

UNIVERSITY OF SILESIA IN KATOWICE

Metrics (Ar+Sc collisions at 150 GeV/c)

DNN classification report:

	precision	recall	f1-score
signal	0.84	0.83	0.84
noise	0.94	0.95	0.94
accuracy			0.92

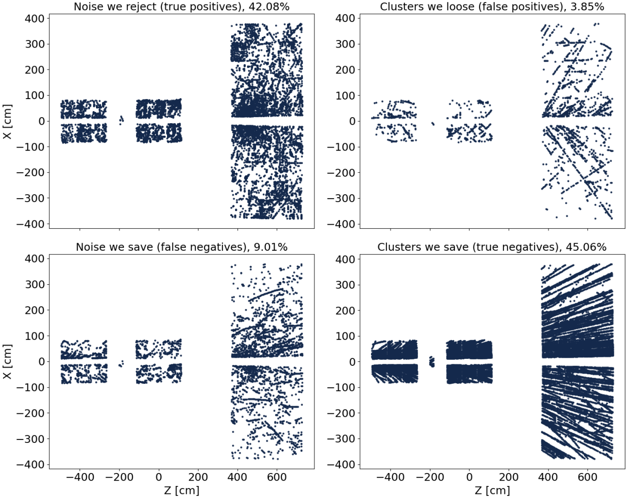
CNN classification report:

	precision	recall	f1-score
signal	0.92	0.85	0.88
noise	0.86	0.93	0.89
accuracy			0.89

- **Precision** – the proportion of positive identifications that were actually correct
- **Recall** - proportion of actual positives identified correctly
- **F1** – harmonic mean of Precision and Recall

UNIVERSITY OF SILESIA IN KATOWICE

CNN confusion matrix (Ar+Sc collisions at 150 GeV/c) thre = 0.5



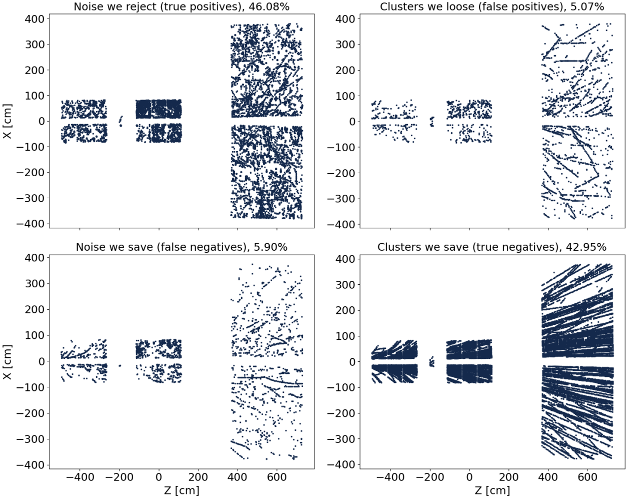
Noisy clusters rejected 85.0%

Clusters lost 7.2%

Noisy clusters saved 14.9%

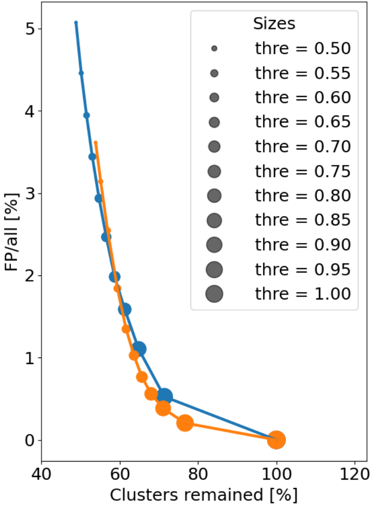
Clusters saved 92.7%

DNN confusion matrix (Ar+Sc collisions at 150 GeV/c) thre = 0.5



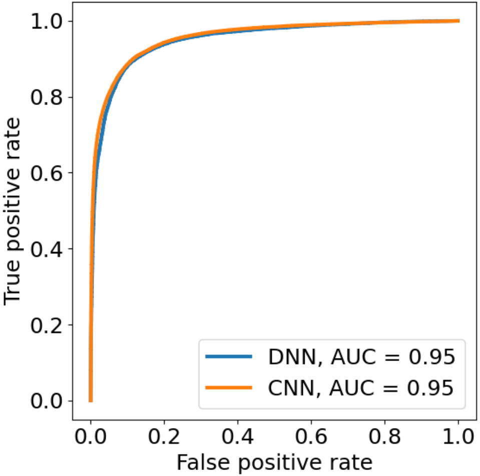
Noisy clusters rejected 88.6%
Clusters lost 10.6%
Noisy clusters saved 11.3%
Clusters saved 89.4%

Parametric cluster classification (Ar+Sc collisions at 150 GeV/c)



False-positive reduction and increase of the stored event size (true negative plus false negative: identified signal plus passed noise) as effect of varying the decision threshold.

Receiver Operating Characteristic (ROC) (Ar+Sc collisions at 150 GeV/c)



Receiver Operating Characteristic (ROC) analysis of CNN-filtered and DNN-filtered data.

Time performance tests (Ar+Sc collisions at 150 GeV/c)

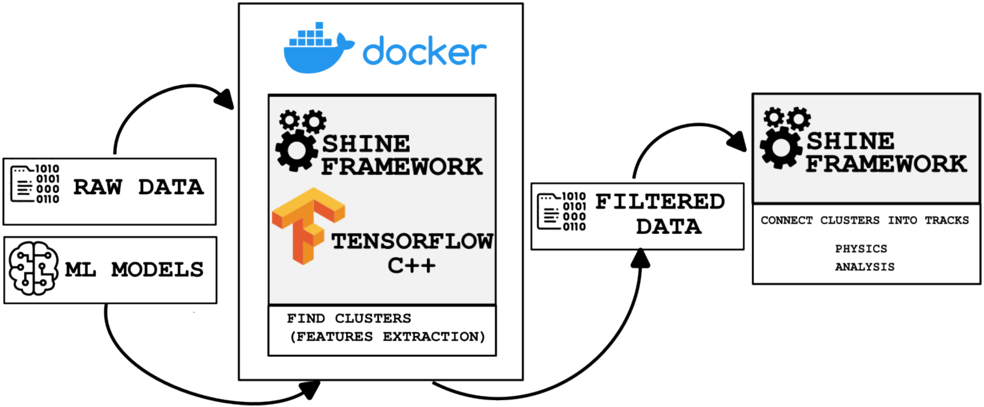
Hardware:

- GPU used: NVIDIA GeForce RTX 3080
- CPU used: Intel(R) Xeon(R) W-1250P CPU @ 4.10GHz

Process	DNN	CNN
Training sets preparation [sec/ev]	1,40	11,37
Training [sec/ev]	14,36	17,01
SUM [sec/ev]	15,76	28,38
Classification [sec/ev]	3,15	7,06
Rejection [sec/ev]	3,65	3,19

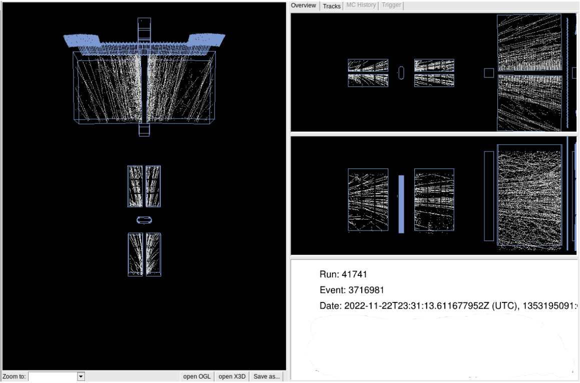
UNIVERSITY OF SILESIA IN KATOWICE

Implementation part 2



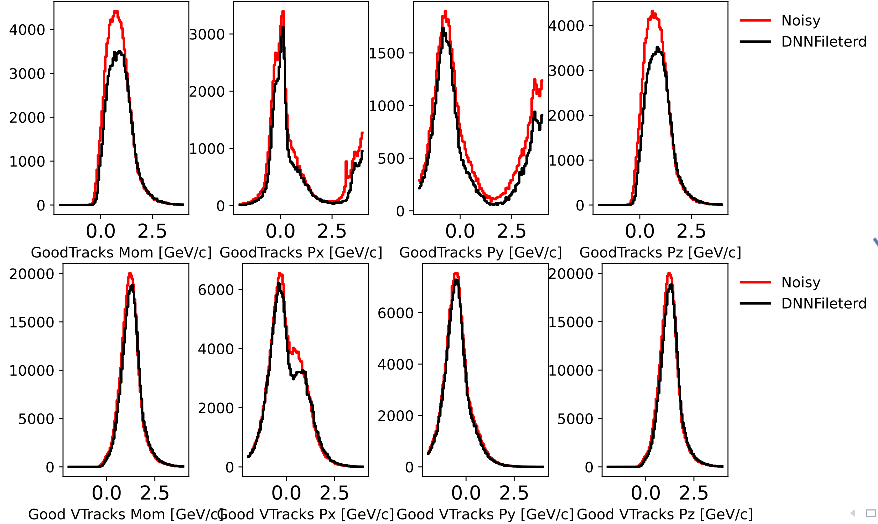
Noisy and filtered data (Ar+Sc collisions at 150 GeV/c)

It actually works!



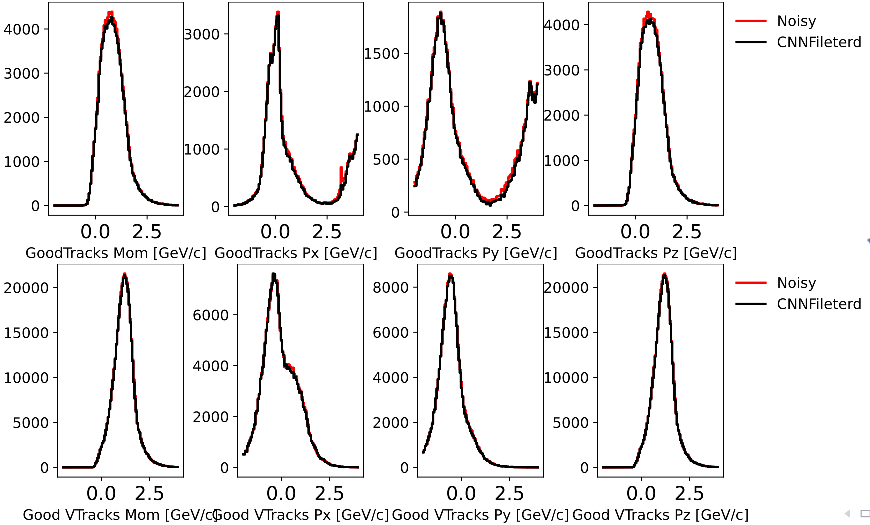
UNIVERSITY OF SILESIA
IN KATOWICE

The Tracks/VTracks momenta distributions (360 Ar+Sc collisions at 150 GeV/c)



WORK IN PROGRESS

The Tracks/VTracks momenta distributions (360 Ar+Sc collisions at 150 GeV/c)



WORK IN PROGRESS

Summary and Conclusion

- Two neural network models: dense and convolutional (DNN, CNN) have been trained to classify noisy clusters.
- Performance analysis for both CNN and DNN was performed.
- Tensorflow C++ was implemented in Shine Offline framework in order to use trained models.
- Classification with the use of CNN model takes longer compared to DNN model.
- A basic comparison of Tracks and VTrack CNN of filtration with native reconstruction shows that using CNN models preserves the properties of multiplicities spectra and tracks momenta distributions.
- DNN filtration causes change in Tracks and VTracks momenta, which is not observed for CNN filtration.
- Careful analysis of the use of a DNN model is needed to understand track multiplicity and momenta discrepancies.
- Comparative analyses of K0s and Lambda particle production for NN filtration models with native reconstruction are being started.
- Further analysis (ex. with higher classification threshold) is required.

UNIVERSITY OF SILESIA IN KATOWICE

Thank You

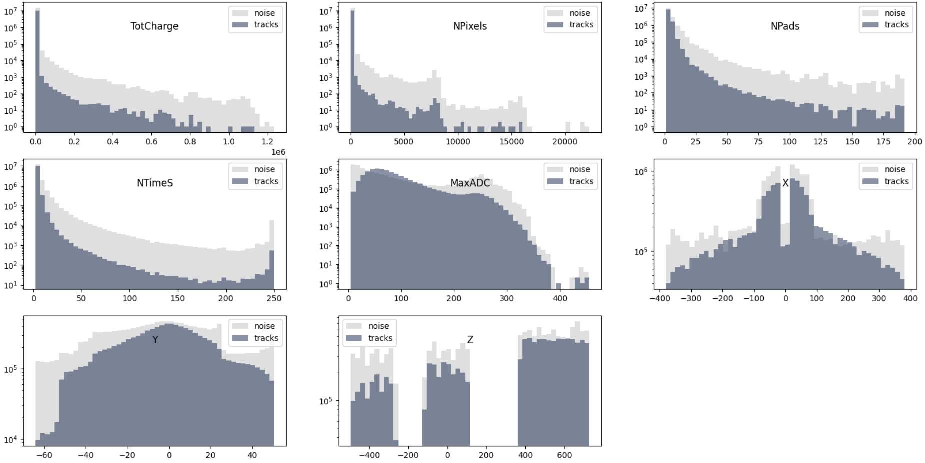
Acknowledgments

We extend our thanks to the NA61/SHINE collaboration for the opportunity to conduct this research. This work was supported by the WUT IDUB and by the National Science Centre Poland grant 2018/31/G/ST2/03910, the Norway Grants in the Polish-Norwegian Research Programme operated by the National Science Centre Poland (grant 2019/34/H/ST2/00585)

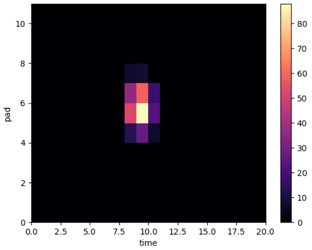
Thank You!

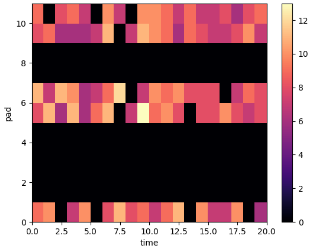
UNIVERSITY OF SILESIA IN KATOWICE

DNN input data: reconstructed properties of the clusters



CNN input data: 2D-Maps (raw cluster data)





- For each cluster find pad ID and timestamp of max charge deposit
- Include signal from ± 5 neighbouring pads and ± 9 timestamps training.