AKADEMIA GÓRNICZO-HUTNICZA
IM. STANISŁAWA STASZICA W KRAKOWIE

AGH UNIVERSITY OF SCIENCE
AND TECHNOLOGY

AGH

Machine Learning for Analysis of Fast Particle Detector Data for Proton Therapy Application

K. Misan[1,2], L. Grzanka[1,3], R. McNulty[4], N. Minafra[5], T. Nowak[3] J.Swakon[3], C. Zacharatou

1. *AGH University of Science and Technology, Kraków, Poland*
2. *Sano Centre for Computational Medicine, Kraków, Poland*
3. *Institute of Nuclear Physics, Krakow, Poland*
4. *School of Physics, University College Dublin, Belfield, Dublin 4, Ireland*
5. *Department of Physics and Astronomy, University of Kansas, Lawrence, KS, United States*
6. *St. Luke's Hospital, Rathgar, Dublin 6, Ireland*

28/05/2023 | FAST2023

Scope

LGAD detectors

- LGAD (Low Gain Avalanche Detector) as an alternative to ionisation chamber in dosimetry
  - High spatial and temporal precision
  - Sensors developed for CMS MTD (MIP Timing Detector) and designed to operate in the LHC environment, Read-out board designed by the University of Kansas
- Sensor consists of 5x5 matrix of pixels with a dimension of 1.3x1.3mm^2 (Fig1)
- Two detectors were placed in the proton beam coming from the AIC-144 cyclotron at the Institute for Nuclear Physics in Cracow, facility was previously used for the proton therapy

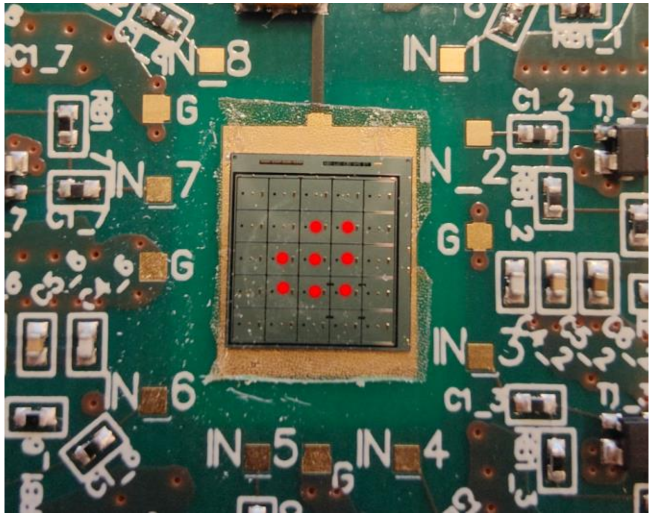


Fig1: LGAD sensor with the connected channels

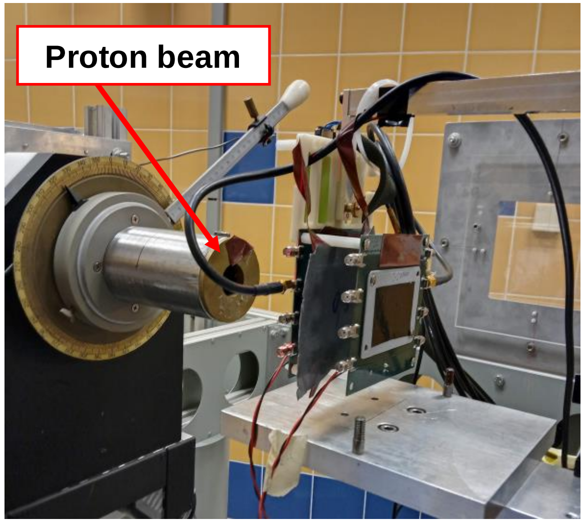


Fig2: Two mounted detector boards

Macro-pulse structure

» AIC-144 cyclotrone in Cracow accelerates protons to an energy of 58 MeV

» The Cyclotrone was producing pulses with the frequency of 50Hz (every 20ms), the length of the pulse was around 0.5ms

» Thanks to the timing precision (~50ps) of the LGAD detector it is possible to analyze the structure of the generated pulse

» A macro-pulse consists of multiple micro-pulses which occur every 38ns and last for around 1ns

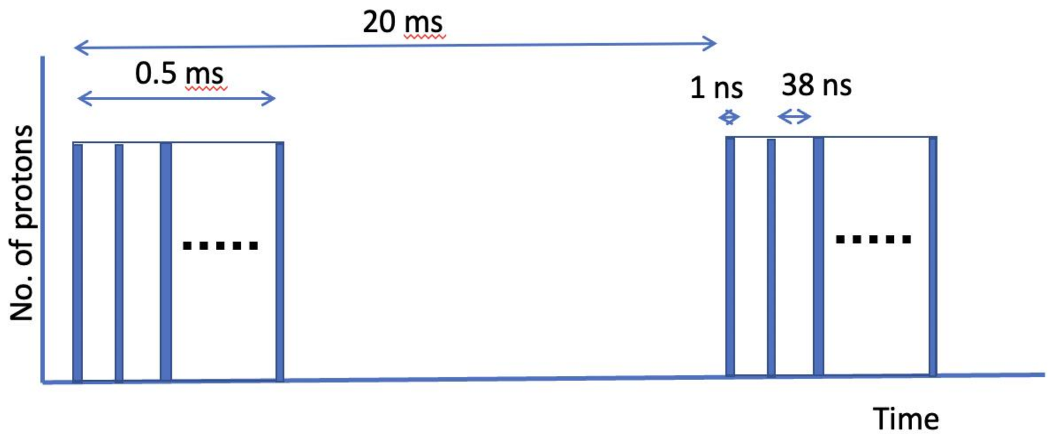


Fig3: Temporal structure of the macro-pulse

Macro-pulse structure

» We found that each micro-pulse could produce a negative (actual signals) or a positive (cross-talk) signal in the detector.

» Splitting data into individual windows allows to perform quantitive analysis and study the shape of the pulses

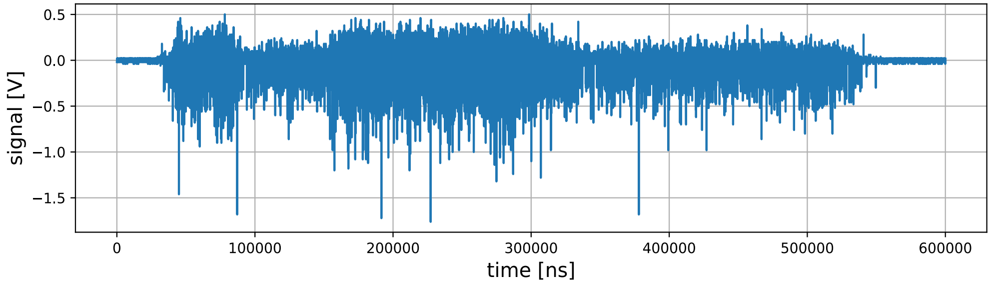


Fig4: Macro-pulse recorded by the detector for 2nA with 8 bit precision

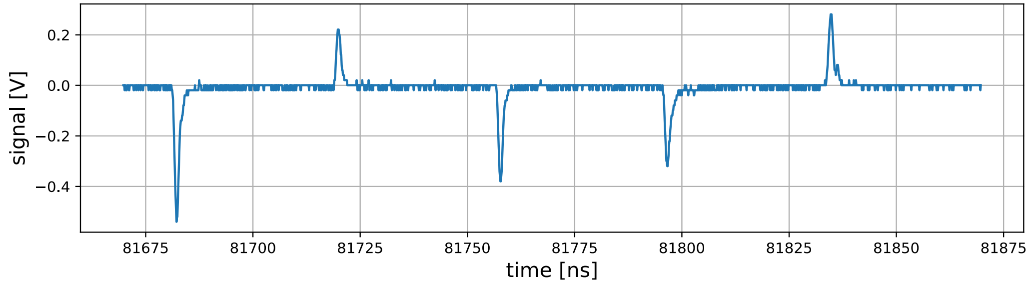


Fig5: Fragment of the macro-pulse with several micro-pulses

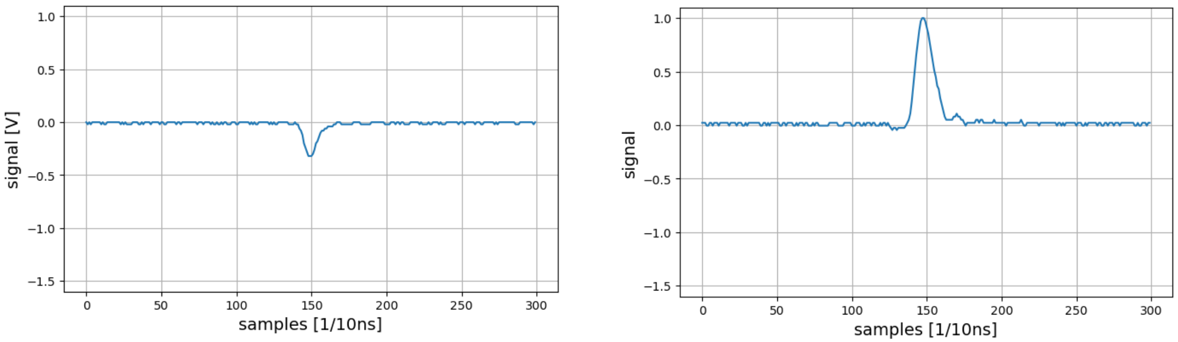


Fig6: Example of the micro-pulse window extracted from the macro-pulse, raw values on the left and normalized on the right

Problem statement

- LGAD detector can record different events depending on the number of particles passing through the sensor and the position of the deposited charge in the pixel matrix
- Analyzing those signals using conventional methods is hard given the number of possible signal shapes and interferences like noise, measurement precision and cross-talk
- Machine learning methods proved to be very powerful at tasks like pattern recognition, outlier detection, noise reduction and signal segmentation

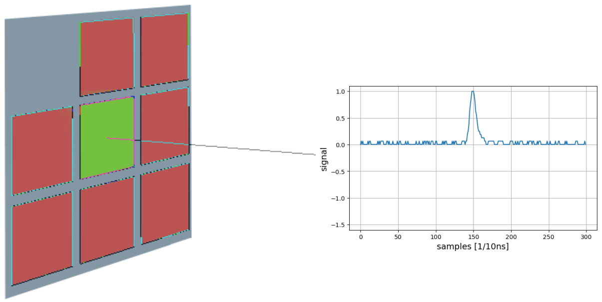


Fig7: Single particle in the middle channel

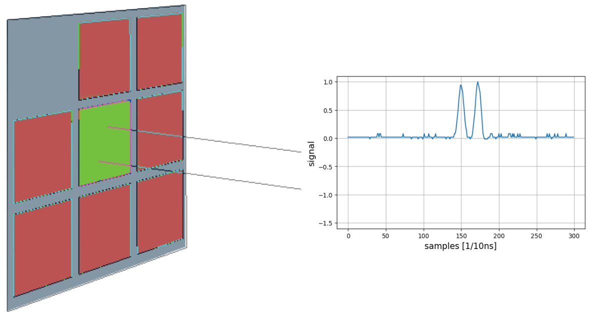


Fig8: Multiple particles in the single channel

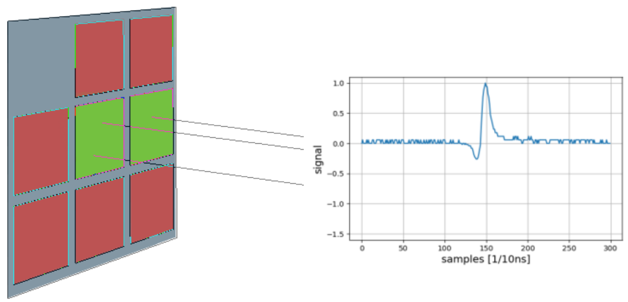
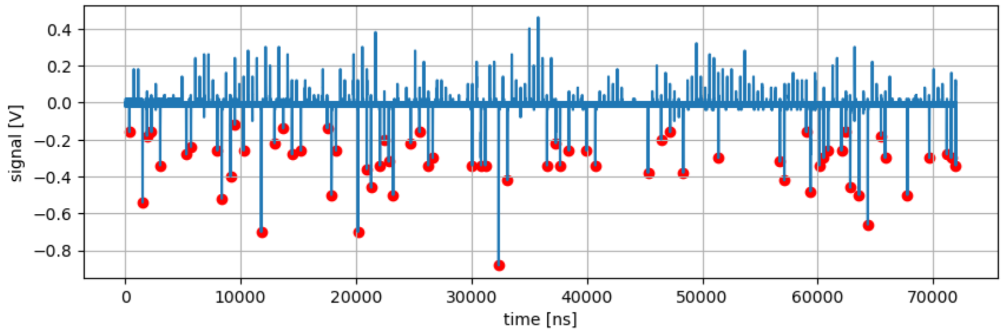


Fig9: Possible crosstalk between two activated channels

AGH

Peak detection algorithm

- Start of the macro-pulse is defined by the first micro-pulse, similarly end of the macropulse is defined by the last micro-pulse
- Baseline of the signal noise is calculated on the fly by the average of N last processed samples
- Peak is defined as the N% change from the baseline
- To compenstate for differences in amplitudes, macro-pulse is normalized:
  - Negative pulses are scaled to 1 while positive pulses remain unchanged



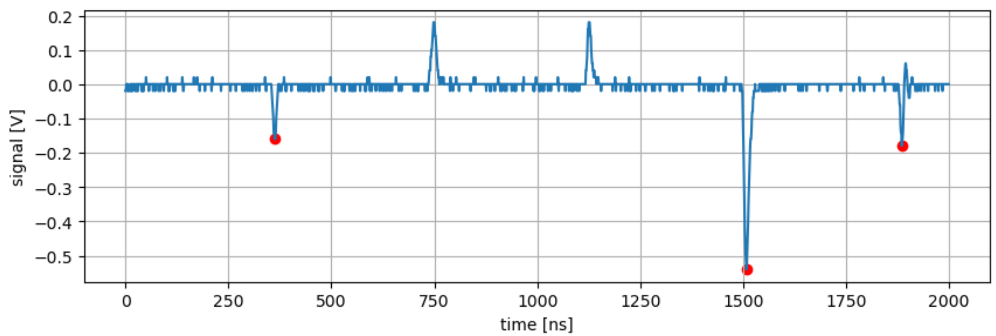


Fig10: Example of the macro-pulse fragment with micro-pulses maxima indicated by the red dot

Windows classification

» Shape of the signal pulse is unique – multiple particles passing through sensor, different levels of noise, different width, etc.

» Individual signal windows can be classified according to its features

» Clustering similar shapes allows for outlier detection and analyzing pileup

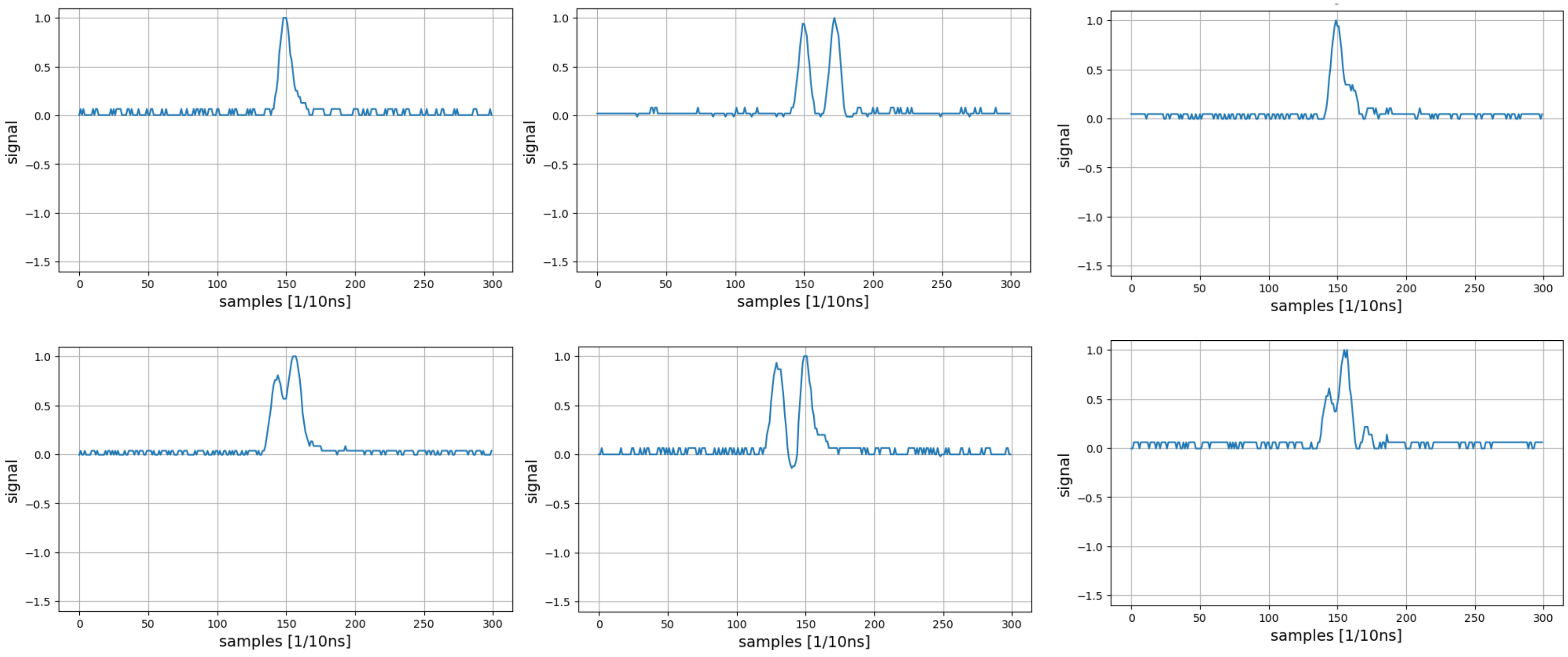


Fig11: Example of the signal shapes

Unsupervised machine learning

- Deep Neural Networks (DNNs) have gained significant attention in various fields, including physics, due to their ability to learn complex patterns and make predictions from large datasets
- Deep Neural Networks learn from training data by adjusting the weights of interconnected neurons through iterative improvement from defined error function.
- Learning can be:
  - Supervised – rely on labeled values (truth values)
  - Unsupervised – focuses on uncovering relationships within data

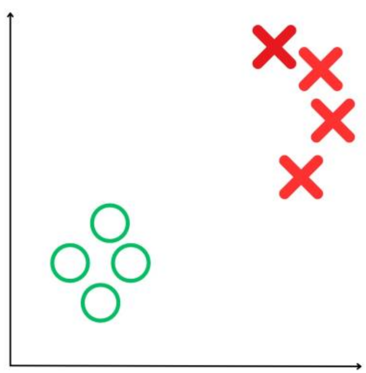


Fig11: Supervised learning (classification)

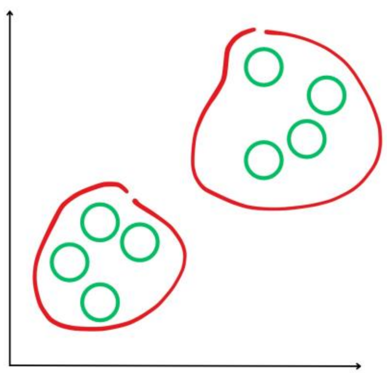


Fig11: Unsupervised learning (clusterization)

Autoencoder Neural Network

- Not every sample and feature within the micro-pulse is important for the purpose of clusterization
- Autoencoder was used to extract most important features of the micro-pulse
- Autoencoders does not require supervision and are excellent tool for extracting data features and noise reduction

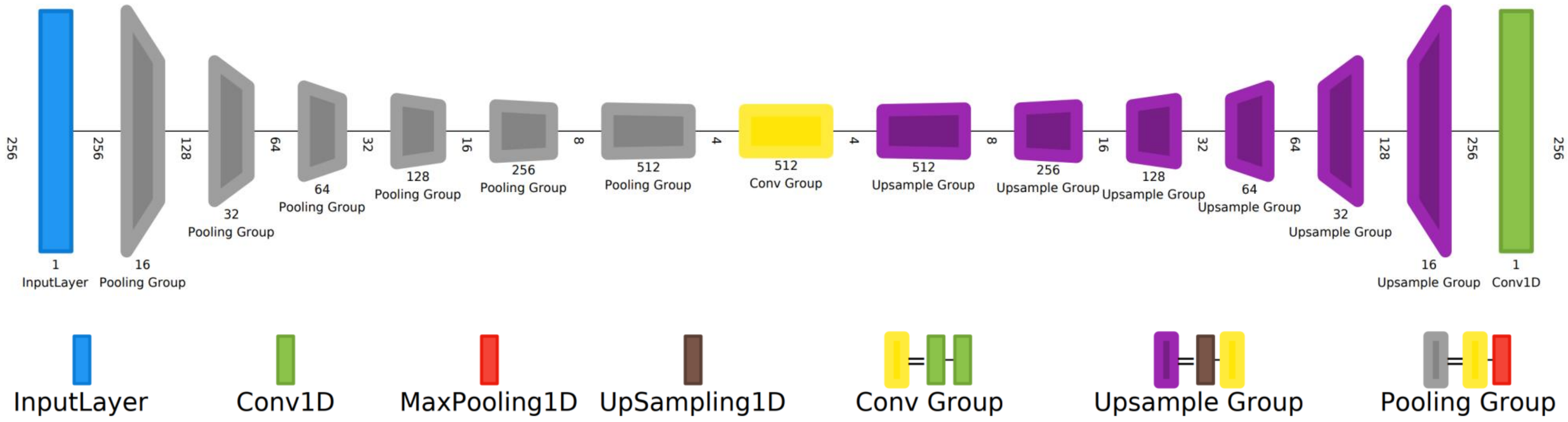


Fig12: Autoencoder architecture

Autoencoder Neural Network

» Trained autoencoder consists of two main components: encoder and decoder

» Decoder can be detached and used individually to reduce dimensionality of the pulses while preserving most important features

» Side effect of the features compression in the encoder is noise suppression

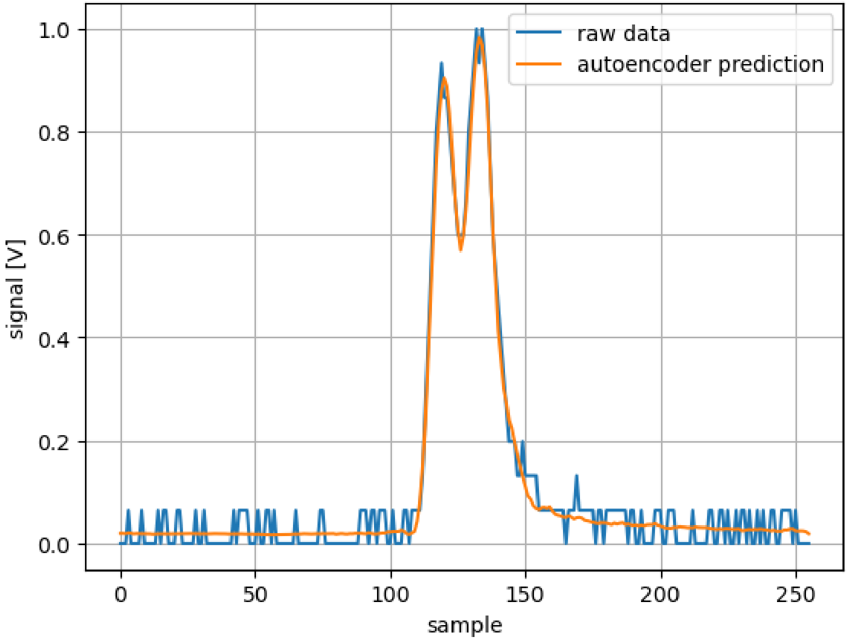


Fig13: Comparison between raw data samples and autoencoder prediction

Clusterization

» Clusterization is a form of unsupervised training that allows to group similar results according to the provided features

» Multiple methods were evaluated – KMeans, DBScan, GaussianMixture

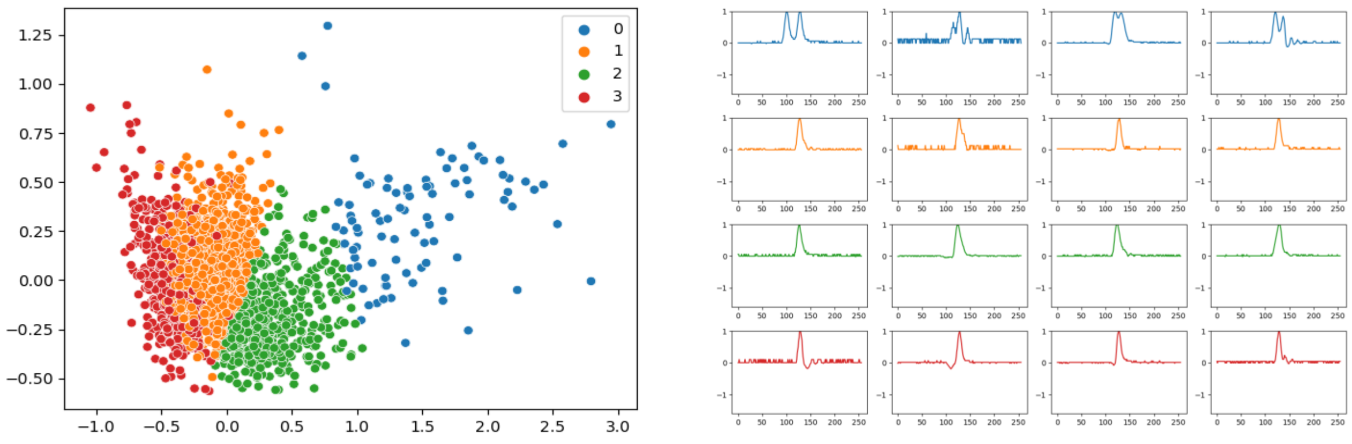


Fig14: Results of the classification on 2nA data with pileup cluster indicated by blue color and PCA projection

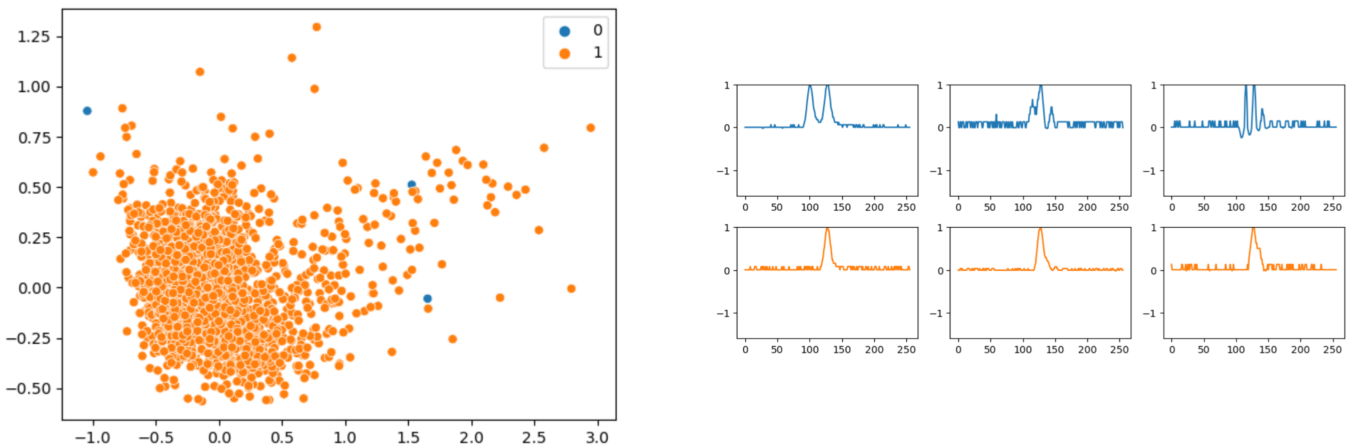


Fig15: Results of the classification on 2nA data with outliers indicated by blue color and PCA projection

Conclusions

- Machine learning techniques can be used in signal processing for analyzing signal shapes
- Separation between signals performs poorly without extracting underlying features
- Autoencoder Neural Networks allow for unsupervised extraction of the signal features, signal interpolation and noise reduction
- After feature extraction, suitable clusterization method can be used to group the data according to the need (we were concentrating on the number of peaks in the signal)
- Further studies concentrated on data with higher pileup are needed

sano

This project has received funding from the European Union's Horizon Europe Research and Innovation programme under Grant Agreement No 101057511 (EURO-LABS).

This project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 857533 and from the International Research Agendas Programme of the Foundation for Polish Science No MAB PLUS/2019/13.

European Funds
Smart Growth

Republic of Poland

FNP Foundation for Polish Science

European Union
European Regional Development Fund

Bibliography

- T. Isidori, P. McCavana, B. McClean, R. McNulty, N. Minafra, N. Raab, L. Rock, and C. Royon. Performance of a low gain avalanche detector in a medical linac and characterisation of the beam profile. Physics in Medicine amp; Biology, 66(13):135002, 2021.
- A. Bäuerle, C. van Onzenoodt and T. Ropinski, "Net2Vis – A Visual Grammar for Automatically Generating Publication-Tailored CNN Architecture Visualizations," in IEEE Transactions on Visualization and Computer Graphics, vol. 27, no. 6, pp. 2980-2991, 1 June 2021, doi: 10.1109/TVCG.2021.3057483.