Development and explainability of models for machine-learning-based signal reconstruction

Kalina Dimitrova[1], Venelin Kozhuharov[1], Peicho Petkov[1]
[1]*Faculty of Physics, Sofia University "St. Kliment Ohridski"*

Introduction

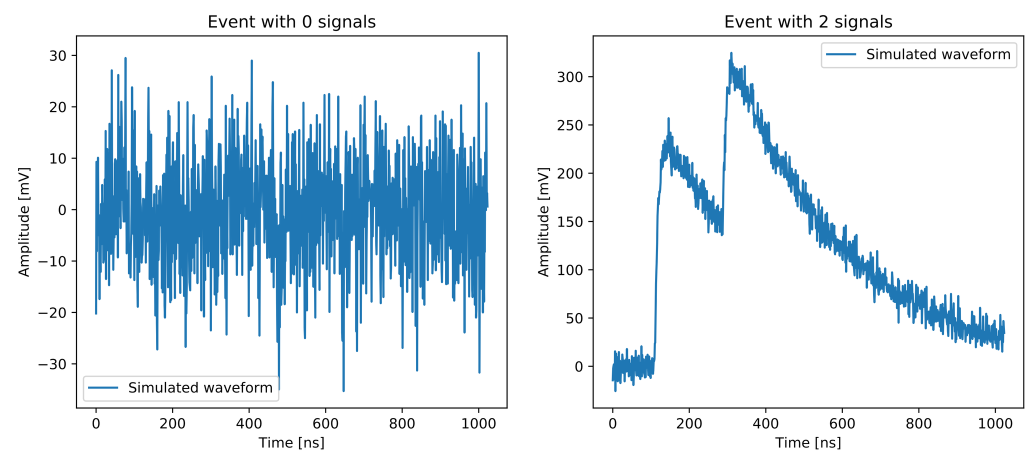
- Many detectors of particle physics experiments suffer from event pileups and need reliable algorithms for the separation of close-in-time signals. One possible solution is applying machine learning methods.
- Autoencoders are widely used for timeseries data description. The desired output can be a modified version of the input data with preserved dimensions. It contains all the information about each individual pulse and turns it into a case of supervised learning.
- Explainability methods (xAI) can provide a deeper look into the network mechanisms. Different approaches can be taken, and the results aid to developing even more accurate models.

Modified autoencoder

Signal simulation

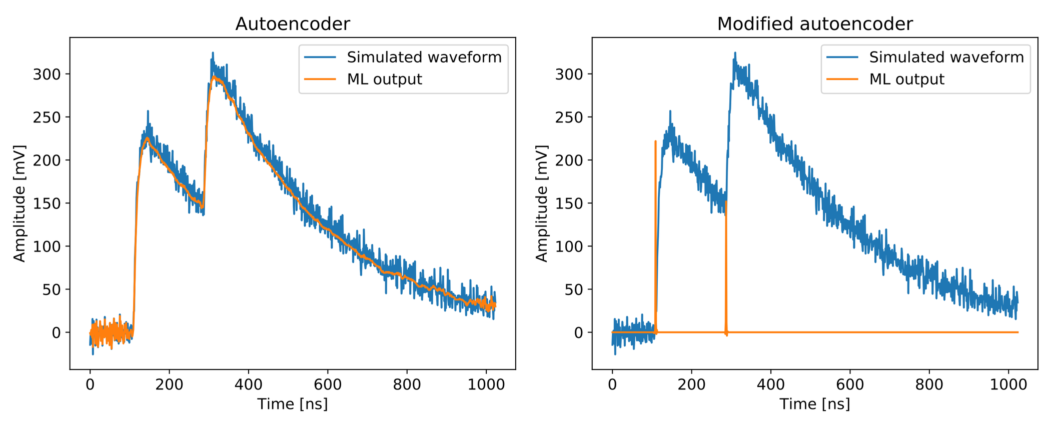
The generated events for training have a fixed length of 1024 ns and include several pulses (with a predefined maximum number). Each pulse has a 10 ns rise time and 300 ns fall time. Each pulse is assigned a random arrival time and has an amplitude above 20 mV with the amplitudes following a Gaussian distribution with a 200 mV mean value and σ=200 mV. Noise is added, taken as white noise – gaussian distribution with 10 mV mean.

$$A(t) = A_0\left(e^{\frac{-(t-t_0)}{\tau_1}} - e^{\frac{-(t-t_0)}{\tau_2}}\right), \quad t \geq t_0, \qquad (1)$$



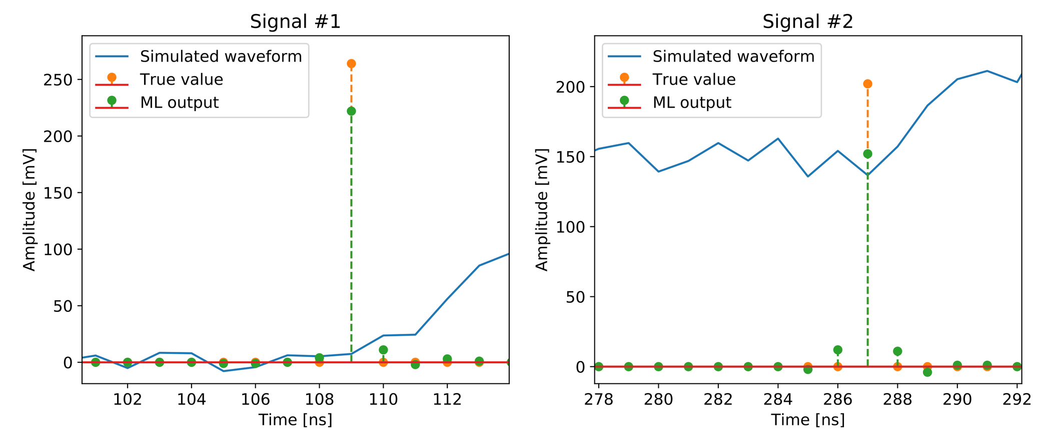
Modified autoencoder networks

The network architecture consists of three convolutional layers, followed by three deconvolutional layers with the same parameters in reverse order. The last layer is a deconvolutional layer with a single filter.



The introduced modification uses labels for each event with the same length as the data. All of the values are set to 0, except for the signal arrival positions. On those, the value is set to the amplitude of the signal.

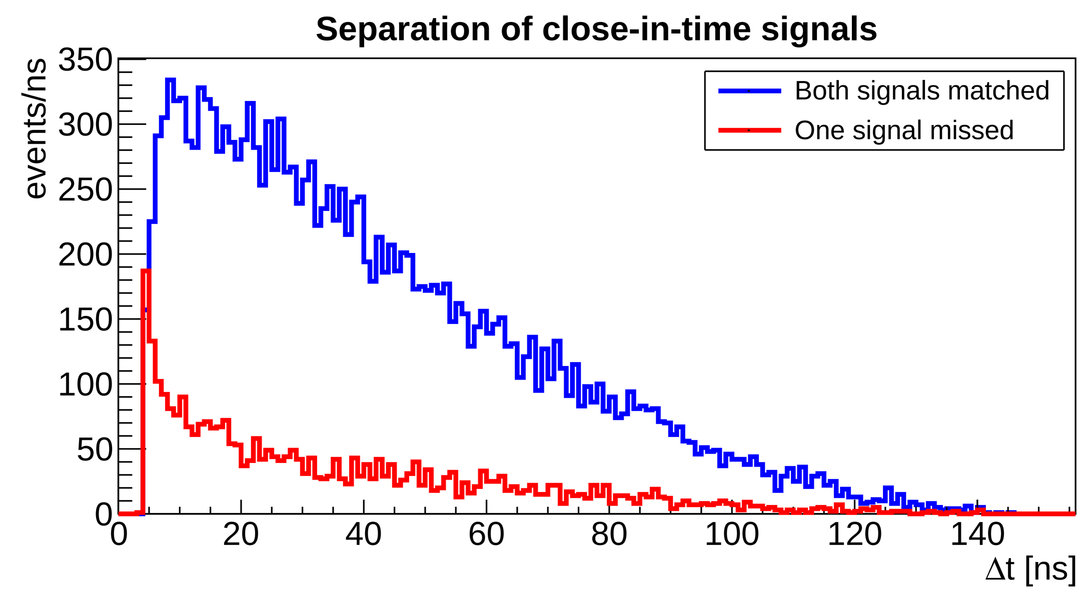
Post-processing

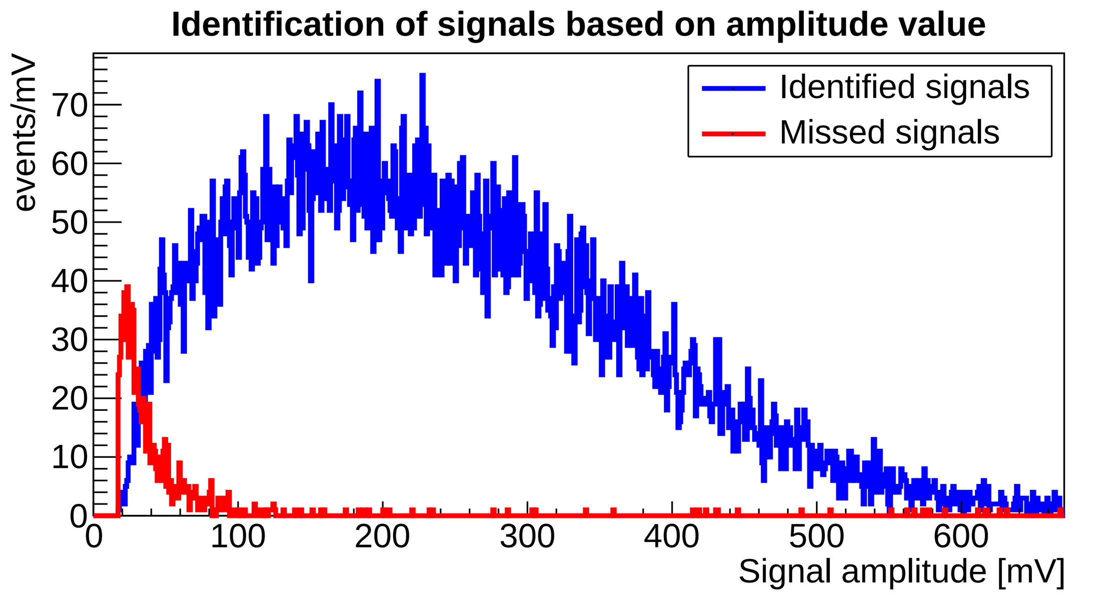


Unlike the labels, the output of the network has non-zero values on the positions before and after the main signal arrival prediction. The post-processing of the result includes defining a merging window, within which all non-zero values are merged into the central (maximum) one.

Network performance

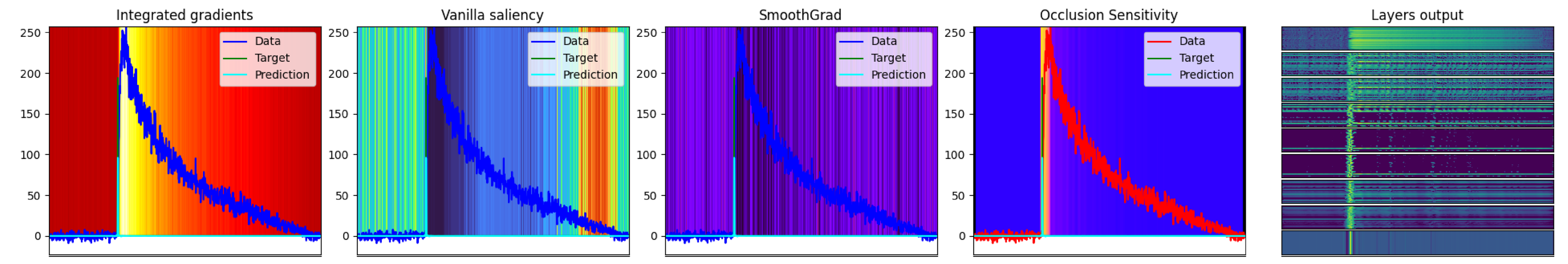
Most of the missed events are either less than 10 ns away from another event, or with amplitudes, smaller than 50 mV. For each recognised pulse, the accuracy for the arrival time prediction is of the order of 0.5 ns. The predicted amplitude value has lower values than the truth.





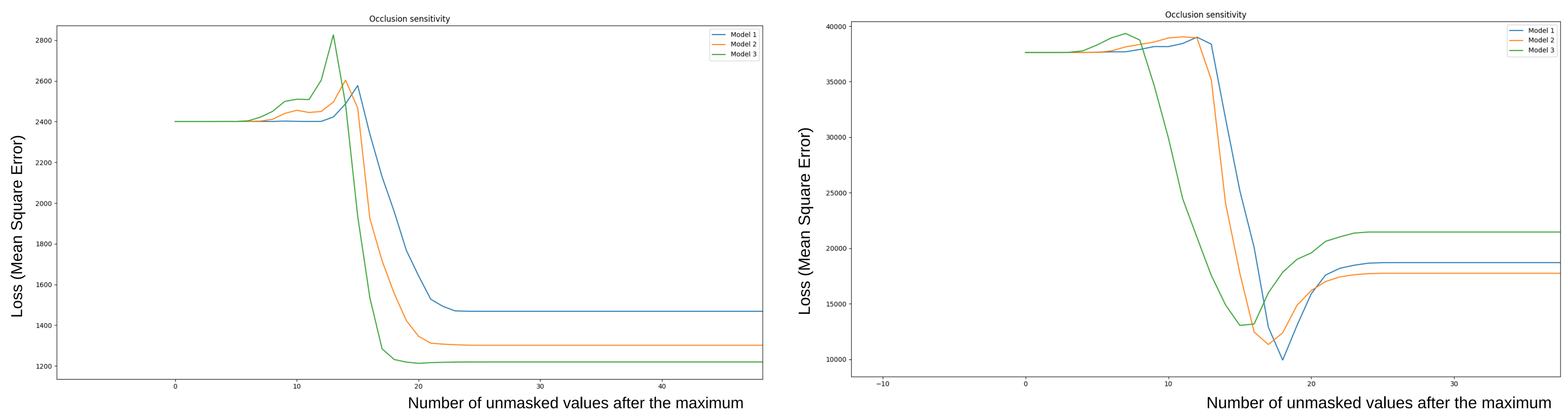
Explainability investigation and upscaling

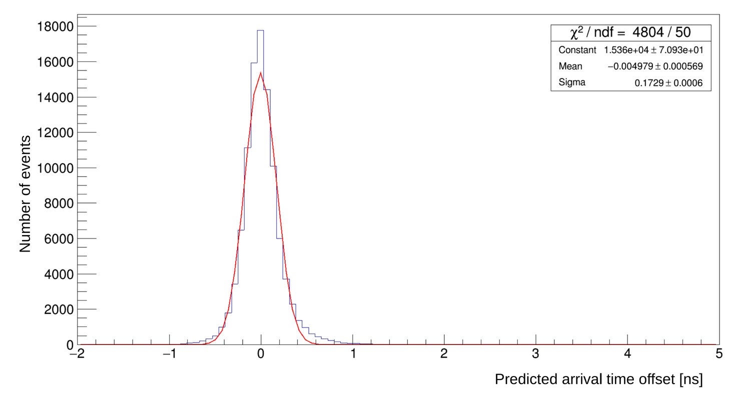
Five different explainability methods were tested on the predictions of the modified autoencoder networks. Plotting the layers output is important for testing the efficiency of different architectures. The integrated gradients, vanilla saliency and SmoothGrad methods don't show promising results, so the study is concentrated on the results of applying the occlusion senstitivity method.



In the case of occlusion sensitivity, different parts of the signal are masked to find the most important regions in the data, without which the algorithm cannot identify the pulse. Results show that the network needs the signal rise and the next 18 values in order to achieve the lowest possible loss.

Two types of events of events emerge: either the loss reaches a minimal value and remains constant or it rises after reaching the minimum.





This points that the exact arrival time of the signal within the 1 ns bin is important for the prediction. Instead of merging into the central bin, the arrival time was determined as a weighted average:

$$t_{arrival} = \frac{\sum A_i t_i}{\sum A_i} \qquad (2)$$

Upsampled models were developed where the output is 4 times larger than the input, allowing for 0.25 ns precision in the arrival time position. This results in a close to 0 deviation of the predicted from the true value, compared to an average of 0.6 ns for non-upsampled models.

Conclusion

Convolutional neural networks with modified autoencoder architectures can successfully separate close-in-time signals in particle detector data. Applying upsampled models can sufficiently improve the accuracy of arrival time reconstruction which can be of great importance for experiments with big event pileups.

Acknowledgements

This work is partially supported by BNSF: KP-06-D002_4/15.12.2020 within MUCCA, CHIST-ERA-19-XAI-009 and by the European Union - NextGenerationEU, through the National Recovery and Resilience Plan of the Republic of Bulgaria, project SUMMIT BG-RRP-2.004-0008-C01