A Convolutional Neural Network Algorithm for Automated Processing of $LaCl_3$ Scintillator Waveforms Incorporating Pulse Shape Discrimination and Pile-up Recovery

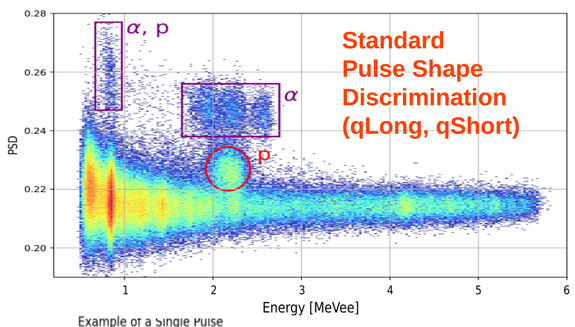
F. Guiotto[2,4,5], G. Guarino[1], D. Rigamonti[2], A. Dal Molin[2], G. Croci[1,2,3], C. Cazzaniga[6], M. Tardocchi[2,3], M. Nocente[1,2], E. Perelli Cippo[2], M. Rebai[2], G. Gorini[1,2,3], P. Franz[4], M. Zuin[4], P. Bettini[4,5], A. Muraro[2]

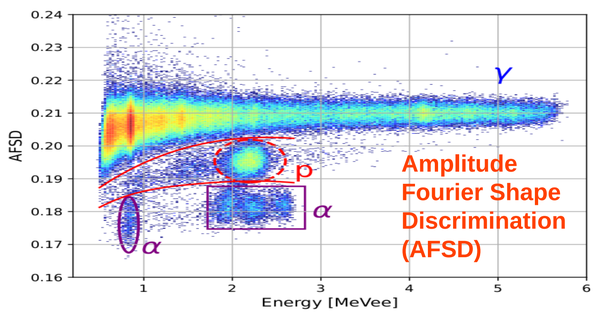
[1] *Dipartimento di Fisica 'G. Occhialini', Università di Milano-Bicocca, Piazza della Scienza 3, Milano, Italia*
[2] *Istituto per la Scienza e Tecnologia dei Plasmi, Consiglio Nazionale delle Ricerche, Via Roberto Cozzi 53, Milano, Italia*
[3] *INFN, Sezione Milano-Bicocca, Piazza della Scienza 3, Milano, 20125, Italia*
[4] *Consorzio RFX (CNR, ENEA, INFN, Università di Padova, Acciaierie Venete), Corso Stati Uniti 4, Padova, Italia*
[5] *Centro Ricerche Fusione (CRF) - University of Padova, C.so Stati Uniti 4, 35127,Padova, Italy*
[6] *UKRI-STFC, Rutherford Appleton Laboratory, Didcot, OX11 0QX, United Kingdom*

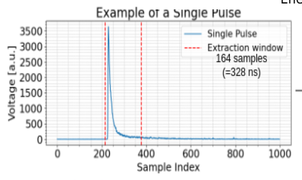
Corresponding author email:
federico.guiotto@phd.unipd.it

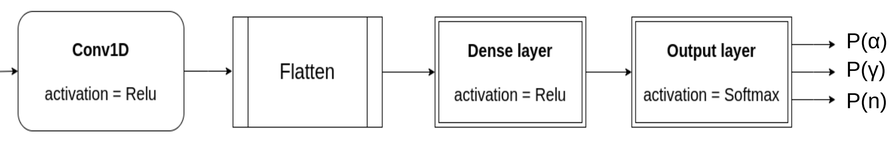
INTERNATIONAL CENTER PIERO CALDIROLA — ICPC

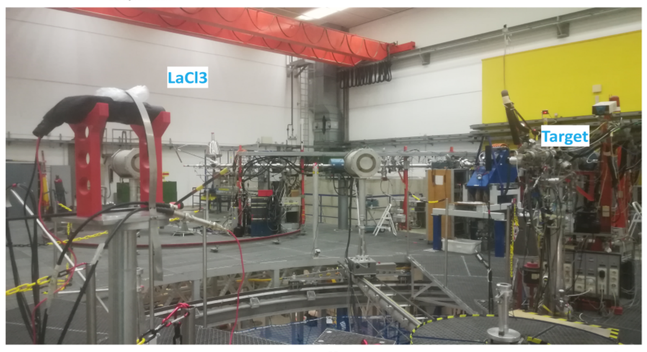
PSD algorithm for $LaCl_3$



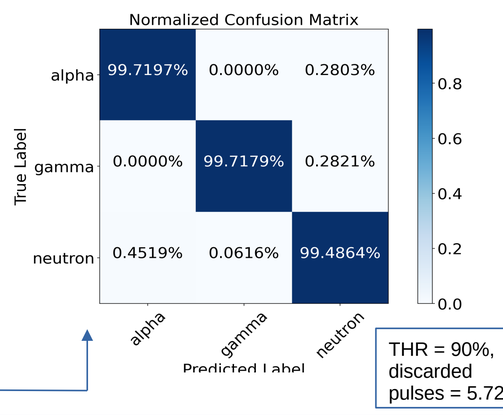






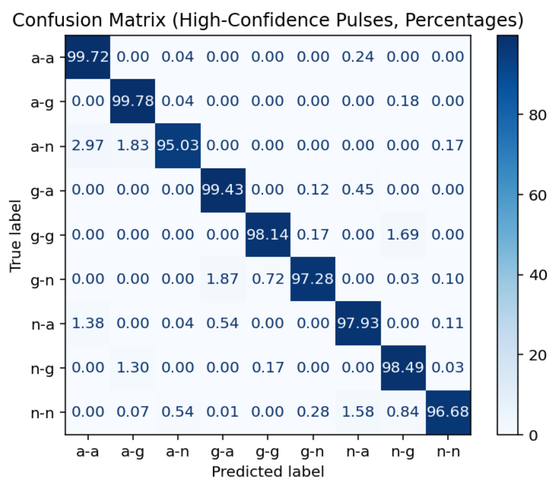


- **Low Scatter Measurement Hall at PTB, Germany → Experimental data used for training and validation.**
- **Testing dataset measured at the NILE facility, UK**



THR = 90%, discarded pulses = 5.72%

Close pulses discrimination



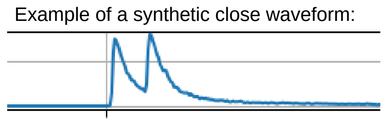
"High confidence" threshold: Predicted probability > 85%

- **Percentage of pulses discarded** (wrt the total): **14.46%**
- **Percentage of neutron-containing wfs discarded** (wrt the total number of neutrons): **19.21%**
- Percentage of alpha-containing wfs discarded(wrt the total number of alphas): 10.93%
- Percentage of gamma-containing wfs discarded(wrt the total number of gammas): 13.58%

Classification Report (High-Confidence Pulses):

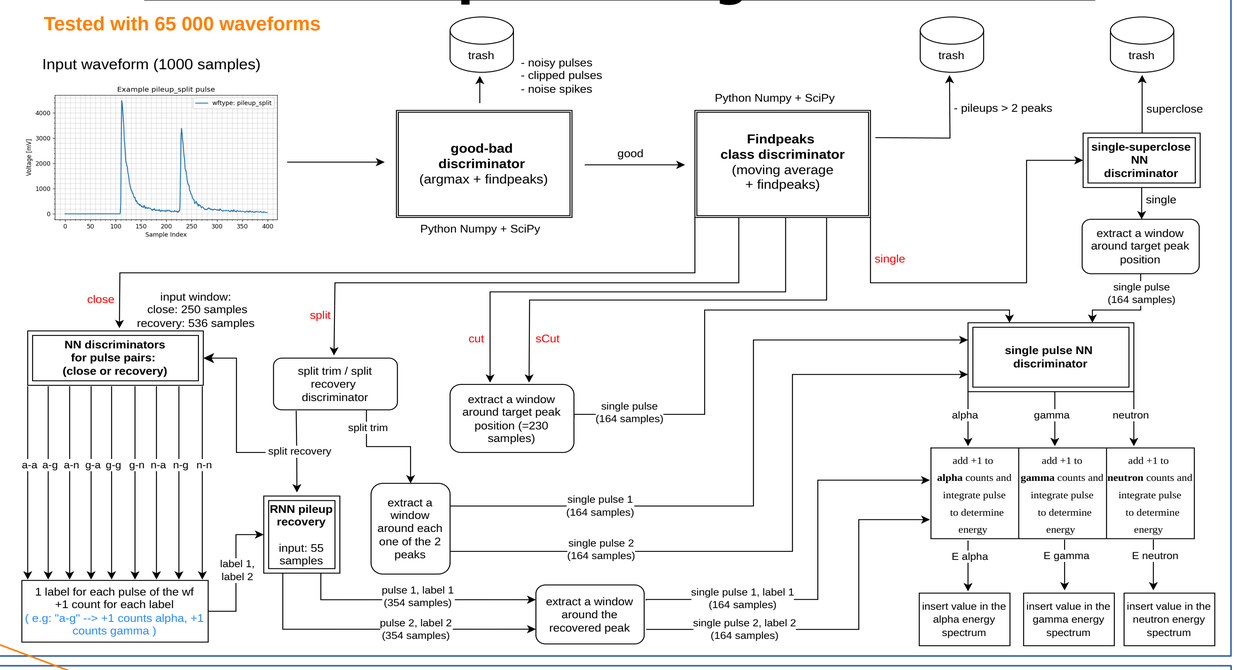
	precision	recall	f1-score	support
a-a	0.9323	0.9972	0.9637	2541
a-g	0.9158	0.9978	0.9551	5082
a-n	0.9821	0.9503	0.9660	4047
g-a	0.9047	0.9943	0.9474	4924
g-g	0.9987	0.9814	0.9900	188638
g-n	0.9857	0.9728	0.9792	26223
n-a	0.9539	0.9793	0.9664	4628
n-g	0.8975	0.9849	0.9392	29368
n-n	0.9961	0.9668	0.9812	12123
accuracy			0.9805	277574
macro avg	0.9519	0.9805	0.9653	277574
weighted avg	0.9819	0.9805	0.9808	277574

Example of a synthetic close waveform:

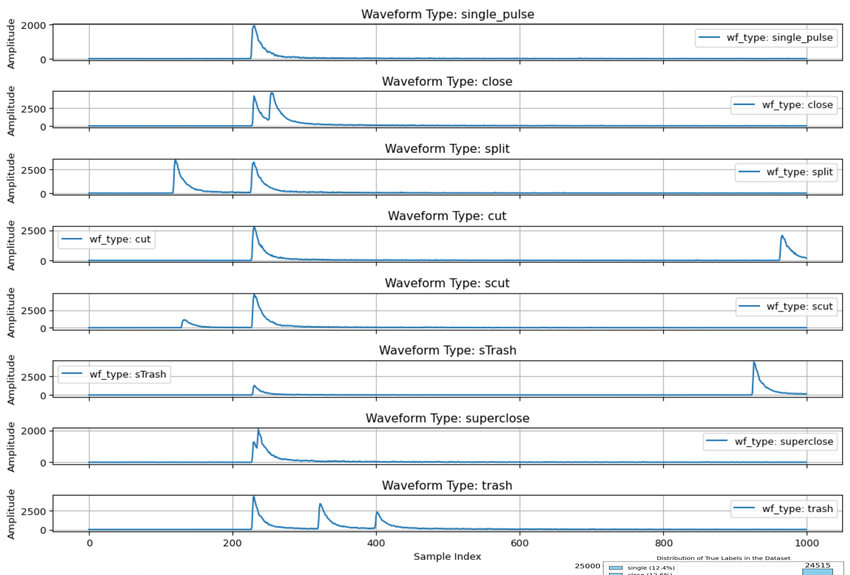


Automated processing architecture

Tested with 65 000 waveforms

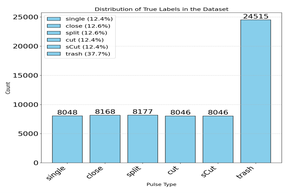


Types of raw waveforms

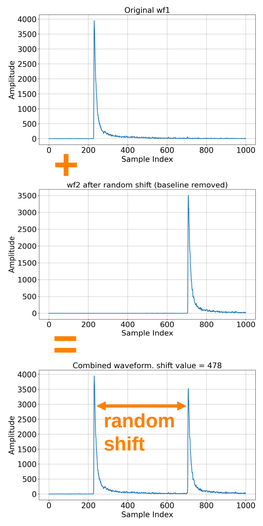


This classification of waveforms and their subsequent processing were inspired by the following work:
Fu, C., Di Fulvio, A., Clarke, S. D., Wentzloff, D., Pozzi, S. A., & Kim, H. S. (2018). "Artificial neural network algorithms for pulse shape discrimination and recovery of piled-up pulses in organic scintillators". Annals of Nuclear Energy, 120, 410–421.

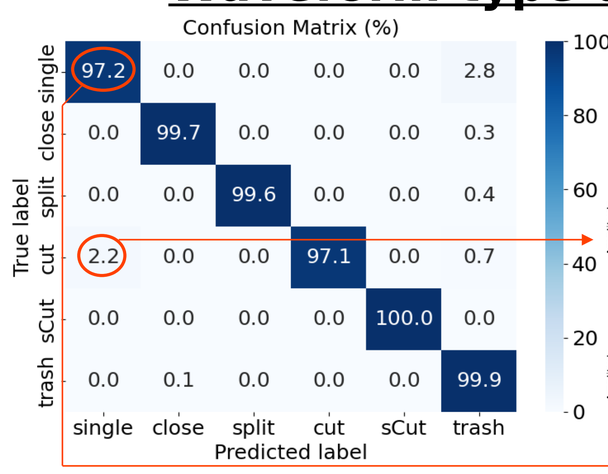
balanced dataset: "trash" contains sTrash, superclose, noisy, clipped, and spike pulses



synthetic dataset created for testing the automated processing architecture

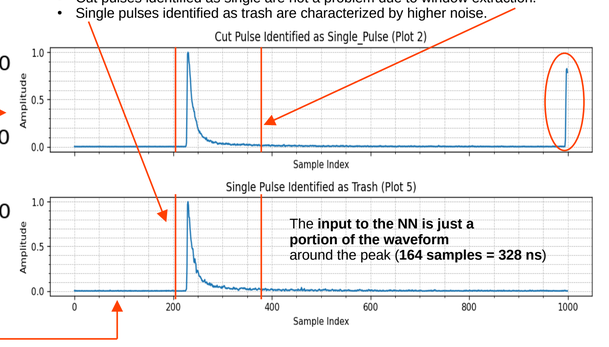


Waveform type discrimination

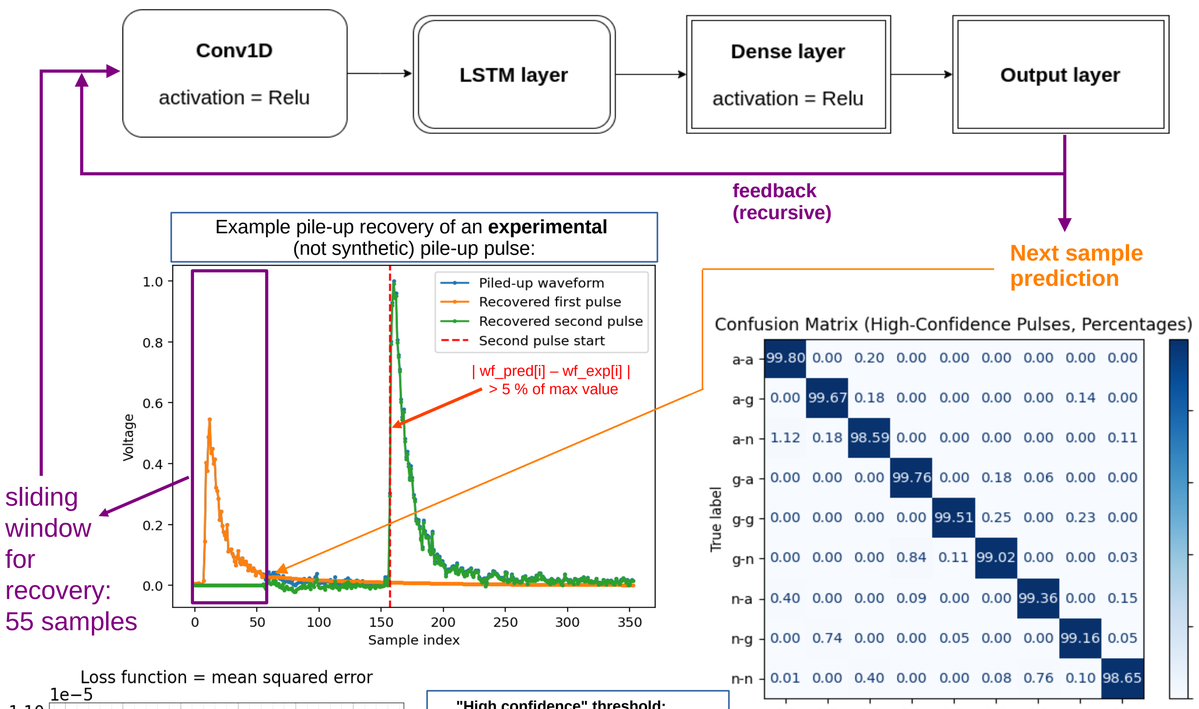


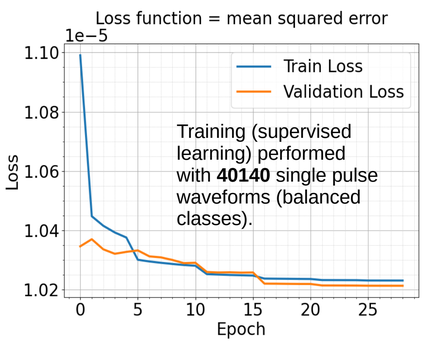
Number of test waveforms = 65 000. Good performance in discriminating the waveform type. Moreover, misclassified events are not problematic:

- Pulses mistakenly identified as trash are simply thrown away without polluting the dataset.
- Trash pulses identified as close are superclose pulses misclassified as close pulses. Maybe they are actually distinguishable (in any case, just 0.1%).
- Cut pulses identified as single are not a problem due to window extraction.
- Single pulses identified as trash are characterized by higher noise.



Pile-up recovery and discrimination





"High confidence" threshold: Predicted probability > 85%

- **Percentage of pulses discarded (wrt the total): 13.15%**
- **Percentage of neutron-containing wfs discarded** (wrt the total number of neutrons): **16.73%**
- Percentage of alpha-containing wfs discarded(wrt the total number of alphas): 8.85%
- Percentage of gamma-containing wfs discarded(wrt the total number of gammas): 12.57%

Classification Report (High-Confidence Pulses):

	precision	recall	f1-score	support
a-a	0.9733	0.9980	0.9855	2555
a-g	0.9563	0.9967	0.9761	4921
a-n	0.9846	0.9859	0.9852	4467
g-a	0.9522	0.9976	0.9744	4993
g-g	0.9998	0.9951	0.9974	188156
g-n	0.9833	0.9902	0.9867	29212
n-a	0.9776	0.9936	0.9856	4705
n-g	0.9843	0.9916	0.9879	29157
n-n	0.9973	0.9865	0.9919	13674
accuracy			0.9938	281840
macro avg	0.9787	0.9928	0.9856	281840
eighted avg	0.9939	0.9938	0.9938	281840

Recovered energy spectra

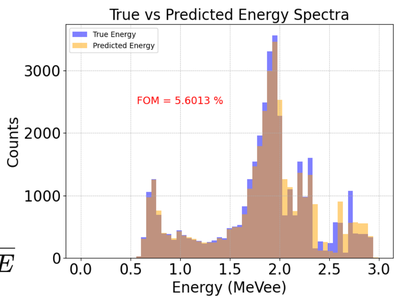
$h_1(E)$: bin counts at energy E of the true histogram

$h_2(E)$: bin counts at energy E of the true histogram

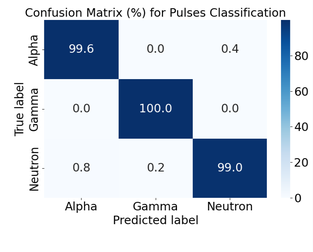
$$\mathrm{FOM}[\%] = \frac{100}{N_{\mathrm{bins}}} \sum_E \left| \widetilde{h}_1(E) - \widetilde{h}_2(E) \right|$$

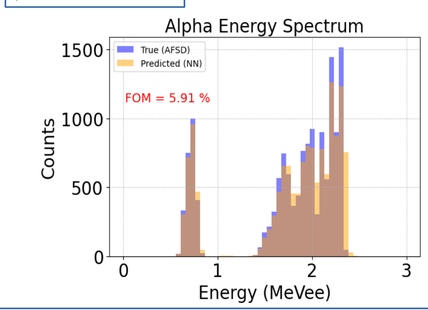
with : $\widetilde{h}_i(E) = \dfrac{h_i(E)}{\sum_E h_i(E)\,\Delta E}$

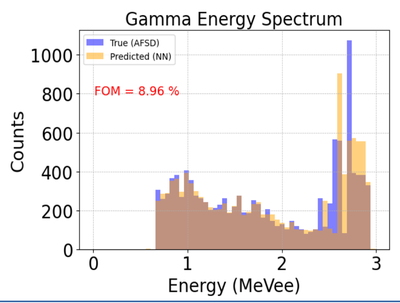
total histogram counts = total number of processed **peaks** with energy information preserved = **37351**

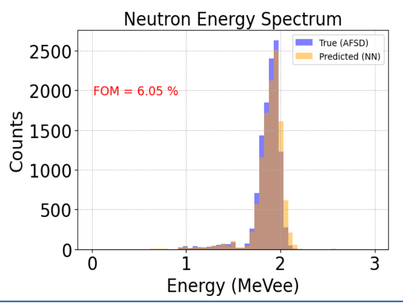


Full architecture discrimination performance:









Conclusions

An algorithm based on Convolutional Neural Networks (CNNs) has been developed for automated processing of data from measurements performed with a $LaCl_3$ scintillator. The algorithm processes raw input waveforms and classify them based on their type, rejecting those identified as potentially dangerous for dataset pollution. Pile-up waveforms are identified and recovered with a Recursive Neural Network to become individual pulses. The only exception are the waveforms of type "close pile-up", which are only labeled without determining the energy information of their pulse pairs. Different CNNs predict the particle type associated to each pulse, selecting the label among "alpha", "gamma", and "neutron". Training and testing of the CNNs were performed using experimental data from two different neutron facilities: PTB (Germany) and NILE (UK). The algorithm will be tested in a nuclear fusion environment using data acquired from the MAST-U experiment (UK).

Project funded under the National Recovery and Resilience Plan (NRRP) in the framework of the European Union – NextGenerationEU, Mission 4, Component 2, Investment 3.1 - Area ESFRI Energy - Call for tender No. 3264 of 28-12-2021 of Italian University and Research Ministry (MUR), Project ID IR0000007, MUR Concession Decree No. 243 del 04/08/2022, CUP B53C22003070006, "NEFERTARI – New Equipment for Fusion Experimental Research and Technological Advancements with Rfx Infrastructure".

Funded by the European Union
NextGenerationEU

UNIVERSITÀ DEGLI STUDI DI PADOVA — INFN Istituto Nazionale di Fisica Nucleare — CONSORZIO RFX Ricerca Formazione Innovazione — ISTP ISTITUTO PER LA SCIENZA E TECNOLOGIA DEI PLASMI — UNIVERSITÀ DEGLI STUDI DI MILANO BICOCCA

International Conference on Diagnostics for Fusion Reactors: The Burning Plasma Era (ICFRD2025)