

A ResNeXt50-based convolution neural network for nuclear reaction classification in an active target TPC detector

HAVERSON, Kristian, SMITH, Robin <<http://orcid.org/0000-0002-9671-8599>>, SHENFIELD, Alex <<http://orcid.org/0000-0002-2931-8077>>, FORTULAN, Raphael, ĆWIOK, M, DOMINIK, W, FIJAŁKOWSKA, M, FILA, M, JANAS, Z, KALINOWSKI, A, KIERZKOWSKI, K, KUICH, M, MAZZOCCHI, C, OKLIŃSKI, W, ZAREMBA, M, GAI, M, SCHWEITZER, D.K., STERN, S.R., FINCH, S, FRIMAN-GAYER, U, JOHNSON, S.R., KOWALEWSKI, T.M., BALABANSKI, D.L., MATEI, C, ROTARU, A, ALLEN, R.A.M., GRIFFITHS, M.R., KOKALOVA, Tz, PIRRIE, S, SANTA RITA ALCIBIA, P and WHELDON, C

Available from Sheffield Hallam University Research Archive (SHURA) at:

<https://shura.shu.ac.uk/34667/>

This document is the Published Version [VoR]

Citation:

HAVERSON, Kristian, SMITH, Robin, SHENFIELD, Alex, FORTULAN, Raphael, ĆWIOK, M, DOMINIK, W, FIJAŁKOWSKA, M, FILA, M, JANAS, Z, KALINOWSKI, A, KIERZKOWSKI, K, KUICH, M, MAZZOCCHI, C, OKLIŃSKI, W, ZAREMBA, M, GAI, M, SCHWEITZER, D.K., STERN, S.R., FINCH, S, FRIMAN-GAYER, U, JOHNSON, S.R., KOWALEWSKI, T.M., BALABANSKI, D.L., MATEI, C, ROTARU, A, ALLEN, R.A.M., GRIFFITHS, M.R., KOKALOVA, Tz, PIRRIE, S, SANTA RITA ALCIBIA, P and WHELDON, C (2024). A ResNeXt50-based convolution neural network for nuclear reaction classification in an active target TPC detector. *Il Nuovo Cimento*, 48 (1): 17. [Article]

Copyright and re-use policy

See <http://shura.shu.ac.uk/information.html>

A ResNeXt50-based convolution neural network for nuclear reaction classification in an active target TPC detector

K. C. Z. HAVERSON⁽¹⁾, R. SMITH^{(1)(2)(*)}, A. SHENFIELD⁽¹⁾, R. FORTULAN⁽³⁾,
M. WIOK⁽⁴⁾, W. DOMINIK⁽⁴⁾, A. FIJAŁKOWSKA⁽⁴⁾, M. FILA⁽⁴⁾, Z. JANAS⁽⁴⁾,
A. KALINOWSKI⁽⁴⁾, K. KIERZKOWSKI⁽⁴⁾, M. KUICH⁽⁴⁾, C. MAZZOCCHI⁽⁴⁾,
W. OKLIŃSKI⁽⁴⁾, M. ZAREMBA⁽⁴⁾, M. GAI⁽²⁾, D. K. SCHWEITZER⁽²⁾,
S. R. STERN⁽²⁾, S. FINCH⁽⁵⁾⁽⁶⁾, U. FRIMAN-GAYER⁽⁵⁾⁽⁶⁾, S. R. JOHNSON⁽⁷⁾⁽⁶⁾,
T. M. KOWALEWSKI⁽⁷⁾⁽⁶⁾, D. L. BALABANSKI⁽⁸⁾, C. MATEI⁽⁸⁾, A. ROTARU⁽⁸⁾,
R. A. M. ALLEN⁽⁹⁾, M. R. GRIFFITHS⁽⁹⁾, Tz. KOKALOVA⁽⁹⁾, S. PIRRIE⁽⁹⁾,
P. SANTA RITA ALCIBIA⁽⁹⁾ and C. WHELDON⁽⁹⁾

⁽¹⁾ *Department of Engineering and Mathematics, Sheffield Hallam University - Howard Street, Sheffield, S1 1WB, Bristol, UK*

⁽²⁾ *Laboratory for Nuclear Science at Avery Point, University of Connecticut - Groton, CT, 06340-6097, USA*

⁽³⁾ *Unconventional Computing Laboratory, University of the West of England - Bristol, UK*

⁽⁴⁾ *Faculty of Physics, University of Warsaw - Warsaw, Poland*

⁽⁵⁾ *Physics Department, Duke University - Durham, NC, USA*

⁽⁶⁾ *Triangle Universities Nuclear Laboratory - Durham, NC, USA*

⁽⁷⁾ *Department of Physics & Astronomy, University of North Carolina Chapel Hill, NC, USA*

⁽⁸⁾ *IFIN-HH / ELI-NP - Bucharest-Magurele, Romania*

⁽⁹⁾ *School of Physics and Astronomy, University of Birmingham - Birmingham, B15 2TT, UK*

received 23 July 2024

Summary. — The Convolutional Neural Network (CNN), ResNeXt50, was used to classify nuclear reactions measured in a gas-filled charged particle detector. This active-target time-projection chamber, built by the University of Warsaw, is optimised for studying photo-dissociation reactions using intense γ -beams. In particular, the $^{16}\text{O}(\gamma, \alpha)^{12}\text{C}$ and $^{12}\text{C}(\gamma, 3\alpha)$ reactions were measured using γ -beams at the HI γ S facility at Duke University. Reactions of interest were distinguished from other background channels using the CNN. For this preliminary study, a small sample of two hundred hand-classified events per category were used to train the model. This was then used to classify an unseen data set comprising one hundred events in each category. This method was extremely effective at removing background events and could differentiate the two main reaction channels with 96% accuracy.

(*) Corresponding author. E-mail: robin.smith@shu.ac.uk

1. – Introduction

The use of high-intensity γ beams is a powerful technique both in astrophysics studies and explorations of nuclear structure. In order to study the processes that govern the evolution of stars, stellar modelling requires accurate knowledge of nuclear reaction cross sections, which decrease rapidly at lower centre-of-mass energies due to the Coulomb barrier. Since astrophysical energies are generally sub-Coulomb barrier, these correspond to very low cross sections, inaccessible through even the most sophisticated modern experimental approaches. Therefore, cross sections are typically measured at higher energies, where the cross section is greater, and then extrapolated to astrophysical energies at the Gamow peak using R -matrix theory.

Key reactions that occur in the core of a star are often direct α and p capture reactions, (α, γ) and (p, γ) . In experiments, the small cross sections that characterise these reactions at low energies can be boosted significantly by capitalising on the principle of detailed balance [1] and studying these reactions through their inverse photo-dissociation reactions. The photo-dissociation cross section is directly related to the capture cross section via the principle of detailed balance: $\sigma(\gamma, \alpha) = k_\alpha^2/2k_\gamma^2 \times \sigma(\alpha, \gamma)$ where $\hbar k_\alpha = \sqrt{2\mu E_{\text{cm}}}$, and $\hbar k_\gamma = \hbar\omega/c = E_\gamma/c$. This means that the photo-dissociation cross section can be significantly larger than the capture cross section.

In astrophysics, the flagship case is determination of the $^{12}\text{C}(\alpha, \gamma)^{16}\text{O}$ reaction cross section through measurement of the inverse $^{16}\text{O}(\gamma, \alpha)^{12}\text{C}$ [2]. This technique has been shown to offer several key advantages over measuring the reaction directly. In addition to higher cross sections (~ 50 times larger in this case), γ -induced reactions are characterised by lower experimental backgrounds and different systematic uncertainties.

On the nuclear structure side, such photo-dissociation reactions are an ideal platform on which to conduct precise measurements of exotic nuclear resonances. Selection rules restrict the spin-parity of states that may be populated via γ absorption, meaning this technique is extremely useful for exploring regions of high level density. In the past, α -particle clustering in ^{12}C was explored via the $^{12}\text{C}(\gamma, \alpha)$ reaction, using an Optical TPC detector [3]. These experiments provided the first unambiguous measurement of the 2^+ rotational excitation of the Hoyle state [4] and its decay branching ratios [5]. The data supported claims of D_{3h} point symmetry in ^{12}C as predicted by the Algebraic Cluster Model [6] and new measurements at higher energies may provide evidence for further rotational states in this nucleus.

In order for this experimental technique to be feasible, high-intensity, quasi-monochromatic γ -beams are required, as well as a sophisticated detection system that is able to accurately measure the momenta of the charged reaction products. In this study, the Warsaw active target TPC detector (Warsaw TPC) was used [7]. Although TPC detectors offer many advantages, as discussed later, they produce large amounts of data. A typical experiment generates upwards of 10 TB of data per week, which must be processed to obtain charge deposition and spatial information. Convolutional Neural Networks (CNNs) have recently shown excellent potential for automating the analysis of particle TPC detector data, for tasks such as event reconstruction, classification and particle identification [8].

Using Artificial Neural Networks for TPC data analysis is not a new concept, with liquid Argon TPC analysis being approached this way as far back as 1995 [9]. Machine Learning techniques in general have also been applied more recently at various facilities, including TexAT [10], JUNO [11], the ATTPC [12] and NeuLAND [13], to identify particle tracks within the data and to reconstruct energy and vertex information. These

techniques include boosted decision trees and using pre-trained deep neural networks for vertex and energy reconstruction. Most recently, residual neural network (ResNet) models were used successfully for event classification and reconstruction of ^{12}C decay in simulated data [8]. The work presented in this paper utilises the ResNeXt50 architecture for classifying photo-dissociation reactions measured using the Warsaw TPC. Initially the experimental facility and TPC detector are described. The types of data produced by the TPC are then discussed, followed by a description of the filtering that is performed on the data before inputting to the neural network. The ResNeXt50 architecture is described and its performance for event classification is quantified.

2. – Experiment details

Intense, quasi mono-energetic photon beams were produced at the High Intensity γ -ray Source (HI γ S) of the Triangle Universities Nuclear Laboratory [14]. As depicted in ref. [2], high energy gamma beams were generated through Compton back-scattering of free electron laser (FEL) light from relativistic electron beams. The gamma beam energy is controlled by the FEL wavelength and the electron energy. In 2022, new data were acquired using beam intensities of $\sim 5 \times 10^8$ γ /s with an energy spread of $\sim 3\%$, at nominal beam energies ranging from 8.6 to 13.9 MeV.

The Warsaw TPC [7] was used for charged particle detection. In the TPC, a low-pressure gaseous CO_2 active volume acts as both the target for the nuclear reactions and as the medium for particle detection, via charge transport and subsequent amplification. The TPC used in this work is the result of a programme led by the University of Warsaw in collaboration with the University of Connecticut, and the Extreme Light Infrastructure Nuclear Physics / IFIN-HH, Bucharest-Magurele, Romania. The detector comprises a drift cage, which permits uniformity of the drift field with $<1\%$ variation. Electrons generated through track ionisation drift in this field towards an anode plane and are amplified by three 50 μm -thick Gas Electron Multiplier (GEM) foils [15] with 3 mm separation. The amplified signals are collected by a segmented multi-layer PCB anode and are processed using GET electronics [16]. A redundant, 3-coordinate system readout is used, arranged along three independent axes oriented at 60° to each other. These form three linear sets of strips with 1.5 mm pitch, that are labelled (U, V, W). This arrangement allows the creation of virtual pixels on a 2D horizontal plane as a function of time, provided that the topologies of the measured tracks are relatively simple, as is the case for the data presented here. A set of example events may be seen in fig. 1.

The total collected charge, track lengths, interaction position, and dE/dx of the particles along each track may be used to differentiate reactions that occur inside the TPC to good effect. These include the reactions of interest, $^{16}\text{O}(\gamma, \alpha)^{12}\text{C}$ and $^{12}\text{C}(\gamma, 3\alpha)$, along with several background channels. The background reactions include, but are not limited to: $^{17/18}\text{O}(\gamma, \alpha)^{13/14}\text{C}$, $^{16}\text{O}(\gamma, p)^{15}\text{N}$, and $^{16}\text{O}(\gamma, n)^{15}\text{O}$. Standard analyses using the metrics described above perform well at separating the reactions. The aim of this work is to demonstrate a modern Convolutional Neural Network (CNN) as a complementary technique for reaction classification, with the possibility of increased speed and convenience over conventional techniques.

3. – Methodology

By analysing photo-dissociation data collected with the Warsaw TPC at the HI γ S facility during a 2022 experimental campaign, this study demonstrates the effectiveness

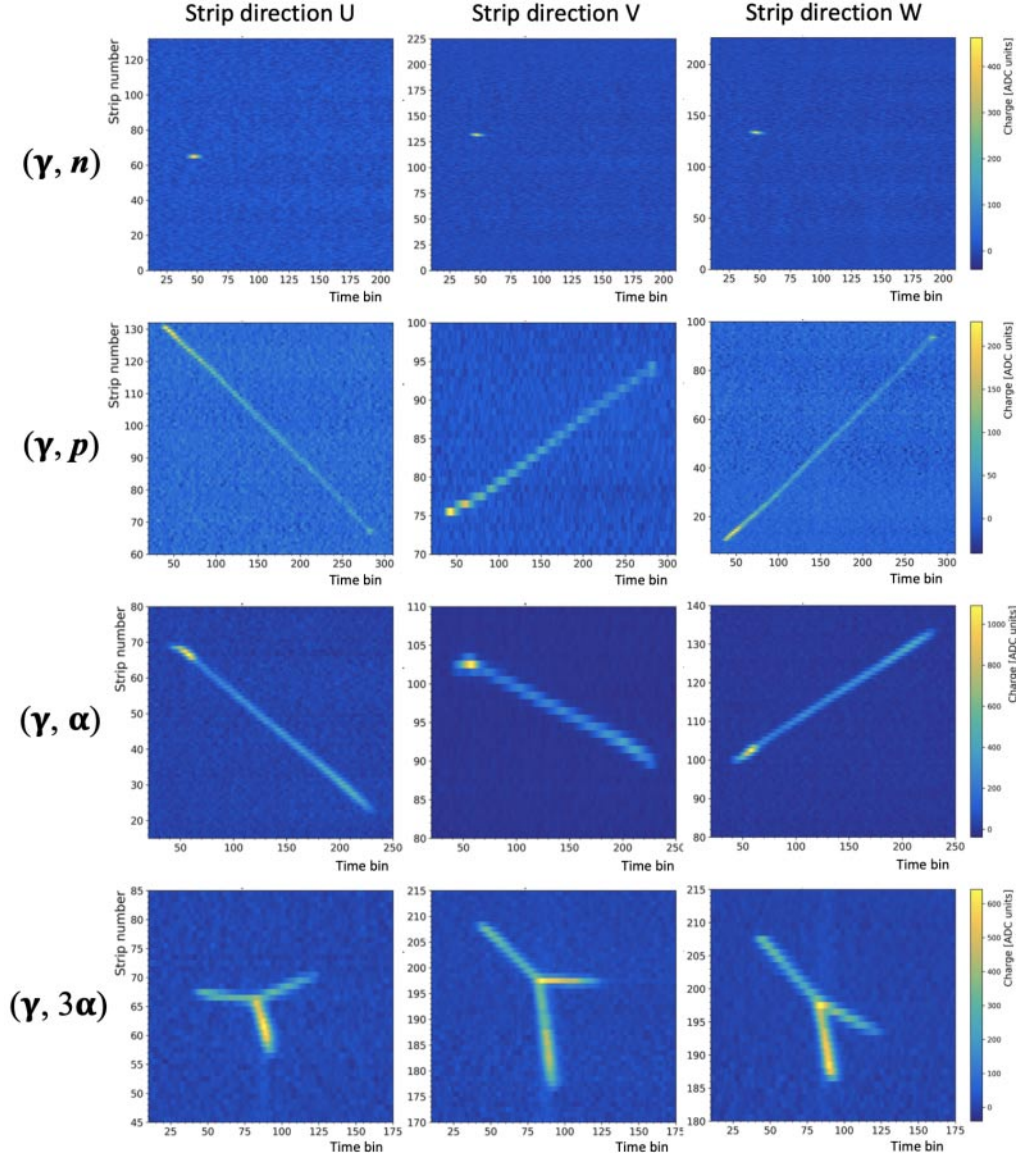


Fig. 1. – The manifestation of different categories of reactions in the U, V, W, t projections.

of machine learning-based classification through the use of a convolutional neural network (CNN) in identifying general nuclear reaction topologies. A robust methodology that leverages transfer learning has been established as an alternative approach for background removal and classification of nuclear decays that are challenging to distinguish through classical methods.

The data considered were sampled from a 13.9 MeV nominal γ beam energy. The TPC was operated with a sampling rate of 12.5 MHz, shaping time of 232 ns, and a pressure of 250 mbar. All data underwent the standard signal processing stages as outlined in ref. [7],

where the raw waveforms were individually pedestal- and fixed pattern noise-corrected. All events considered in the analysis had the same trigger delay. This study also applied two further corrections: a gain normalisation across all channels and a deconvolution. The channel-by-channel time deconvolution was implemented to account for the shaping caused by the electronics [19]. The response functions were determined by using an internal pulser with a known input signal and examining the corresponding output. A comparison has been made of the classifier’s performance when analysing the corrected data *versus* the non-corrected data.

As demonstrated in fig. 1, each recorded event is stored as a collection of three images, which show the charge collection on a family of strips (U, V, W) on the vertical axis, against drift time in 512 time bins on the horizontal axis. The number of channels is 264 for U, 376 for V and 378 for W. To eliminate remnant switching effects from the read-write electronics, only signals from the first 490 time bins were considered. This defines the length of our horizontal axis. The vertical axis is then zero-pad to create three 490×490 pixel images. The maximum charge collected in any projection is used to normalise the data, giving pixel values ranging from zero to unity. Finally, the normalised images are stacked each into a different RGB channel, resulting in a single $490 \times 490 \times 3$ image per event. An example event is shown in fig. 2.

The sample dataset used in this study was manually classified and reconstructed by experts, with checks applied for consistency. These were then further checked for eligibility and full containment within the TPC by placing fiducial cuts on the reconstructed vertex distance from the beam, and a 10 mm maximum distance from any track endpoint to the edge of the drift volume.

The training classes for the classification problem were defined as: (γ, n) , (γ, p) , (γ, α) , and $(\gamma, 3\alpha)$. No distinction was made between 3α decays that proceed via the ground state of ${}^8\text{Be}$ and first excited state, even though these types of events can manifest quite

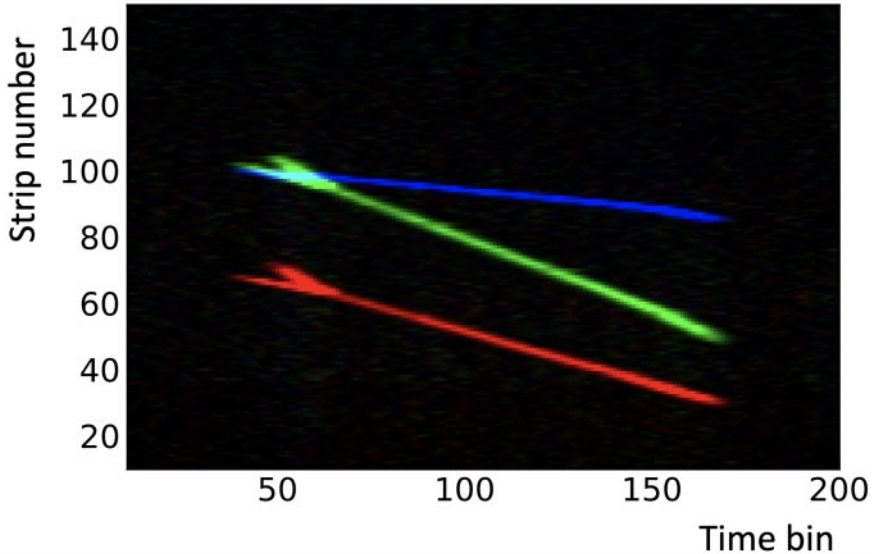


Fig. 2. – An example of the RGB image input to the CNN. This event is one of the ${}^{12}\text{C}(\gamma, \alpha)$ reactions that were ultimately mis-classified as ${}^{16}\text{O}(\gamma, \alpha)$.

differently in the data. This choice was made due to low statistics caused by a low 3α reaction cross-section at the energy considered. Further event classes that sometimes appear in the data include α particles from natural radioactivity and pile-up events, which should be included as separate categories in future work.

Each class examined in this work consisted of 300 events. A 2:1 split was used to train and validate the classifier, where 2/3 of each class were used for training and 1/3 (100 events) were reserved for validation. Transfer learning techniques leveraging a pre-trained ResNeXt-50 model were employed, which has already been trained on a large image database to be effective at detecting common visual features, such as edges and curves. ResNeXt is a deep convolutional neural network architecture [17] that is based upon the ResNet model, but uniquely incorporates groups of multiple parallel pathways within each layer. This multi-path group structure enhances the model’s ability to learn a broader range of more complex features [18]. This model was then fine-tuned using our training data, allowing it to adapt and learn to extract our application-specific features, which is typical in computer vision classification problems with low statistics. All specific hyper-parameters used for the training of the model can be found in table I.

4. – Results and discussion

The performance of the model for both standard and time-deconvolved data is shown in table II. The performance, when considering time-deconvolved data, shows an increase in terms of precision, recall, and F1 score, which was achieved at epoch 4, where validation accuracy first reached its maximum value.

Using the ResNeXt model to classify the validation set, 99% of neutron, proton, and alpha particle events were correctly classified for the deconvolved data, as captured in the confusion matrix of fig. 3. However, the classification performed less well for triple alpha events (3α). Here, 96% were correctly classified, with 2% misclassified as single alpha events. A lower 94% were correctly classified without the deconvolution. The kinematics of 3α decays, where particles can emerge with a small opening angle, makes certain distributions inherently difficult to distinguish from single alpha events in the TPC. See fig. 2 for an example of such an event. To overcome this, training on larger

TABLE I. – *Model hyper-parameters.*

Parameters	Values
Loss	Cross Entropy Loss
Optimiser	Stochastic Gradient Descent
Learning Rate	0.001
Momentum	0.9
Batch Size	16

TABLE II. – *Model performance.*

Model	Precision	Recall	F1 Score
Standard model	0.978	0.978	0.978
Deconvolved model	0.983	0.983	0.983

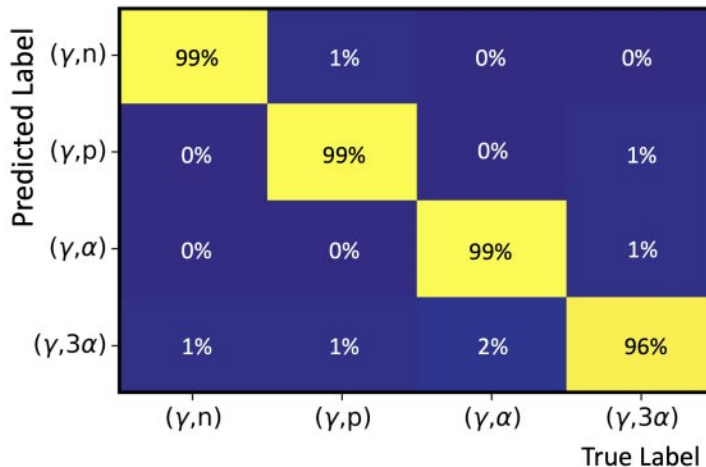


Fig. 3. – The confusion matrix resulting from the classification of unseen time-deconvolved data. Almost all (γ,n) and (γ,p) are correctly classified. Two percent of $^{12}\text{C}(\gamma,\alpha)$ reactions are mis-classified as $^{16}\text{O}(\gamma,\alpha)$.

pre-classified datasets is required. A promising approach could be to train the network on Monte-Carlo simulated data that accurately capture the physics of the events, as well as background noise, and detector effects. The generation of sufficiently realistic training data remains an area of active pursuit.

5. – Summary

The Convolutional Neural Network, ResNeXt50, was used to classify nuclear reactions measured in the gas-filled Warsaw active-target time-projection chamber. This detector was used to study photo-dissociation reactions of ^{16}O and ^{12}C using intense γ -beams at the HI γ S facility at Duke University. Reactions of interest were distinguished from other background channels using ResNeXt50. For this preliminary study, a small sample of two hundred hand-classified events were used to train the model, which was then used to classify one hundred unseen events. This method was effective at identifying background events with 99% accuracy and could differentiate two key reaction channels with 96% accuracy. For this sample of events, time-deconvolved events were classified most accurately.

* * *

The authors would like to thank the staff of HI γ S and for providing high-quality beams during the experiments. This work is supported by: the UK Science and Technology Facilities Council (STFC) (grant No. ST/V001086/1); the National Science Centre, Poland (contract No. UMO-2019/33/B/ST2/02176); the University of Warsaw, Poland (Interdisciplinary Centre for Mathematical and Computational Modelling – computational allocation No. G89-1286 and Excellence Initiative Research University – IDUB program); the US Department of Energy, Office of Science, Office of Nuclear Physics (grant No. DE-FG02-94ER40870, DE-FG02-97ER41033, and DE-FG02-97ER41041); and by the Romanian Ministry of Research, Innovation and Digitalization (contract No. PN 23.21.01.06).

REFERENCES

- [1] ĆWIOK M. *et al.*, *EPJ Web of Conferences*, **279** (2023) 04002.
- [2] SMITH R. *et al.*, *Nat. Commun.*, **12** (2021) 5920.
- [3] GAI M. *et al.*, *JINST*, **5** (2010) P12004.
- [4] ZIMMERMAN W. R. *et al.*, *Phys. Rev. Lett.*, **110** (2013) 152502.
- [5] SMITH R. *et al.*, *Phys. Rev. C*, **101** (2020) 021302.
- [6] BIJKER R. and IACHELLO F., *Ann. Phys. (N.Y.)*, **298** (2002) 334.
- [7] ĆWIOK, M. *et al.*, *EPJ Web of Conferences*, **290** (2023) 01004.
- [8] WU H. *et al.*, *Nucl. Instrum. Methods Phys. Res. A*, **1055** (2023) 168528.
- [9] CENNINI P. *et al.*, *Nucl. Instrum. Methods Phys. Res. A*, **356** (1995) 507.
- [10] KIM, C. *et al.*, *Nucl. Instrum. Methods Phys. Res. A*, **1048** (2023) 168025.
- [11] QIAN Z. *et al.*, *Nucl. Instrum. Methods Phys. Res. A*, **1010** (2021) 165527.
- [12] KUCHERA M. P. *et al.*, *Nucl. Instrum. Methods Phys. Res. A*, **940** (2019) 156.
- [13] MAYER J. *et al.*, *Nucl. Instrum. Methods Phys. Res. A*, **1013** (2021) 165666.
- [14] WELLER H. R. *et al.*, *Prog. Part. Nucl. Phys.*, **62** (2009) 257.
- [15] SAULI F., *Nucl. Instrum. Methods Phys. Res. A*, **386** (1997) 531.
- [16] POLLACCO E. C. *et al.*, *Nucl. Instrum. Methods Phys. Res. A*, **887** (2018) 81.
- [17] XIE S. *et al.*, *Aggregated residual transformations for deep neural networks*, in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (IEEE)* 2017.
- [18] ERDOGAN A., *ResNeXt: A New Paradigm in Image Processing*, <https://medium.com/@atakanerdogan305/resnext-a-new-paradigm-in-image-processing-ee40425aea1f> (2017), accessed: 11 March 2024.
- [19] GIOVINAZZO J. *et al.*, *Nucl. Instrum. Methods Phys. Res. A*, **840** (2016) 15.