
Sparse Methods for Vector Embeddings of TPC Data

Tyler Wheeler¹ Michelle P. Kuchera² Raghuram Ramanujan²
Ryan Krupp¹ Chris Wrede¹ Saiprasad Ravishankar¹
Connor L. Cross² Hoi Yan Ian Heung² Andrew J. Jones² Benjamin Votaw²
¹ Michigan State University ² Davidson College
wheele56@msu.edu {mikuchera, ramanujan}@davidson.edu
{krupprya, wrede, ravisha3}@msu.edu
{cocross, iaheung, anjones1, bevotaw}@davidson.edu

Abstract

Time Projection Chambers (TPCs) are versatile detectors that reconstruct charged-particle tracks in an ionizing medium, enabling sensitive measurements across a wide range of nuclear physics experiments. We explore sparse convolutional networks for representation learning on TPC data, finding that a sparse ResNet architecture, even with randomly set weights, provides useful structured vector embeddings of events. Pre-training this architecture on a simple physics-motivated binary classification task further improves the embedding quality. Using data from the Gaseous Detector with GERmanium Tagging (GADGET) II TPC, a detector optimized for measuring low-energy β -delayed particle decays, we represent raw pad-level signals as sparse tensors, train Minkowski Engine ResNet models, and probe the resulting event-level embeddings which reveal rich event structure. As a cross-detector test, we embed data from the Active-Target TPC (AT-TPC)—a detector designed for nuclear reaction studies in inverse kinematics—using the same encoder. We find that even an untrained sparse ResNet model provides useful embeddings of AT-TPC data, and we observe improvements when the model is trained on GADGET data. Together, these results highlight the potential of sparse convolutional techniques as a general tool for representation learning in diverse TPC experiments.

1 Introduction

Time Projection Chambers (TPCs) are widely used in nuclear physics experiments for their ability to reconstruct three-dimensional charged-particle tracks. By recording ionization signals in a detection medium (typically a gas or liquid), TPCs enable sensitive studies of rare decays and reaction dynamics across a broad range of energies and beam conditions [Hilke, 2010]. However, these benefits come with substantial computational challenges: TPCs produce massive, high-dimensional datasets that are inherently sparse and vary significantly across detector geometries and experimental goals.

Traditional approaches to analyzing TPC data often rely on detector-specific pipelines, limiting the development of generalizable tools that can be applied across experiments. Deep learning approaches have been successful for TPC data, with deep convolutional networks being a common choice for TPC “images” [Kuchera et al., 2019, Solli et al., 2021, Abratenko et al., 2021, Carloni et al., 2022, Wheeler et al., 2025, Dey et al., 2025], but have thus far required independent training for each task of interest. They also suffer from computational inefficiency when applied to sparse high-dimensional inputs.

In this work, we explore the use of sparse convolutional neural networks, implemented via the Minkowski Engine framework [Choy et al., 2019], to process raw TPC data efficiently and extract

meaningful event representations. We are particularly interested in whether such models can transfer across different TPC systems—rather than purely optimizing classifier performance—thereby motivating the eventual development of general-purpose TPC foundation models. We focus primarily on the Gaseous Detector with GERmanium Tagging (GADGET II) TPC [Mahajan et al., 2024], detailing how raw pad-level signals are represented as sparse tensors, trained with ResNet models, and analyzed in a latent space. To demonstrate the richness of the embedding space, we also embed data from the $^{16}\text{O} + \alpha$ experiment that used the Active-Target TPC (AT-TPC) [Ayyad et al., 2020]. The two detectors differ markedly in geometry and scientific goals—GADGET II is compact and optimized for low-energy β -delayed charged particle decays, whereas the AT-TPC provides wide solid-angle coverage for inverse-kinematics reactions and is surrounded by a solenoid magnet that provides up to 2 T of magnetic field—allowing us to assess how well sparse tensor methods carry across TPC designs. Example data from both detectors are shown in Figure 1. We would like to draw the reader’s attention to the visual contrast between these event topologies: tracks from GADGET II and the AT-TPC look like data from entirely unrelated domains. It is thus remarkable that transfer learning between such distinct detectors is effective. A concise summary of our findings is provided in the Results and Conclusion sections.

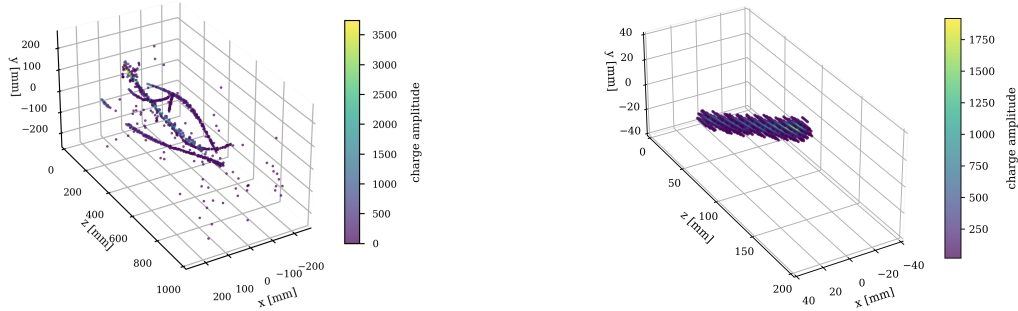


Figure 1: Left: an example of a 4-track AT-TPC event. Right: an example of a 1600 keV proton event from the GADGET II TPC.

2 Methods

2.1 Network Architecture

We employ sparse convolutional neural networks implemented with the Minkowski Engine to process four-dimensional point cloud data from TPCs, where each point consists of spatial coordinates and an associated charge (x, y, z, q) . Convolutions operate only in the three spatial dimensions, while q is treated as a feature channel. This preserves pad-level sparsity and avoids the cost of dense voxelization.

The backbone is a shallow residual network (ResNet14) [He et al., 2015] adapted for sparse inputs. Its main components are summarized in Table 1. The design includes an initial stem, four residual stages, and a final pre-pooling block followed by global max pooling and a fully connected head. This configuration uses ten convolutional layers and one linear layer. We found that deeper variants (e.g., ResNet50 with bottleneck blocks) increased training time without clear performance gains.

A key advantage of this architecture is its ability to handle variable-length inputs natively. Unlike many implementations (e.g., typical PointNet [Charles et al., 2017] training pipelines) that require padding or truncation, sparse ResNets flexibly accommodate events with widely varying numbers of active channels. This is particularly important for TPC data, where track lengths and hit multiplicities can vary substantially both within and across detectors.

Table 1: Summary of the sparse ResNet14 architecture. Convolutions are sparse $3\times 3\times 3$ operations unless otherwise noted.

Stage	Description	Key Operations
Stem	Input processing	Sparse conv (stride 2), BN, ReLU, max pool (2^3)
Residual Stage 1–4	Feature extraction	$2\times$ BasicBlock per stage (ReLU); stride 2 at start
Downsampling	Dimension alignment	1^3 sparse conv when required
Pre-pooling block	Feature refinement	Dropout (0.8), sparse conv (stride 3), BN, GELU
Global pooling	Feature aggregation	Global max pool
Head	Output	Fully connected linear layer

Table 2: Key training and optimization settings.

Item	Setting
Loss	Cross-entropy
Optimizer	Adam (LR 5×10^{-4} , weight decay 10^{-4})
Scheduler	Cosine annealing ($T_{\max} = 13$)
Batching	Variable-length events; batch size 64
Regularization	Gradient clipping (max-norm = 1.0), dropout $p = 0.8$
Epochs	15 (best model by validation loss)
Class imbalance	Weighted random sampler
Metrics	Accuracy, precision, recall, macro-F1

2.2 Training Setup

The sparse ResNet14 model is trained with four input channels (x, y, z, q) to perform binary proton–alpha classification on GADGET II data; we refer to this pretrained model as $\text{ResNet}_{\text{train}}$. Labels are derived by gating distinct bands in the range–energy plane (Figure 2). Inputs are raw pad-level hits quantized with a voxel size of 0.05. Batching uses Minkowski Engine’s `batched_coordinates` without padding. The best checkpoint is selected using the lowest validation loss, and training converges within ~ 15 epochs, achieving ≈ 0.995 accuracy and ≈ 0.99 F1. Key hyperparameters and optimization settings are summarized in Table 2.

2.3 Latent Embedding and Analysis

We extract latent vectors from the penultimate layer of our ResNet14 model trained solely on the binary proton–alpha classification task ($\text{ResNet}_{\text{train}}$). Additionally, we also investigate the embeddings produced by a randomly initialized ResNet14 that undergoes no further training ($\text{ResNet}_{\text{rand}}$). Prior research with dense CNNs has uncovered that such randomly weighted networks produce surprisingly effective embeddings [Ulyanov et al., 2018, Saxe et al., 2011, Rosenfeld and Tsotsos, 2018] and we use these as a strong baseline for comparison.

To investigate the information encoded in the latent space of our models, we train a linear probe [Belinkov, 2022]—specifically, a linear support vector machine—on various classification tasks that are distinct from the training task. Deliberately choosing a low capacity model as the probe allows us to test whether the embeddings make useful information about events accessible in a straightforward, readily interpretable form.

To visualize the structure of the embedding space, we apply principal component analysis (PCA) [Jolliffe, 2002] to reduce the high-dimensional latent vectors to their first two principal components. This linear projection captures the directions of maximum variance, providing a view of the embeddings that is legible to humans and allowing us to directly see whether architectural choices and/or pretraining induce separation among event types.

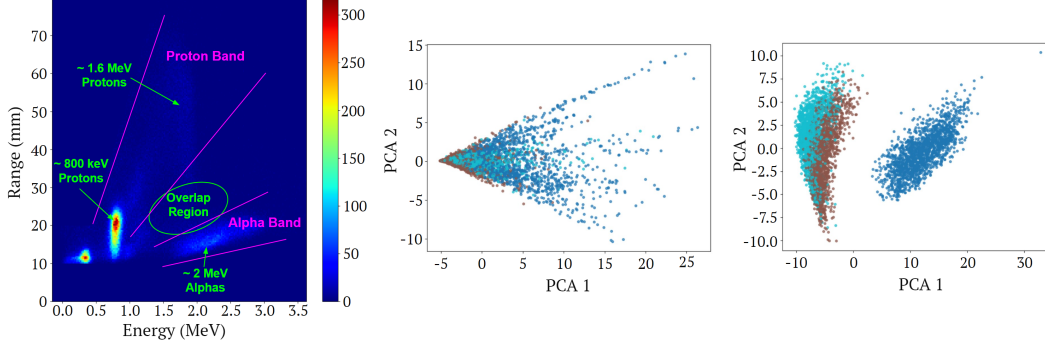


Figure 2: Left: Range vs energy plot for FRIB experiment 21072 (1 hr run). Distinct proton and alpha bands are visible; gating cuts on these bands provided training labels for models, while events in the overlap region were excluded. Middle: PCA of embeddings from ResNet_{rand} . Right: PCA of embeddings from ResNet_{train} . In PCA plots: 800 keV p , 1600 keV p , and 2 MeV α ; brown, light blue, and dark blue, respectively.

3 Results

3.1 Linear Probing

To probe the information embedded in the latent space of our models, we use a linear classifier to distinguish between categories that were not used in training. We probe embedded representations of data both from GADGET II and from the AT-TPC using both ResNet_{train} and ResNet_{rand} .

We train a linear support vector machine on classification tasks that are distinct from the training task for ResNet_{train} . For GADGET II data, the linear probe achieved high performance on a three-class separation task—distinguishing between 800 keV protons, 1600 keV protons, and 2 MeV alphas—that is more challenging than the binary classification task used for pretraining. Accuracy and F1 both reached 0.97 when using the ResNet_{train} embeddings, compared to 0.85 for ResNet_{rand} (see Table 3). Both numbers are significantly higher than the corresponding figures for the naïve model that always predicts the mode: accuracy of 0.33 and an F1 score of 0.17. This indicates that the embeddings capture physically meaningful substructure beyond the binary training labels, enabling linear separability of multiple decay channels. Moreover, the higher performance of the ResNet_{train} embeddings suggests that there is value in pretraining on in-domain data, even if it’s on a different task.

For AT-TPC data, embeddings from ResNet_{rand} and ResNet_{train} were probed using a track counting task ($\{0, 1, 2\}$ vs. $\{3\}$ vs. $\{4, 5\}$ tracks). The results (see Table 3) show that latent features learned on one detector retain their utility when transferred to a different system with vastly different geometry and physics goals.

It is important to note that our objective in these experiments is not to realize the highest-performing classifier. With further optimization and engineering effort, higher classification accuracy could undoubtedly be achieved. Rather, our intent is to demonstrate that this form of transfer learning is viable in practice and that it provides a basis for developing broadly useful TPC foundation models.

3.2 PCA Analysis

PCA projections of the GADGET II embeddings are shown in Figure 2. Even without training (ResNet_{rand}), the architecture produces embeddings with some degree of ordering: classes are partially separate in the 2D projection, reflecting the architectural bias of sparse convolutions toward learning from local spatial continuity.

Embeddings from ResNet_{train} , which was optimized only for binary proton–alpha discrimination, exhibit an even cleaner separation of all three particle classes (800 keV p , 1600 keV p , 2 MeV α). This shows that pretraining on a simple physics-driven task yields representations with richer event structure than the labels alone, producing a latent space where multi-class tasks can become linearly separable.

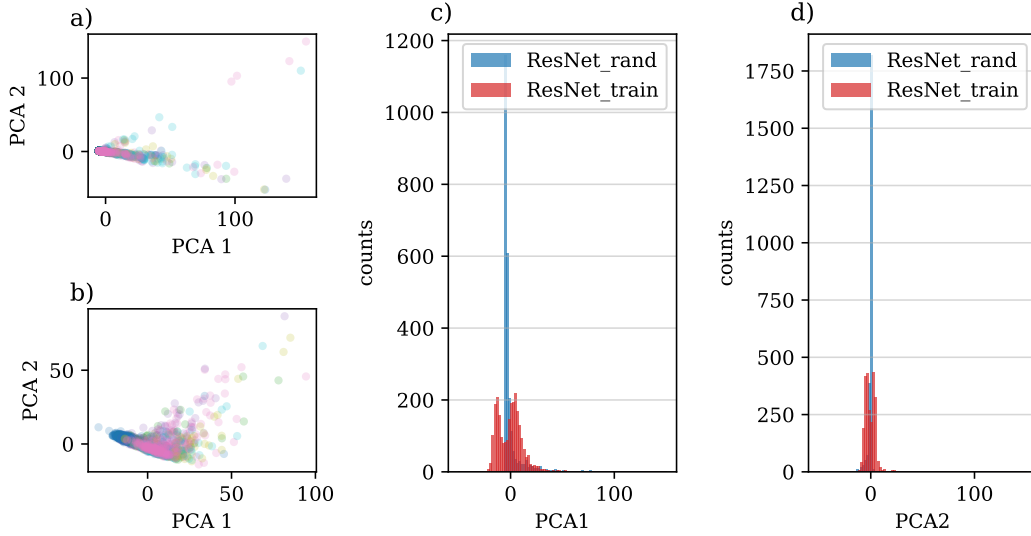


Figure 3: Visualization of AT-TPC ResNet embeddings. Plot a) shows the first two principal components (PCA1 and PCA2 respectively) of embeddings from ResNet_{rand}. Plot b) depicts PCA2 vs. PCA1 from ResNet_{train}. Plots c) and d) present the distribution of latent embeddings for PCA1 and PCA2, respectively.

We further applied PCA to embeddings of AT-TPC data (Figure 3). For ResNet_{rand}, projections again show weak but nontrivial organization, with narrow distributions and limited separation. In contrast, embeddings obtained by running AT-TPC events through the GADGET-trained encoder (ResNet_{train}) show more dispersed embeddings in PCA space, allowing for more successful separation. This demonstrates that features learned from GADGET II not only enhance latent separation in-domain but also transfer effectively across detectors.

Domain	Accuracy			F1 Score		
	ResNet _{train}	ResNet _{rand}	Naïve	ResNet _{train}	ResNet _{rand}	Naïve
GADGET II	0.97	0.85	0.33	0.97	0.85	0.17
AT-TPC	0.74	0.58	0.48	0.70	0.53	0.31

Table 3: Linear probe results on GADGET II three-class particle classification and AT-TPC track counting tasks. The untrained architecture (ResNet_{rand}) provides non-trivial structure; pretraining on GADGET II data (ResNet_{train}) improves both in-domain and out-of-domain transfer.

4 Conclusion

We showed that embeddings learned by sparse convolutional networks trained on GADGET II data not only enable accurate classification of decay events within that detector, but also transfer effectively to new tasks, and even to different TPCs with distinct geometries and physics goals (AT-TPC). While even an untrained architecture yields embeddings with useful structure, pretraining on a straightforward GADGET II proton-alpha classification task substantially improves the utility of the embeddings. These results highlight the value of sparse tensor methods for analyzing and processing TPC data and their promise for building cross-domain foundation models for TPCs.

5 Acknowledgments

This work was supported by the U.S. Department of Energy, Office of Science, under award No. DE-SC0024587 and the National Science Foundation Office of Advanced Cyberinfrastructure award No. OAC-2311263.

References

- P. Abratenko, M. Alrashed, R. An, J. Anthony, J. Asaadi, A. Ashkenazi, S. Balasubramanian, B. Baller, C. Barnes, G. Barr, V. Basque, L. Bathe-Peters, O. Benevides Rodrigues, S. Berkman, A. Bhandari, A. Bhat, M. Bishai, A. Blake, T. Bolton, L. Camilleri, D. Caratelli, I. Caro Terrazas, R. Castillo Fernandez, F. Cavanna, G. Cerati, Y. Chen, E. Church, D. Cianci, J. M. Conrad, M. Convery, L. Cooper-Troendle, J. I. Crespo-Anadón, M. Del Tutto, A. Devitt, R. Diurba, L. Domine, R. Dorrill, K. Duffy, S. Dytman, B. Eberly, A. Ereditato, L. Escudero Sanchez, J. J. Evans, G. A. Fiorentini Aguirre, R. S. Fitzpatrick, B. T. Fleming, N. Foppiani, D. Franco, A. P. Furmanski, D. Garcia-Gamez, S. Gardiner, G. Ge, S. Gollapinni, O. Goodwin, E. Gramellini, P. Green, H. Greenlee, W. Gu, R. Guenette, P. Guzowski, E. Hall, P. Hamilton, O. Hen, G. A. Horton-Smith, A. Hourlier, E.-C. Huang, R. Itay, C. James, J. Jan de Vries, X. Ji, L. Jiang, J. H. Jo, R. A. Johnson, Y.-J. Jwa, N. Kamp, G. Karagiorgi, W. Ketchum, B. Kirby, M. Kirby, T. Kobilarcik, I. Kreslo, R. LaZur, I. Lepetic, K. Li, Y. Li, B. R. Littlejohn, D. Lorca, W. C. Louis, X. Luo, A. Marchionni, S. Marcocci, C. Mariani, D. Marsden, J. Marshall, J. Martin-Albo, D. A. Martinez Caicedo, K. Mason, A. Mastbaum, N. McConkey, V. Meddage, T. Mettler, K. Miller, J. Mills, K. Mistry, A. Mogan, T. Mohayai, J. Moon, M. Mooney, A. F. Moor, C. D. Moore, J. Mousseau, M. Murphy, D. Naples, A. Navrer-Agasson, R. K. Neely, P. Nienaber, J. Nowak, O. Palamara, V. Paolone, A. Papadopoulou, V. Papavassiliou, S. F. Pate, A. Paudel, Z. Pavlovic, E. Piasetzky, I. D. Ponce-Pinto, D. Porzio, S. Prince, X. Qian, J. L. Raaf, V. Radeka, A. Rafique, M. Reggiani-Guzzo, L. Ren, L. Rochester, J. Rodriguez Rondon, H. E. Rogers, M. Rosenberg, M. Ross-Lonergan, B. Russell, G. Scanavini, D. W. Schmitz, A. Schukraft, M. H. Shaevitz, R. Sharankova, J. Sinclair, A. Smith, E. L. Snider, M. Soderberg, S. Söldner-Rembold, S. R. Soleti, P. Spentzouris, J. Spitz, M. Stancari, J. St. John, T. Strauss, K. Sutton, S. Sword-Fehlberg, A. M. Szelc, N. Tagg, W. Tang, K. Terao, C. Thorpe, M. Touns, Y.-T. Tsai, S. Tufanli, M. A. Uchida, T. Usher, W. Van De Pontseele, B. Viren, M. Weber, H. Wei, Z. Williams, S. Wolbers, T. Wongjirad, M. Wospakrik, W. Wu, T. Yang, G. Yarbrough, L. E. Yates, G. P. Zeller, J. Zennamo, and C. Zhang. Convolutional neural network for multiple particle identification in the microboone liquid argon time projection chamber. *Phys. Rev. D*, 103:092003, May 2021. doi: 10.1103/PhysRevD.103.092003. URL <https://link.aps.org/doi/10.1103/PhysRevD.103.092003>.
- Y. Ayyad, N. Abgrall, T. Ahn, H. Álvarez Pol, D. Bazin, S. Beceiro-Novo, L. Carpenter, R.J. Cooper, M. Cortesi, A.O. Macchiavelli, W. Mittag, B. Olaizola, J.S. Randhawa, C. Santamaria, N. Watwood, J.C. Zamora, and R.G.T. Zegers. Next-generation experiments with the active target time projection chamber (at-tpc). *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 954:161341, 2020. ISSN 0168-9002. doi: <https://doi.org/10.1016/j.nima.2018.10.019>. URL <https://www.sciencedirect.com/science/article/pii/S0168900218313342>. Symposium on Radiation Measurements and Applications XVII.
- Y. Belinkov. Probing classifiers: Promises, shortcomings, and advances. *Computational Linguistics*, 48(1):207–219, March 2022. doi: 10.1162/coli_a_00422. URL <https://aclanthology.org/2022.cl-1.7/>.
- K. Carloni, N.W. Kamp, A. Schneider, and J.M. Conrad. Convolutional neural networks for shower energy prediction in liquid argon time projection chambers. *Journal of Instrumentation*, 17(02):P02022, feb 2022. doi: 10.1088/1748-0221/17/02/P02022. URL <https://dx.doi.org/10.1088/1748-0221/17/02/P02022>.
- R. Qi Charles, Hao Su, Mo Kaichun, and Leonidas J. Guibas. Pointnet: Deep learning on point sets for 3d classification and segmentation. In *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pages 77–85, 2017. doi: 10.1109/CVPR.2017.16.
- C. Choy, J. Gwak, and S. Savarese. 4d spatio-temporal convnets: Minkowski convolutional neural networks. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, June 2019.
- P. Dey, Adam K. Anthony, Curtis Hunt, Michelle P. Kuchera, Raghuram Ramanujan, William G. Lynch, ManYee Betty Tsang, Joseph M. Wieske, Jessica W. Ajongbah, Saul Beceiro-Novo, Kyle W. Brown, Zbigniew Chajecski, Kaitlin J. Cook, Skyler Gangestad, Tom Ginter, Bergen Kendzioriski, Fanurs Chi Eh Teh, and HoTing Wong. Point-cloud based machine learning for classifying rare events in the active-target time projection chamber. *Nuclear Instruments and Methods in Physics*

Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment, 1072:170002, 2025. ISSN 0168-9002. doi: <https://doi.org/10.1016/j.nima.2024.170002>. URL <https://www.sciencedirect.com/science/article/pii/S0168900224009288>.

K. He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. Deep residual learning for image recognition, 2015. URL <https://arxiv.org/abs/1512.03385>.

H. J. Hilke. Time projection chambers. *Reports on Progress in Physics*, 73(11):116201, oct 2010. doi: 10.1088/0034-4885/73/11/116201. URL <https://dx.doi.org/10.1088/0034-4885/73/11/116201>.

I. T. Jolliffe. *Principal Component Analysis*. Springer, 2nd edition, 2002.

M. P. Kuchera, R. Ramanujan, J.Z. Taylor, R.R. Strauss, D. Bazin, J. Bradt, and Ruiming Chen. Machine learning methods for track classification in the at-tpc. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 940:156–167, 2019. ISSN 0168-9002. doi: <https://doi.org/10.1016/j.nima.2019.05.097>. URL <https://www.sciencedirect.com/science/article/pii/S0168900219308046>.

R. Mahajan, T. Wheeler, E. Pollacco, C. Wrede, A. Adams, H. Alvarez-Pol, A. Andalib, A. Anthony, Y. Ayyad, D. Bazin, T. Budner, M. Cortesi, J. Dopfer, M. Friedman, B. Jain, A. Jaros, D. Pérez-Loureiro, B. Mehl, R. De Oliveira, S. Ravishankar, L. J. Sun, and J. Surbrook. Time projection chamber for gadget ii. *Phys. Rev. C*, 110:035807, Sep 2024. doi: 10.1103/PhysRevC.110.035807. URL <https://link.aps.org/doi/10.1103/PhysRevC.110.035807>.

A. Rosenfeld and John K. Tsotsos. Intriguing properties of randomly weighted networks: Generalizing while learning next to nothing. *CoRR*, abs/1802.00844, 2018. URL <http://arxiv.org/abs/1802.00844>.

A. M. Saxe, Pang Wei Koh, Zhenghao Chen, Maneesh Bhand, Bipin Suresh, and Andrew Y. Ng. On random weights and unsupervised feature learning. In *Proceedings of the 28th International Conference on International Conference on Machine Learning*, ICML’11, page 1089–1096, Madison, WI, USA, 2011. Omnipress. ISBN 9781450306195.

R. Solli, D. Bazin, M. Hjorth-Jensen, M.P. Kuchera, and R.R. Strauss. Unsupervised learning for identifying events in active target experiments. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 1010:165461, 2021. ISSN 0168-9002. doi: <https://doi.org/10.1016/j.nima.2021.165461>. URL <https://www.sciencedirect.com/science/article/pii/S0168900221004460>.

D. Ulyanov, Andrea Vedaldi, and Victor S. Lempitsky. Deep image prior. In *2018 IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2018, Salt Lake City, UT, USA, June 18-22, 2018*, pages 9446–9454. Computer Vision Foundation / IEEE Computer Society, 2018. doi: 10.1109/CVPR.2018.00984. URL http://openaccess.thecvf.com/content_cvpr_2018/html/Ulyanov_Deep_Image_Prior_CVPR_2018_paper.html.

T. Wheeler, S. Ravishankar, C. Wrede, A. Andalib, A. Anthony, Y. Ayyad, B. Jain, A. Jaros, R. Mahajan, L. Schaedig, A. Adams, S. Ahn, J.M. Allmond, D. Bardayan, D. Bazin, K. Bosmpotinis, T. Budner, S.R. Carmichael, S.M. Cha, A. Chen, K.A. Chipps, J.M. Christie, I. Cox, J. Dopfer, M. Friedman, J. Garcia-Duarte, E. Good, T.J. Gray, A. Green, R. Grzywacz, K. Hahn, R. Jain, E.A.M. Jensen, T. King, S.N. Liddick, B. Longfellow, R. Lubna, C. Marshall, Y. Mishnayot, A.J. Mitchell, F. Montes, T.H. Ogunbeku, J. Owens-Fryar, S.D. Pain, J. Pereira, E. Pollacco, A.M. Rogers, M.Z. Serikow, K. Setoodehnia, L.J. Sun, J. Surbrook, A. Tsantiri, and L.E. Weghorn. Object detection with deep learning for rare event search in the gadget ii tpc. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 1080:170659, 2025. ISSN 0168-9002. doi: <https://doi.org/10.1016/j.nima.2025.170659>. URL <https://www.sciencedirect.com/science/article/pii/S0168900225004607>.