

2020 年臺灣國際科學展覽會 優勝作品專輯

作品編號 160044

參展科別 物理與天文學

作品名稱 **Improving Particle Classification In Wimp
Dark Matter Detection Using Neural
Networks**

得獎獎項 大會獎：二等獎

國 家 Canada

就讀學校 Lo-Ellen Park Secondary School

作者姓名 Brendon Matusch

關鍵詞 Dark Matter, Neural Networks,
Deep Learning

作者照片



1 Introduction

In all experiments for detection of WIMP dark matter, it is essential to develop a classifier that can distinguish potential WIMP events from background radiation. Most often, classifiers are developed manually, via physical modeling and empirical optimization. This is problematic for two reasons: it takes a great deal of time and effort away from developing the experiment, and the resulting classifiers often perform suboptimally (which means that a greater amount of expensive run time is required to obtain a confident experimental result).

Machine learning has the potential to automate this and accelerate experimentation, and also to detect patterns that humans cannot. However, two major challenges, which are shared among several dark matter experiments, stand in the way: impure calibration data, which hinders training of models, and unpredictable physical dynamics within the detector itself.

My objective was to develop a set of machine learning techniques that address these two problems, and thus more efficiently generate highly accurate classifiers.

I was able to obtain raw data for two dark matter experiments which exhibit these challenges: the PICO-60 bubble chamber [2], and the DEAP-3600 liquid argon scintillator [1]. For each experiment, I developed and compared three general-purpose algorithms intended to resolve its inherent challenge (impurity and unpredictable dynamics, respectively).

In PICO-60, background alpha and WIMP-like neutron calibration datasets are used for training; however, there is an impurity of 10% alphas in the neutron set. While a conventional classifier was developed (and is believed to be 100% accurate), machine learning in the form of a supervised neural network (NN) has also been previously explored, because of the benefits of automation. Unfortunately, it achieved a mean accuracy of only 80.2% – not usable as a practical replacement for conventional methods in future iterations of the experiment.

In DEAP-3600, photons are absorbed by a wavelength shifting medium and re-emitted in an unpredictable direction, before being detected by one of 255 photomultiplier tubes (PMTs) around the spherical detector. The randomness severely limits the accuracy of conventional classifiers; in a simulation, the best so far removes 99.6% of alpha background, while also (undesirably) removing 91.0% of WIMP events. Because of physical limitations, simulated data is used for calibration, with 30 real-world experimental events available for testing.

I have written a research paper [11] about my work on PICO-60, which has been approved by the PICO collaboration and pre-published at <https://arxiv.org/abs/1811.11308>. It is currently undergoing peer review for publication in *Computer Physics Communications*.

All PICO researchers are listed on my paper for their work on the original PICO-60 experiment. They did not contribute to this study; I completed and documented it independently.

2 Procedure

2.1 Resolving Impurities (PICO-60)

In the context of PICO-60, my intent was to develop a highly accurate classification algorithm

that is tolerant to substantial impurities in calibration data. I developed and compared two sets of classification algorithms: supervised learning, and semi-supervised learning.

2.1.1 Supervised Learning

I experimented with supervised learning, exploring the following data formats to determine which facilitates the best-performing solution:

1. An 8-band Fourier transform of audio recorded in the detector. I applied a dense NN with dropout [6] and L2 regularization (which reduce overfitting to impure data).
2. A full-resolution Fourier transform of the audio (all 50,001 data points). I once again applied a dense neural network with the Adam [9] optimizer.
3. A raw audio waveform, with a very deep 1D convolutional neural network (CNN) inspired by Dai et al. [5].
4. Event images, with a 2D CNN (to determine whether they contain useful information).

2.1.2 Semi-Supervised Learning

To better handle impurities, I developed two novel semi-supervised learning techniques. The concept of both algorithms is to iteratively generate more accurate labels for a set of inaccurately labeled training data, permitting less overfitting and greater validation accuracy than simple regularization would allow.

1. Iterative Clustering is inspired by unsupervised clustering algorithms, and takes advantage of the “confidence” of the neural network’s predictions (that is, how close the predictions are to 0 or 1).

First, a neural network is trained on a labeled set. After 30 epochs, it runs predictions on the unlabeled set. As the network trains, the confidence of these predictions increases. When $p < t$ or $p > 1 - t$ for a given prediction p and a (small) confidence threshold t , the example is labeled correspondingly and added to the labeled set.

It is hypothesized that accuracy will gradually *improve* as the neural network’s most confident predictions are fed back into the network as newly labeled training data.

2. Gravitational Differentiation is an analog alternative to Iterative Clustering. Using a novel piecewise exponential function for calculating final-layer derivatives, called $\text{GravDiff}(p, \psi, g)$, unlabeled examples are caused to “gravitate” toward more accurate predictions based on the confidence of the neural network’s predictions.

The function is a transformation of the hyperbolic tangent that flattens the central range and comparatively exaggerates the asymptote on either side, defined as follows:

$$\text{GravDiff}(p, \psi, g) = g \cdot \text{sgn}(p) \cdot |\tanh(2p - 1)|^\psi$$

where p is the network’s prediction, ψ determines the rate at which the output approaches 0 as the confidence of the prediction decreases, and g is the learning rate multiplier. The output is the final-layer gradient to be used in backpropagation.

2.2 Resolving Unpredictable Photon Dynamics (DEAP-3600)

In DEAP-3600, the key challenge is the aforementioned unpredictable photon dynamics. The major source of background radiation is “neck alphas”, which strike atoms in the gas filling the neck above the spherical detector body; because of the unpredictable direction of photon emission, it is non-trivial to identify this class of events.

Because of the inherently spatial nature of this problem, I hypothesized it would be important to most effectively process the arrangement of PMTs in the DEAP-3600 detector: a hexagonal lattice on the surface of a sphere. I evaluated three approaches to this problem (two of which are original), listed below. Techniques 2 and 3 (Figure 1) take advantage of geometric and spatial patterns in PMT activations associated with neck alpha events, by extending the CNN concept for use in spherical and arbitrary topologies.

1. I inputted photon counts from each PMT into a multi-layer perceptron.
2. I developed and applied a Mercator-like cylindrical projection, which maps a sphere onto a rectangle. Cubic spline interpolation was used to map latitude lines on the hexagonal lattice (which vary in length) to equal-length rows on the rectangular image. Once again, photon counts were used as input data.
3. I developed a new type of CNN, called a topological CNN. Rather than a 2D image of square pixels, kernels convolve over an arbitrary topology.

While a conventional convolutional layer downscales a rectangular image to a smaller rectangular image by multiplying square regions by a set of corresponding weights, a topological CNN (as applied to this problem) downscales the approximately spherical hexagonal mesh to a smaller but topologically equivalent hexagonal mesh.

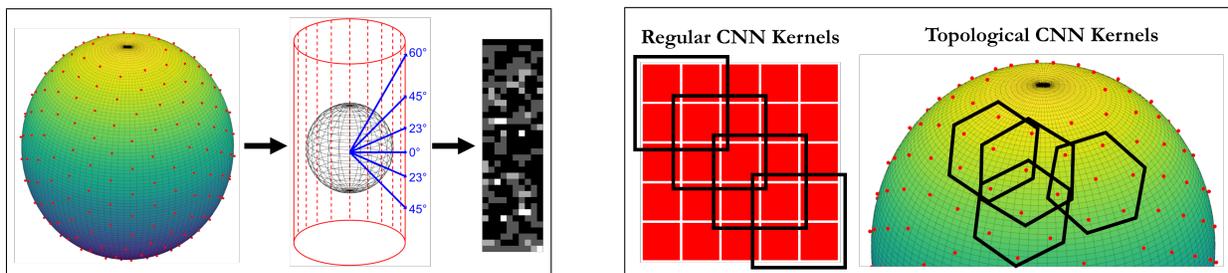


Figure 1: Diagrams of the cylindrical projection (left) and topological CNN (right) systems.

2.3 Statistical Practices

Models were optimized using grid searches, in which every configuration within a hyperparameter space is trained and evaluated. The best-performing technique and hyperparameters were selected based on repeated random sub-sampling of validation sets.

Finally, the best-performing model was retrained on a newly selected training set, and a test set was used to verify its performance. In DEAP-3600, two test sets were used: one consisting of simulated data, and another consisting of 30 examples collected from the real-world detector (neutron calibration data, and background radiation).

3 Results

3.1 Resolving Impurities (PICO-60)

As seen in Figure 2, the banded Fourier transform format was the most effective of the supervised learning models (including the previous study), achieving 97.3% validation accuracy. The image data was shown to contain no useful information, reaching only 63.0% accuracy.

Semi-supervised learning produced better accuracy than supervised learning. Gravitational Differentiation was the best overall technique, achieving 99.7% accuracy on validation data and 98.3% on test data. This confirms my hypothesis; it is indeed beneficial to incorporate the neural network’s most confident predictions into its training data.

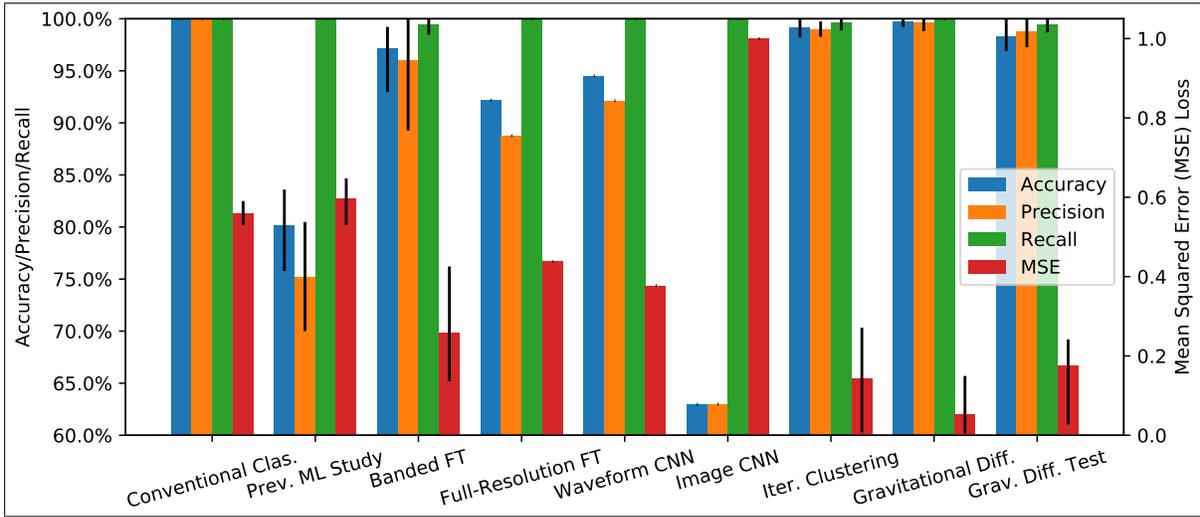


Figure 2: Validation accuracy, precision, recall, and loss for each technique in PICO-60. Error bars represent the 8th and 92nd percentiles.

3.2 Resolving Unpredictable Photon Dynamics (DEAP-3600)

The three I developed to resolve the problem of unpredictable dynamics in DEAP-3600 were evaluated based on:

- Identification of neck alphas with a minimum of 99.6% precision (the same as the previously developed conventional classifier).
- Reduction in the rate of false positives (simulated WIMP events misidentified as neck alphas) compared to the 91.0% result achieved with a conventional classifier.

All techniques were validated on simulated data. Both the dense neural network and topological CNN techniques were unsuccessful; they produced higher false positive rates than the conventional classifier, meaning that they misclassified more WIMPs as alpha particles.

However, as seen in Figure 3, the cylindrical projection method retained a 99.6% neck alpha identification rate, and also successfully *reduced* the false positive rate from 91.0% to 75.7%.

This successful network model was further tested on the limited amount of available real-world experimental data. It removed 73.9% of WIMP-like (neutron calibration) events,

which is in line with the 75.7% obtained on simulated data.

However, only 91.2% of background events (expected to be neck alphas) were removed, below the 99.6% reached on simulated data. This indicates one of three possibilities:

1. The simulation is not an accurate representation of the real-world detector (some physical dynamic has not been simulated), and thus the network did not generalize.
2. There is WIMP dark matter present in the set of background events.
3. There is an unanticipated form of background radiation present.

Technique	Data Source	Background Removal Rate	False Positive Rate (WIMP Misclassifications)
Conventional Classifier	Simulation (Val.)	99.6%	91.0%
Dense Neural Network	Simulation (Val.)	100%	100%
Cylindrical Projection	Simulation (Val.)	99.9%	76.7%
Topological CNN	Simulation (Val.)	99.9%	93.0%
Cylindrical Projection	Simulation (Test)	99.6%	75.7%
Cylindrical Projection	Real-World	91.2%	73.9%

Figure 3: Average background and WIMP/neutron removal rates for the highest-accuracy hyperparameter configuration of each technique in DEAP-3600.

4 Conclusion

These results confirm that I met my objective: my new techniques successfully addressed the two fundamental challenges I identified, permitting significantly improved performance while requiring little manual analysis or optimization.

On PICO-60, the new semi-supervised learning algorithms performed extremely well, exceeding 98% accuracy while nonetheless permitting impure calibration data.

On DEAP-3600, my cylindrical projection, combined with a 2D CNN, performed better than the conventional classifier on simulated data, reducing the WIMP misclassification rate from 91.0% to 75.7% while matching the conventional classifier’s 99.6% background removal rate. Additionally, the predictions made on real-world data demonstrate clear potential. Differences between the simulated and real-world data provide potentially important diagnostic information on the DEAP-3600 simulation (assuming that WIMPs are not present).

Fundamentally, the problems I have solved are not specific to PICO-60 and DEAP-3600; they are common across many dark matter detection experiments. Thus, my algorithms are broadly applicable, and demonstrate promise in the field as a whole.

5 Acknowledgements

Many thanks to Dr. Nigel Smith, Dr. Ken Clark, Dr. Carsten Krauss, Dr. Scott Fallows, Dr. Eric Vázquez-Jáuregui, and Dr. Pierre Godel for introducing me to PICO-60 and DEAP-3600, and generously providing access to data for training.

6 References

- [1] P. A. Amaudruz et al. “First results from the DEAP-3600 dark matter search with argon at SNOLAB”. In: *Phys. Rev. Lett.* 121.7 (2018), p. 071801. DOI: 10.1103/PhysRevLett.121.071801. arXiv: 1707.08042 [astro-ph.CO].
- [2] C. Amole et al. “Dark Matter Search Results from the PICO–60 C₃F₈ Bubble Chamber”. In: *Phys. Rev. Lett.* 118 (25 June 2017), p. 251301. DOI: 10.1103/PhysRevLett.118.251301. URL: <https://link.aps.org/doi/10.1103/PhysRevLett.118.251301>.
- [3] Rene Brun and Fons Rademakers. “ROOT - An Object Oriented Data Analysis Framework”. In: *AIHENP’96 Workshop, Lausanne*. Vol. 389. 1996, pp. 81–86.
- [4] François Chollet et al. *Keras*. <https://keras.io>. 2015.
- [5] Wei Dai et al. “Very Deep Convolutional Neural Networks for Raw Waveforms”. In: *CoRR* abs/1610.00087 (2016). arXiv: 1610.00087. URL: <http://arxiv.org/abs/1610.00087>.
- [6] Geoffrey E. Hinton et al. “Improving neural networks by preventing co-adaptation of feature detectors”. In: *CoRR* abs/1207.0580 (2012). arXiv: 1207.0580. URL: <http://arxiv.org/abs/1207.0580>.
- [7] J. D. Hunter. “Matplotlib: A 2D graphics environment”. In: *Computing In Science & Engineering* 9.3 (2007), pp. 90–95. DOI: 10.1109/MCSE.2007.55.
- [8] Eric Jones, Travis Oliphant, Pearu Peterson, et al. *SciPy: Open source scientific tools for Python*. 2001–. URL: <http://www.scipy.org/>.
- [9] Diederik P. Kingma and Jimmy Ba. “Adam: A Method for Stochastic Optimization”. In: *CoRR* abs/1412.6980 (2014). arXiv: 1412.6980. URL: <http://arxiv.org/abs/1412.6980>.
- [10] Martín Abadi et al. *TensorFlow: Large-Scale Machine Learning on Heterogeneous Systems*. Software available from tensorflow.org. 2015. URL: <http://tensorflow.org/>.
- [11] B. Matusch et al. “Developing a Bubble Chamber Particle Discriminator Using Semi-Supervised Learning”. In: (2018). arXiv: 1811.11308 [physics.comp-ph].
- [12] Travis Oliphant. *A guide to NumPy*. 2006.
- [13] F. Pedregosa et al. “Scikit-learn: Machine Learning in Python”. In: *Journal of Machine Learning Research* 12 (2011), pp. 2825–2830.
- [14] Guido van Rossum. *Python Reference Manual*. Tech. rep. Amsterdam, The Netherlands, 1995.
- [15] Stéfan van der Walt et al. “scikit-image: image processing in Python”. In: *PeerJ* 2 (June 2014), e453. ISSN: 2167-8359. DOI: 10.7717/peerj.453. URL: <http://dx.doi.org/10.7717/peerj.453>.

7 Bibliography

- All programming for this study was done in Python 3 [14].
- Keras [4], running on a TensorFlow [10] backend, was used for all machine learning tasks.
- NumPy [12] and SciPy [8] were used for linear algebra and signal processing.
- ROOT [3], scikit-image [15], and scikit-learn [13] were used for data loading and storage.
- Matplotlib [7] was used for data visualization.

【評語】 160044

The project studies a very advanced topic in current scientific research. The author has participated in two large experiments and his grasp of the physics and techniques are impressive. His contribution in developing new methods to enhance the data analysis is also well recognized. This is an excellent project.