Novel XENON1T Event Localization Using Graph Neural Networks

Alejandro Oranday

Advisor: Christopher Tunnell

2021-04-26

1 Introduction

Dark matter constitutes 85% of the matter in our Universe and is through observations on galaxy formation, gravitational lensing, and the cosmic microwave background [4]. Images such as the Bullet Cluster in Figure 1 show the effects of gravitational lensing and presence of dark matter by visualizing the separation of baryonic matter and high mass concentrations. Such an image is only made possible through the presence of dark matter and discussed in further detail in Sec. 2.1.

However, detecting dark matter experimentally is exceedingly difficult with particle physics detectors. They clearly interact with gravitational force, but do not interact with electromagnetic force. One candidate, the Weakly Interacting Massive Particle (WIMP), would interact through the electroweak force [4]. For us to detect WIMPs directly, one would need a detector that is sensitive to electroweak forces at keV levels of energy [4].

The XENON1T detector operated at this level of sensitivity as the most sensitive dark matter detector [1], while the newly operating XENONnT detector is currently the most sensitive dark matter detector [5]. In these detectors, there are two important elements that need to be reconstructed: the energy and the position of particle interactions. By reconstructing these key elements, we are able to accept or reject numerous observations if the reconstructed position is within the detector's fiducial volume and if the reconstructed energy is within a rejection threshold [3].

This is made possible by a defined fiducial volumes and energy thresholds in XENON1T [3]. The placement of the fiducial volume is to reject extraneous interactions that take place near the walls of the detector [3]. The shape of the fiducial volume is shown in Figure 4 as the purple line [2]. This paper will focus on the use of the fiducial volume and reconstructed positions in XENON1T.

The problem of position reconstruction, or *event localization*, is an inverse problem. Generally, the data we collect is the result of an event. Using this data, we aim to reconstruct the position of the event. This reconstructed position can then be used with the fiducial volume to filter the extraneous events. Details on the inverse problem for the XENON1T detector are given in Sec. 2.2.

However, we know the limitations of this inverse problem while machines do not. All events must take place within the detector and, for a uniform position distribution of events, the reconstructed positions should be uniform as well. This is not the case for previous position reconstruction implementations, such as multilayer perceptrons, weighted sum position, and maximum photomultiplier tube (PMT) [8]. Events are reconstructed outside of the detector, shown by the grey points outside the black line in Figure 4, and have an inward reconstruction bias [8].

In this paper, we show the novel application of a graph neural network for position reconstruction in XENON1T. Specifically, we will make use of a graph convolutional neural network (GCNN). These networks have primarily been used for classification, while the inverse problem we aim to solve is instead regression [6]. The results of this paper will show the acceptability of GCNNs for this class of inverse problems.

2 Background

2.1 Dark Matter

The effects of dark matter are apparent through various observations, such as the gravitational lensing within the Bullet Cluster [7]. This cluster is the result of the collision between two galaxy clusters. The baryonic matter is visible through X-ray images and acts as we'd expect: there is a drag between the two clusters that resulted in the Bullet Cluster's distinct shape [7]. However, the massive portions of these clusters continue to move past each other as shown by the gravitational lensing map in Figure 1. These results are indicative of both the "dark" and weakly interacting nature of dark matter. The particles are not visible through electromagnetism, thus receiving the name dark matter. The particles have a mass that causes the gravitational lensing. This upper limit on the related mass and collision cross

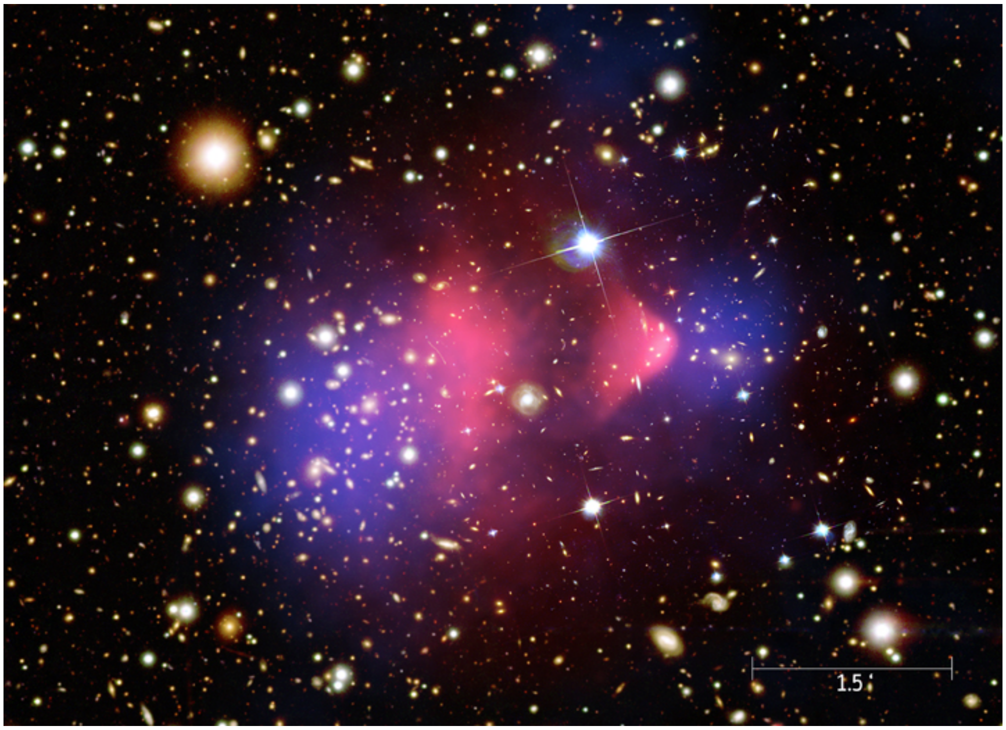


Figure 1: Composite image of the Bullet Cluster that shows the presence of dark matter through gravitational lensing. Pink regions are the X-rays of baryonic matter. Blue regions are from the gravitational lensing map. [9]

section is given by Markevitch et al. [4] as:

$$\frac{\sigma}{m} < 5 \text{ cm}^2\text{g}^{-1} \tag{1}$$

where σ is the collision cross section and m is the particle mass [7].

One candidate for dark matter particles is the weakly interacting massive particle (WIMP) [4]. This particle interacts by the electroweak force and has a lower bound of mass at the keV level [4]. Thereby, a direct detection experiment must be sensitive to nuclear recoils at the keV level. The XENON Collaboration produced the XENON1T detector for this exact purpose. The collaboration has since built and started commissioning the successor XENONnT, but the topics discussed in this paper will be with regard to XENON1T.

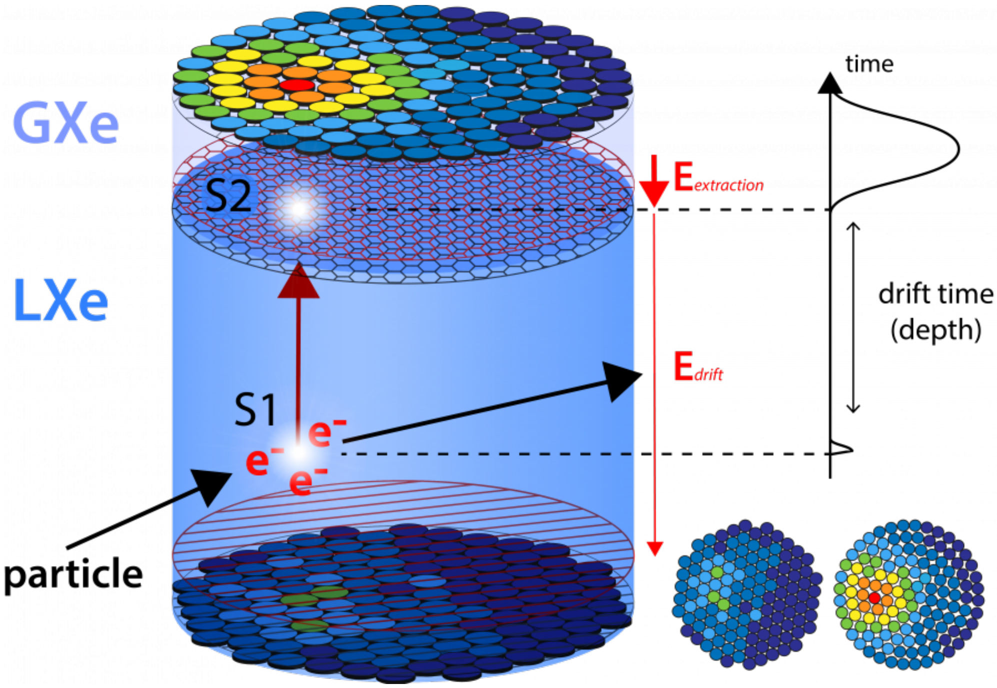


Figure 2: Diagram of the XENON1T detector and event scintillation and ionization. The first signal is the S1 where electrons and photons are emitted by the particle interaction. The electrons from the S1 are drifted upward by a constant electric field and extracted from the liquid phase to the gas phase using a stronger electric field. This produces the second signal S2 where more photons are emitted. The plane of circles at the bottom and top of the detector represent the photomultiplier tubes in the detector. The color of these circles indicate the photoelectrons observed for an S1 or S2 and is shown more clearly in Figure 3.

2.2 XENON1T Detector

The XENON1T detector is a dual phase xenon time projection chamber (TPC) located in the *Laboratori Nazionali del Gran Sasso* (LNGS) in central Italy. As stated in Sec. 2.1, it was a requirement for the detector to be sensitive to keV energy levels in order to detect WIMPs. To meet these requirements, the detector's placement in LNGS, use of xenon, and water shielding all serve to reduce the presence of extraneous signals that would be appear from muons and more radioactive media [1]. It contains a target of $\sim$1.3 tons of liquid xenon and a fiducial volume with a maximum radius of 42.84 cm [3]. The detector features 258 photomultiplier tubes (PMTs): 127 on the top array and 121 on the bottom array.

There is a constant electric field of 116.7 V/cm that causes electrons to drift towards the liquid-gas interface at the top of the detector, where they are extracted by a stronger electric field and produces a proportional scintillation signal [1].

An event in the detector is defined by a particle scattering off xenon atoms. The recoiling nuclei of the xenon atoms results in a scintillation and ionization. The scintillation is observed by the PMTs and regarded as the S1 signal. The signal produced by the ionization at the liquid-gas interface is regarded as the S2 signal. An example of this signal is shown in Figure 3. A diagram of the detector, S1, and S2 are shown in Figure 2. The resulting light pattern from the S2 signal is the observation that is used for our inverse problem. The inverse problem exactly is to use this light pattern from the top array of PMTs to reconstruct the (x, y) position of the event in reference to the plane of the PMTs. It is unnecessary to include the depth, or z position, of the event as that is achieved using the time separating the S1 and S2 and the known applied electric field. This is the problem for position reconstruction.

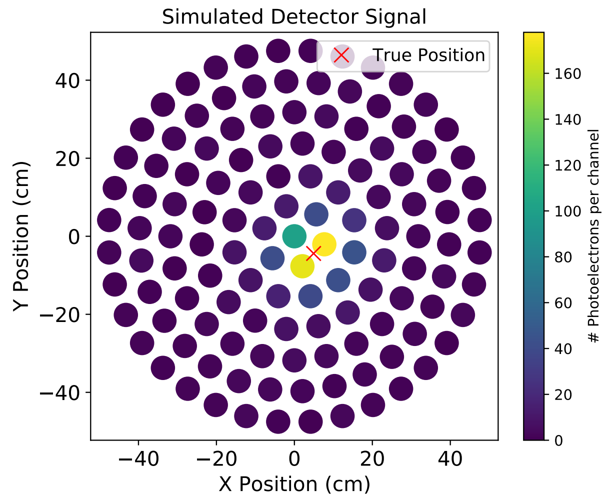


Figure 3: Example simulated hit pattern. Each circle represents a PMT in the top array of the detector. The position of each circle is the real position of the PMT in the detector. The simulation has a true position marked by the red x and created the shown hit pattern.

Using the reconstructed positions, the data from the experiment can be filtered according to the position relative to the fiducial volume. As shown in Figure 4, many of these events are reconstructed near and beyond the wall of the detector. However, reconstruction of events outside the detector is clearly not possible for this experiment. We aim to learn the detector by using a GCNN to minimize these outside reconstructions and minimize the number of reconstructions that are greater than 1 cm away from the true position when applied to

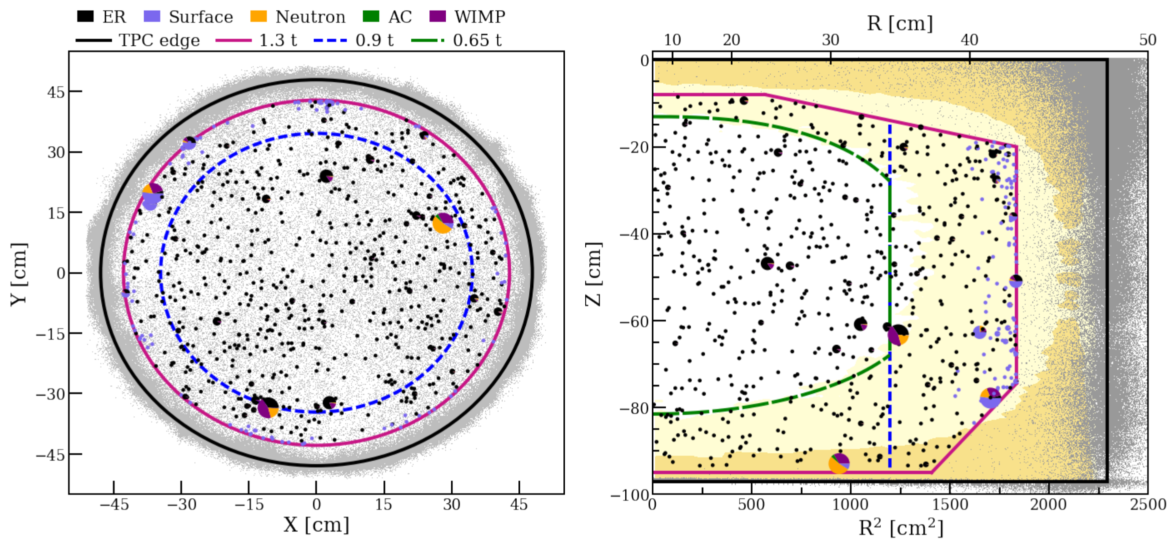
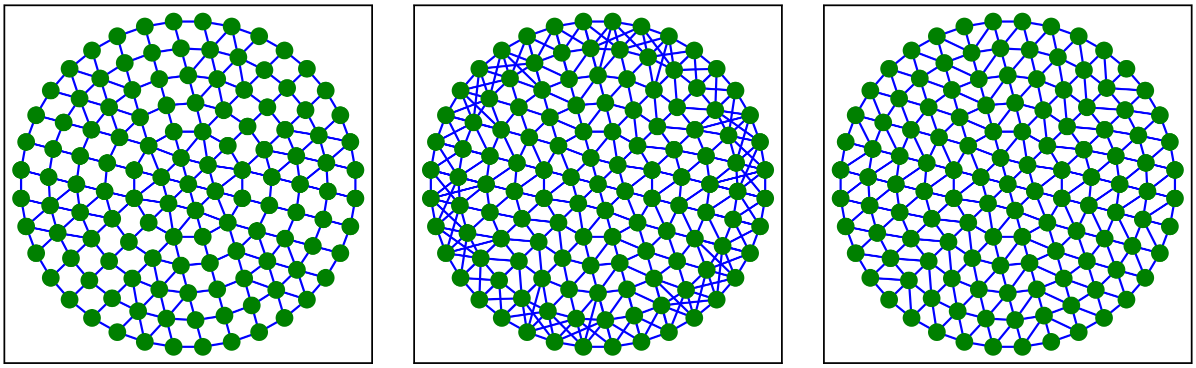


Figure 4: Spatial distribution of dark matter search data. The fiducial volume is shown in the magenta and the TPC edge is shown in black. Grey points are reconstructed events that are outside the fiducial volume. Many of these points are reconstructed outside the bounds of the detector. [2]

simulation. At the same time, this will be the novel implementation of a GCNN for regression and in particle physics.

2.3 Machine Learning

The maximum PMT and weighted sum position algorithms were two cases that did well to keep reconstructions within the detector walls. However, these algorithms did not do well for their overall accuracy. One of the initial machine learning algorithms that was applied a neural network. Recently, this neural network takes the form of a multilayer perceptron. These neural networks have performed significantly better in terms of their overall accuracy, but have difficulty keeping the reconstructions within the bounds of the detector. It is difficult to encode the detector walls into the neural network while also maintaining their good accuracy. This difficulty at the walls is due to the hard cut off of information. There are no more PMTs and thus there is no information present to inform the network that nothing is occurring outside the detector. At the same time, experimental data will hold many problems. The observed S2 signal is not clean, as in background events are present



(a) (b) (c)

Figure 5: Three of the considered graph structures: Radius $R = 10$ cm neighbors (a), k-nearest-neighbors $k = 6$ (b), Delaunay triangulation neighbors (c). Each node here is a photomultiplier tube in the top array of the XENON1T detector. The positions of each tube in the detector was used for each of the explored graph structure approaches.

within the S2 signal. Over the course of the experiment, PMTs will break and the signal they would see can no longer be used.

3 Neural Networks

Recent experiments have been collecting a significant amount of data. Machine learning has become a very useful tool when processing this data. A neural network is one section of machine learning that is heavily based off the structure of animal brains. Where an animal brain is able to learn by activating specific neurons for thoughts and actions, the neurons in a machine learn in a similar vein through various kernels and learnable parameters. These sets of neurons are divided into layers and generally show up as an input layer, some amount of hidden layers, and an output layer. For our problem, the input layer will be the signal seen by the photomultiplier tubes in the top array of the detector, and the output layer will be the (x, y) position of the interaction. We are mainly concerned with graph convolutional neural networks (GCNNs) which are a result of the success with convolutional neural networks (CNNs).

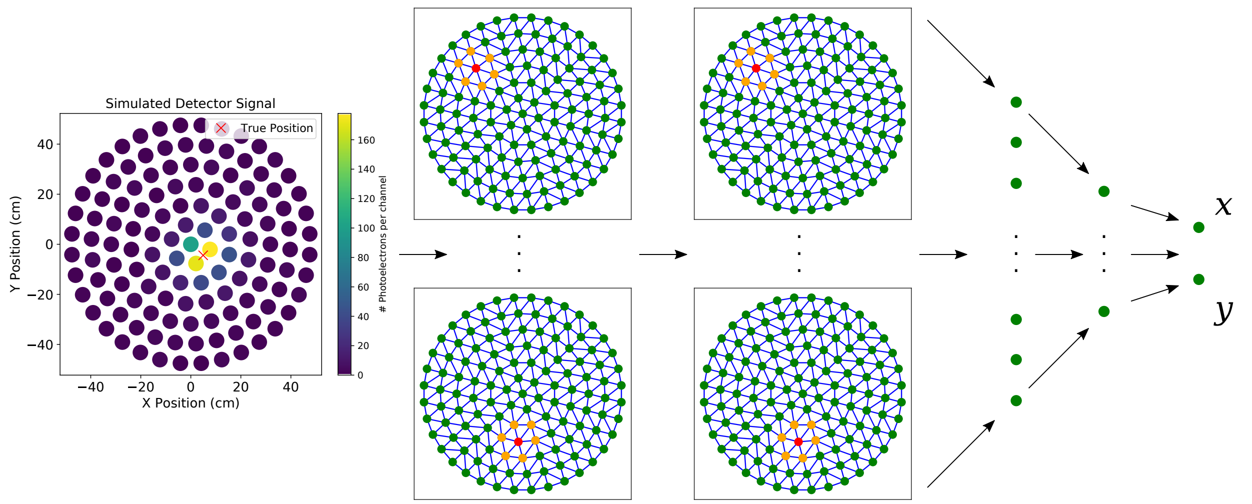


Figure 6: The graph convolutional neural network structure that the results of this paper are based on. The final structure features the signal input layer, two propagation (graph convolution) layers, two fully connected layers, and a final output layer for the (x, y) position of the interaction. Between each layer is a ReLU activation function. There are a total of 57,486 trainable parameters.
TODO: Replace the image with a sharper PDF version with less white space between the first propagation layer and the second.

3.1 Convolutional Neural Networks

A convolutional neural network is a specific kind of neural network that features a convolution layer. The input to a CNN typically comes in the form of a matrix where there is a locality between the elements in the matrix. The convolution layer makes use of a kernel that takes information from submatrices of the input matrix and summarizes the values within these submatrices. These summaries maintain the locality of the information for the respective submatrices. However, the structure of our dataset does not have a form that can be easily made into a matrix. Therefore, we would need an algorithm that is capable of making use of datasets with any possible structure.

3.2 Graph Convolutional Neural Networks

The design of the graph convolutional neural network (GCNN) came by considering how a convolution could be applied to a graph structured dataset [6]. The graph convolution layer that we use and that was proposed by Kipf and Welling propagates the values of nodes according to the edges [6]. The value of connected nodes will increase according to the values of the connected nodes while nodes that are disconnected will see no change. The exact propagation rule is given by the equation:

$$H^{(l+1)} = \sigma\left(\tilde{D}^{-1/2}\tilde{A}\tilde{D}^{-1/2}H^{(l)}W^{(l)}\right) \tag{2}$$

where $H^{(0)}$ will be our initial input signal, $\tilde{A}$ and $\tilde{D}$ are respectively modified, unweighted adjacency and modified degree matrices, $W^{(l)}$ is the trainable weights matrix for the l^{th} layer, and $\sigma(\cdot)$ is an activation function [6]. The modified adjacency and degree matrices are to include self-loops. This means that nodes will connect to themselves and propagate values to themselves when passed through the graph convolution layers. The modified adjacency and degree matrices take the form:

$$\tilde{A} = A + I_N$$
$$\tilde{D} = D + I_N$$

where I_N is a $N \times N$ identity matrix where N is the number of nodes.

We considered the PMTs of the XENON1T detector as our nodes and their quantity of light collected as their primary value. We also included the (x, y) position of the PMT at the top of the detector as addition values. When translating to equation (2), the matrices take the form:

$$\tilde{A}, \tilde{D} \in \mathbb{R}^{127\times127}$$
$$H^{(0)} \in \mathbb{R}^{127\times3}$$
$$H^{(1)}, W^{(0)} \in \mathbb{R}^{127\times w}$$

where w will depend on our choice of network structure. We chose to use rectified linear units (ReLU) as our activation function for all layers.

With regard to the network structure, we were influenced by the success of image classifiers, such as AlexNet. However, convolutional neural networks of this structure and graph convolutional neural networks in general are more often used for classification, while we have a regression problem. Our reasoning to use a GCNN for predicting the position of interactions came from being able to encode the local structure of the XENON1T detector into the dataset. By treating the nodes of our graph as the PMTs at the top of the detector, we understood that the connections or edges that we put in place would maintain the local structure if done carefully. We considered several graph structures shown in Figure 5. We ultimately chose the Delaunay Triangulated graph for it's consistent connection density throughout the graph: only PMTs that are immediately near each other are connected and resulted in most nodes having 6 edges. Only nodes that represent PMTs near the wall of the detector had degrees less than 6. There is potential for finding the graph structure that describes the detector's data best, but for this we chose to go with a heuristic approach.

4 Results

4.1 During Training

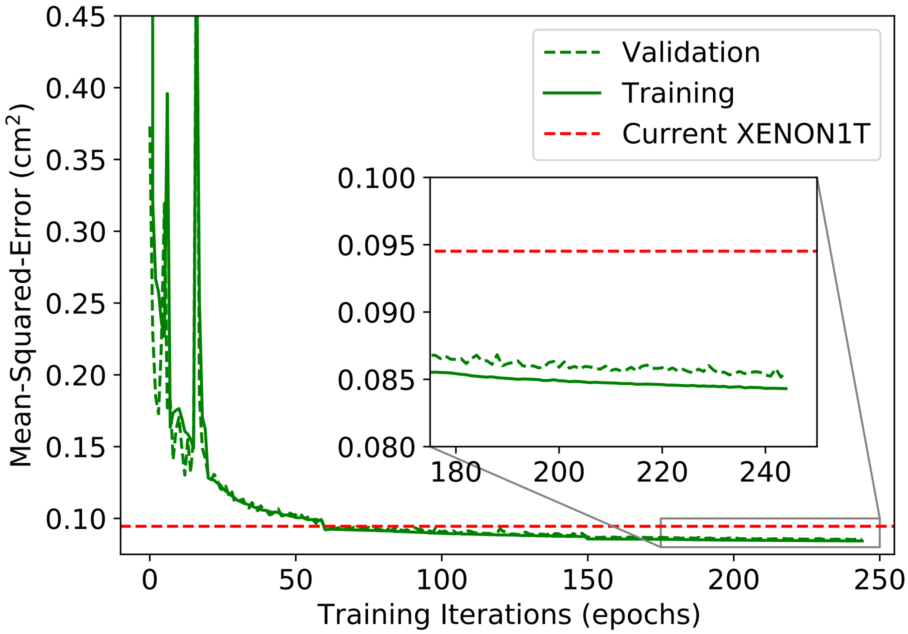


Figure 7: Mean-Squared-Error (MSE) in square centimeters of our algorithm for the training set and validation set during the training process. The minimum MSE for the current state-of-the-art is given in red at 0.0945 cm^2. This is the benchmark that our algorithm had to pass during training. Spikes within the first 20 epochs occur due to large step size during gradient descent. Learning rate was lowered at epochs 20, 60, and 150. The minimum MSE achieved by our GCNN on the validation set is 0.0852 cm^2.

The performance of our algorithm was compared to the current state-of-the-art in XENON1T during training as a benchmark and early warning system. If the GCNN did not approach a comparable performance to that of the state-of-the-art swiftly enough, training would typically stagnate and not surpass this benchmark. By not performing better here, it was generally indicative that the GCNN would also perform worse when we gave attention to our performance metrics. After a few iterations of this, we chose to only look at the performance metrics if the GCNN model produced a lower mean-squared-error during training and would restart training in cases where it was clear that the current iteration would not perform better within a reasonable number of epochs. An example of when it was clear a prospective model would not do better is if the mean-squared-error was not below 0.25 cm^2 within the first 50 epochs.

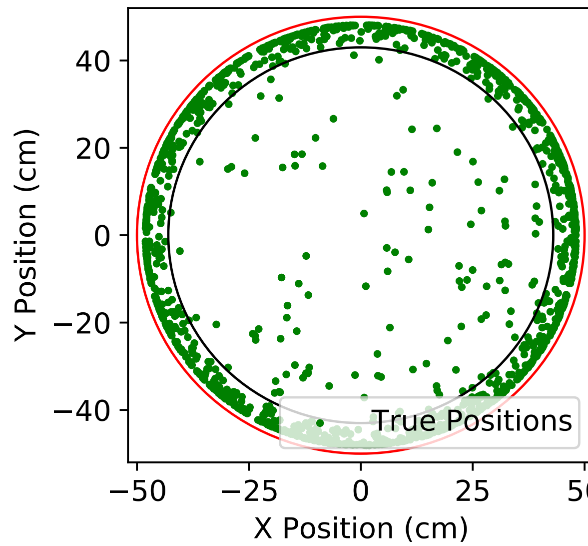


Figure 8: True positions of GCNN mis-reconstructions. 1,680 of the 197,975 simulated events were mis-reconstructed and are shown here. Red circle is the wall of the detector (50 cm); black circle is the largest radius of the fiducial volume (43 cm). 123 of the 1680 mis-reconstructions are within the fiducial volume.

We used an optical Monte Carlo simulation of 989,875 events for training as an attempt to assume a "perfect" detector. This is to say that no spurious events, such as single electrons, dark counts, or PMT after-pulses, were within our simulation. The observations by the PMTs are as if every part of the detector ran perfectly. By using a simulation like this, we were able to input the data into our model without normalization or standardization.

Our algorithm was able to outperform the state-of-the-art in training, which is a good indicator for the overall performance. Much of the work for this stage was in optimizing the learning rate used for gradient descent. Our solution was to lower the learning rate at specific epochs based on the performance of previous results. Specifically, the learning rate was lowered at epochs 20, 60, and 150. This caused notable dips within Figure 7 and resulted in a much smoother curve after epoch 20. However, a better solution would have the learning rate lower based on the model's performance during training instead of milestones set by the attentive user.

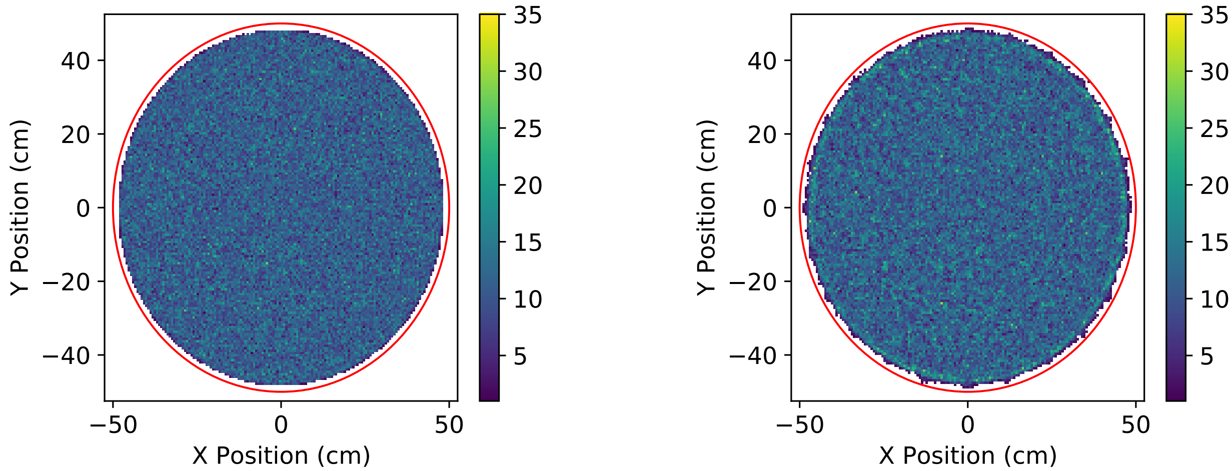


Figure 9: 2D histograms of the true positions (left) and reconstructed (right) positions at 150 bins. Red circle is the wall of the detector (50 cm). The edge of the reconstructed positions is noticeably more jagged.

4.2 Validation Set Performance

As previously stated, the two performance metrics we focused on are to have no reconstructions outside the detector and to minimize the number of reconstructions that are 1 cm away from the true position. For best practice in machine learning, we focused on the results of the validation set which is made of 197,975 simulated events.

Since we are hard set on having no reconstructions outside of the detector, this was the first metric we would check. As it turned out, we counted zero reconstructions outside of the detector for our latest version of the GCNN. At this point, our algorithm has successfully surpassed the state-of-the-art training benchmark and made no exceedingly erroneous reconstructions, a rule that previous implementations had difficulty passing.

As for the further than 1 cm reconstructions, these too performed well. Of the 197,975 events, 1,680 were reconstructed at greater than 1 cm away from the true position, about 0.85% of the validation set. As can be seen in Figure 8, many of the mistakes are made along the walls of the detector and explains the jagged edge found in Figure 9. If we reduce the area we count on to the maximum radius of the fiducial volume ($R = 43$ cm), we find only 123 of the 197,975 events mis-reconstructed, 0.06% of the validation set.

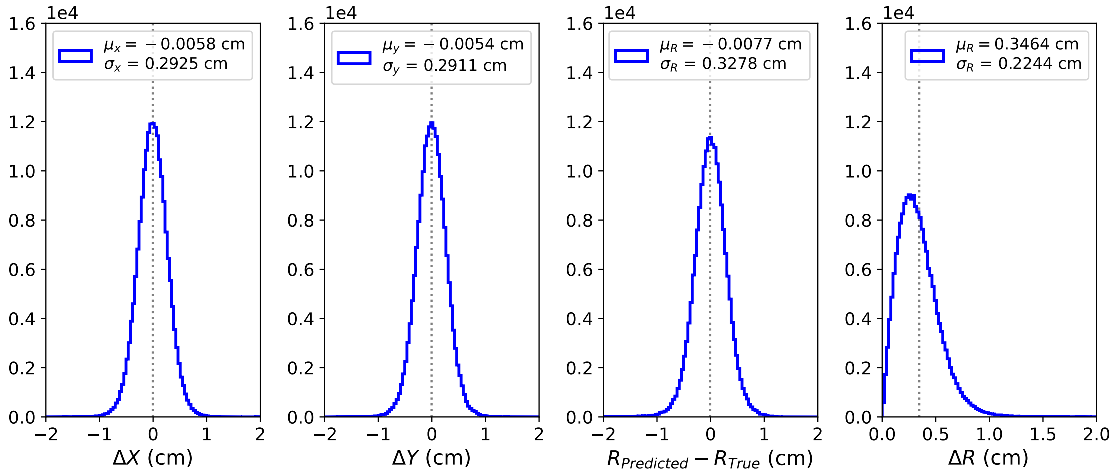


Figure 10: 1D histograms of the reconstructed position minus the true positions. There are 100 bins between -2 cm and 2 cm on $\Delta X, \Delta Y, (R_{\text{Predicted}} - R_{\text{True}})$ and 100 bins between 0 cm and 3 cm on ΔR. The left three histograms are near Gaussian curves of the same statistics.

The last important performance check, as for any experiment, is to produce the most accurate measurements or reconstructions, in our case. For this we use the resolution metrics ΔX, ΔY, and ΔR:

$$\Delta X \equiv X_{\text{Reconstructed}} - X_{\text{Simulated}}, \quad \Delta Y \equiv Y_{\text{Reconstructed}} - Y_{\text{Simulated}},$$
$$\Delta R \equiv \sqrt{\Delta X^2 + \Delta Y^2}$$

where X and Y are the x and y positions of the reconstructions and the associated simulation. The means and standard deviations produced by our algorithm are shown in Figure 10. This too outperformed the state-of-the-art which had standard deviations greater than 3 cm. From previous observations of the results of our GCNN, the mean and standard deviation of ΔR follows suit with what we expect: most of the reconstructions are within 1 cm of the true, simulated position. At the same time, the approximately-Gaussian curves of ΔX, ΔY, and radius difference further confirms the positive performance of our GCNN.

5 Conclusion & Future Work

We presented one of the first applications of a graph neural network – more specifically a graph *convolutional* neural network – in particle physics as well as one of the first regression applications of such a network. Our algorithm worked well enough to compete with the state-of-the-art position reconstruction algorithm for XENON1T. However, we do not think that this algorithm is the best approach to take in the future. We suspect that a broader graph neural network would work better for position reconstruction for the upcoming XENONnT experiment. The use of a graph neural network has shown the ability to maintain locality between PMTs, our graph's nodes, but is not attributed to the use of convolutions. We believe there is a better approach that can be more specific to the XENON1T detector that does not necessarily use convolutions. The problem we realized when applying to experimental data is the overbearing presence of background events. The graph convolutions compound the presence of these background events and create more difficulties than initially anticipated.

We have shown that graph neural networks are applicable to particle physics detectors. They are able to preserve the locality of sensors in the detector when necessary for the application. Constructing these graphs will be unique for the problem to be solved and the detector itself, and creates a potential area of research to find the most optimal graph structure. We have shown that graph neural networks can be applied as both classifiers and regressors with significant effectiveness.

6 Appendix A: Graph Convolutional Neural Networks

Place holder.

References

[1] E. Aprile, J. Aalbers, F. Agostini, et al. First dark matter search results from the xenon1t experiment. *Physical Review Letters*, 119(18), Oct 2017. ISSN 1079-7114. doi: 10.1103/physrevlett.119.181301. URL `http://dx.doi.org/10.1103/PhysRevLett.119.181301`.

[2] E. Aprile, J. Aalbers, F. Agostini, et al. Dark matter search results from a one ton-year exposure of xenon1t. *Phys. Rev. Lett.*, 121:111302, Sep 2018. doi: 10.1103/PhysRevLett.121.111302. URL `https://link.aps.org/doi/10.1103/PhysRevLett.121.111302`.

[3] E. Aprile, J. Aalbers, F. Agostini, et al. Xenon1t dark matter data analysis: Signal reconstruction, calibration, and event selection. *Physical Review D*, 100(5), Sep 2019. ISSN 2470-0029. doi: 10.1103/physrevd.100.052014. URL `http://dx.doi.org/10.1103/PhysRevD.100.052014`.

[4] Gianfranco Bertone and Dan Hooper. History of dark matter. *Reviews of Modern Physics*, 90(4), Oct 2018. ISSN 1539-0756. doi: 10.1103/revmodphys.90.045002. URL `http://dx.doi.org/10.1103/RevModPhys.90.045002`.

[5] The XENON collaboration, E. Aprile, J. Aalbers, et al. Projected wimp sensitivity of the xenonnt dark matter experiment, 2020.

[6] Thomas N. Kipf and Max Welling. Semi-supervised classification with graph convolutional networks. *CoRR*, abs/1609.02907, 2016. URL `http://arxiv.org/abs/1609.02907`.

[7] M. Markevitch, A. H. Gonzalez, D. Clowe, et al. Direct constraints on the dark matter self-interaction cross section from the merging galaxy cluster 1e 0657-56. *The Astrophysical Journal*, 606(2):819–824, May 2004. ISSN 1538-4357. doi: 10.1086/383178. URL `http://dx.doi.org/10.1086/383178`.

[8] B. E. J. Pelssers. Position reconstruction and data quality in xenon. Master's thesis,

University of Utrecht, July 2015. URL https://dspace.library.uu.nl/handle/1874/322783.

[9] X ray: NASA/CXC/CfA/M.Markevitch et al.; Optical: NASA/ST-ScI; Magellan/U.Arizona/D.Clowe et al.; Lensing Map: NASA/STScI; ESO WFI; Magellan/U.Arizona/D.Clowe et al. Bullet cluster, 2006. URL https://chandra.harvard.edu/photo/2006/1e0657/more.html.