Graph Convolutional Neural Networks for Position Reconstruction in the XENON1T Experiment

Alejandro Oranday

Advisor: Christopher Tunnell

aeo3@rice.edu

March 23, 2021

Contents

Introduction

Dark matter constitutes 85% of the matter in our Universe and was made evident by observations on galaxy formation, gravitational lensing, and the cosmic microwave background [3]. However, detecting dark matter experimentally is exceedingly difficult with particle physics detectors. These particles are suspected to be weakly interacting [2] such that any detection attempt would need to be sensitive to recoils at keV levels of energy.

The XENON1T detector operated as the most sensitive dark matter detector [1], and the soon-to-be-active XENONnT detector plans to overtake that title [4]. In these detectors, there are two important elements that need to be reconstructed: energy and position. 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 [2].

Previous machine learning implementations for position reconstruction perform well enough [5], but still run into issues with reconstruction outside the bounds of the detector as well as having an inward reconstruction bias. Finding the type of machine learning algorithm that is the most appropriate for these problems is essential for the best use of the detector's fiducial volume. This project produced the first application of graph convolutional neural networks (GCNNs) for position reconstruction in the dark-matter field, and one of the first applications of a GCNN for use in regression.

Background

Dark Matter

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. The detector aimed to observe weakly interacting massive particles (WIMPs) as the primary candidate for dark matter particles. It was a requirement for the detector to be sensitive to keV energy levels in order to observe these particles and was made possible through a combination of the stable xenon 136 isotope, water shielding, and depth within the Gran Sasso massif.

Machine Learning

Results

During Training

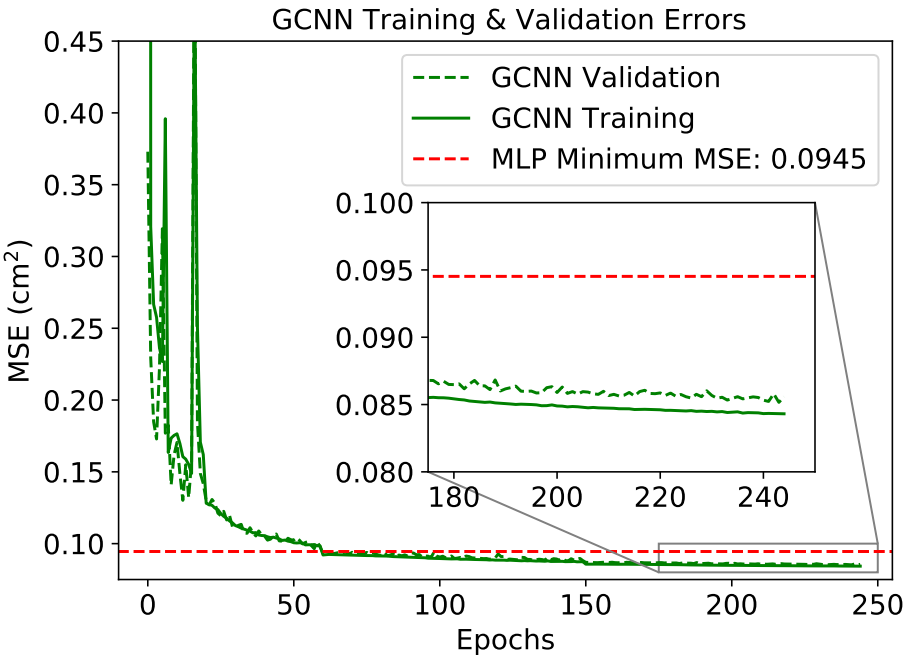

Figure 1: Mean squared error (MSE) in square centimeters of the GCNN for the training set and validation set during the training process. The minimum MSE for the MLP is given in red at 0.0945 cm^2. This is the benchmark that the GCNN has to pass during training. Spikes within the first 20 epochs occur due to large step size during gradient descent. This was temporary solved by lowering the step size at epochs 20, 60, and 150. These milestones were placed by observations of previous trainings, but the permanent solution will be based on the MSE as it's training. The minimum MSE achieved by our GCNN on the validation set is 0.0852 cm^2.

The performance of our GCNN was compared to a MLP during training as a benchmark and early warning system. If the GCNN did not approach a comparable performance to that of the MLP swiftly enough, training would typically stagnate and not surpass this benchmark. By not performing better than the MLP 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 MSE during training and would restart training in cases where it was clear that the GCNN would not perform better

within a reasonable number of epochs. An example of when it was clear the GCNN would not do better is if the MSE was not below 0.25 cm^2 within the first 50 epochs.

We used an optical Monte Carlo simulation 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.

The GCNN was able to outperform the MLP 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. We ended up lowering the learning rate at epochs 20, 60, and 150. This caused notable bumps within Figure 1 and resulted in a much smoother curve after epoch 20. However, a better solution would have the learning rate lower based on the GCNN's performance during training instead of milestones set by the attentive user.

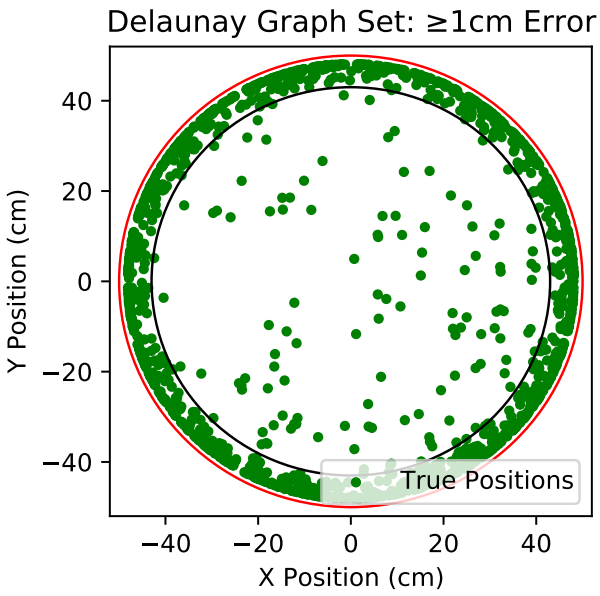

Figure 2: True positions of GCNN mis-reconstructions. Only 1680 of the 197,975 simulated events were mis-reconstructed and are shown here. Red circle is the wall of the detector (50 cm); black circle is the largest radius of the fiducial volume (43 cm). It's clear that the majority of these mis-reconstructions occur near the wall and are beyond the largest radius of the fiducial volume. Only 123 of the 1680 mis-reconstructions are within the fiducial volume.

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

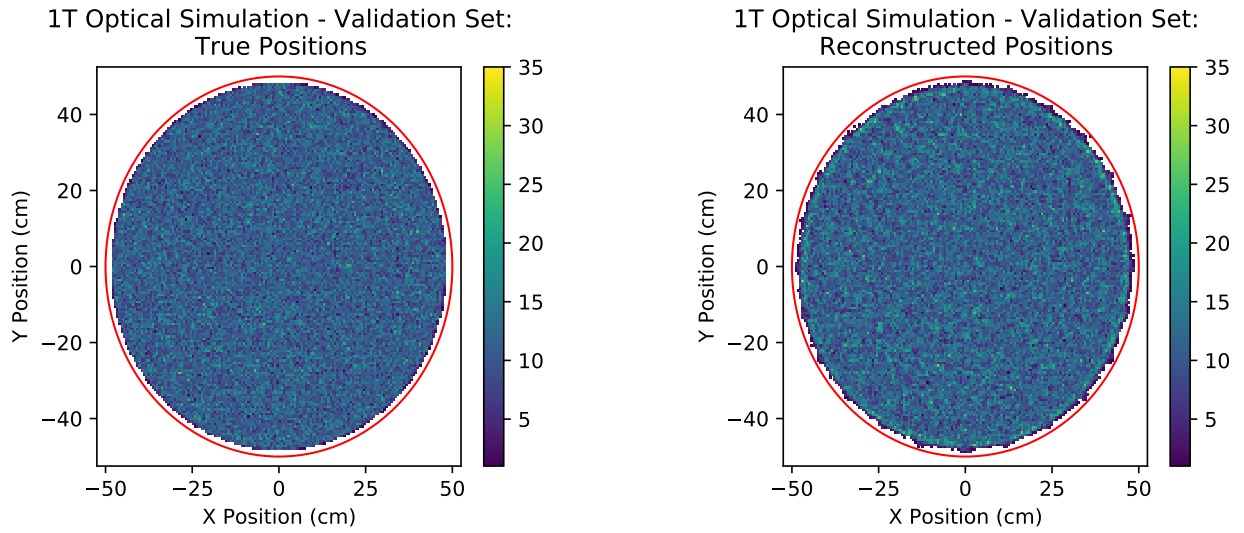

Figure 3: 2D histograms of the true positions and reconstructed positions at 150 bins. Red circle is the wall of the detector (50 cm). The edge of the reconstructed positions is notably jagged, but entirely contained within the walls of the detector. This agrees with our expectations of doing poorly near the wall of the detector while still succeeding in containing the reconstructions within the detector. Otherwise, the uniform reconstruction is a good indicator that there aren't any particular spots of inaccuracy nor is there any outstanding inward reconstruction bias.

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 GCNN has successfully surpassed the MLP 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. Only 1680 of the 197,975 events were reconstructed at greater than 1 cm away from the true position. That is about 0.85% of the validation set. As can be seen in Figure 2, many of the mistakes are made along the walls of the detector and explained the jagged edge found in Figure 3. 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. Only 0.06% of the validation set!

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:

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

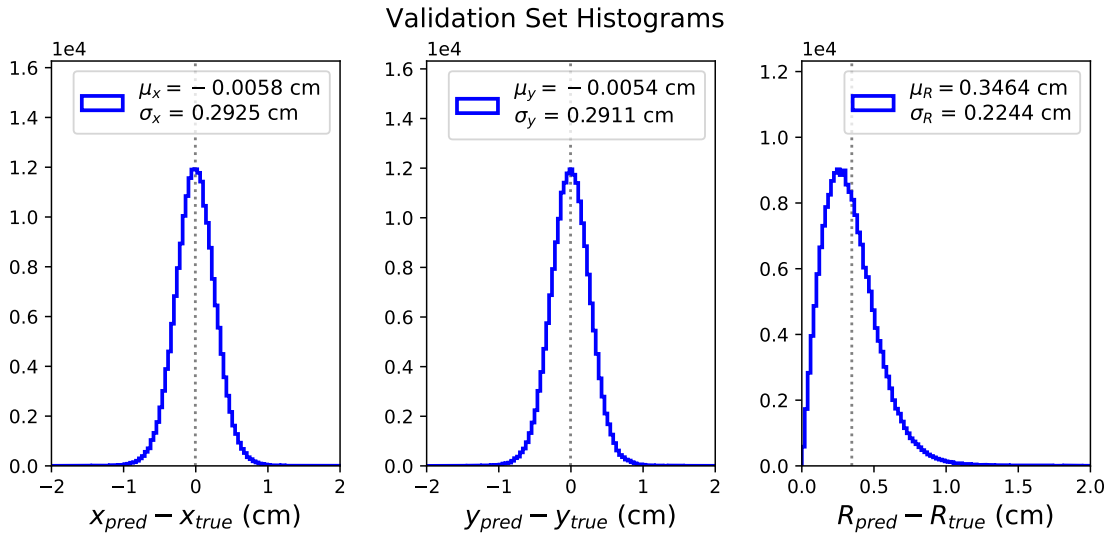

Figure 4: 1D histograms of the reconstructed position minus the true positions. There are 100 bins between -2 cm and 2 cm on $\Delta X, \Delta Y$ and 100 bins between 0 cm and 3 cm on ΔR. For ΔX we achieved a mean value of -0.00578 cm and a standard deviation of 0.292 cm. For ΔY we achieved a mean value of -0.00536 cm and a standard deviation of 0.291 cm. For ΔR we achieved a mean value of 0.346 cm and a standard deviation of 0.224 cm. Both the ΔX and ΔY histograms are near Gaussian curves of the same statistics.

where X and Y are the x and y positions of the reconstructions and the associated simulation. As for the means and standard deviations produced by our GCNN are shown in Figure 4. We found ΔX with a mean of -0.00578 cm and standard deviation of 0.292 cm, ΔY with a mean of -0.00536 cm and standard deviation of 0.291 cm, and ΔR with a mean of 0.346 cm and a standard deviation of 0.224 cm. This too outperformed the MLP 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 and ΔY further confirms the positive performance of our GCNN.

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. 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.

[3] 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.

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

[5] 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.