CS230

Position reconstruction in the CDMS HV detector

To Chin Yu
Department of Physics
Stanford University
ytc@stanford.edu

Shridevi S. Muthkhod
Department of Computer Science
Stanford University
shridevi@stanford.edu

Abstract

We demonstrate that neural network can an effective tool to use for position reconstruction in cryogenic phonon detectors. With simulated data we achieve a 2.5 mm z resolution in 33 mm thick CDMS HV100mm detector. This beats the nominal resolution goal of 10 mm by a factor of 4.

1 Introduction

CDMS (Cryogenic Dark Matter Search) is an experimental collaboration looking for dark matter particle candidates. The detector consists of a semiconductor crystal instrumented with superconducting sensors placed under a cryogenic enviroment (T≈30mK). The resulting device is extremely sensitive to any kind of energy deposited within the crystal. Therefore it is extremely important to be able to reject any uninteresting background events.

Currently a majority of the background comes from surface events such as electrons or lead atoms interacting with the detector surface (figure 1). These surface events greatly degrade detector performance and effective surface rejection (fiducialization) is extremely valuable.

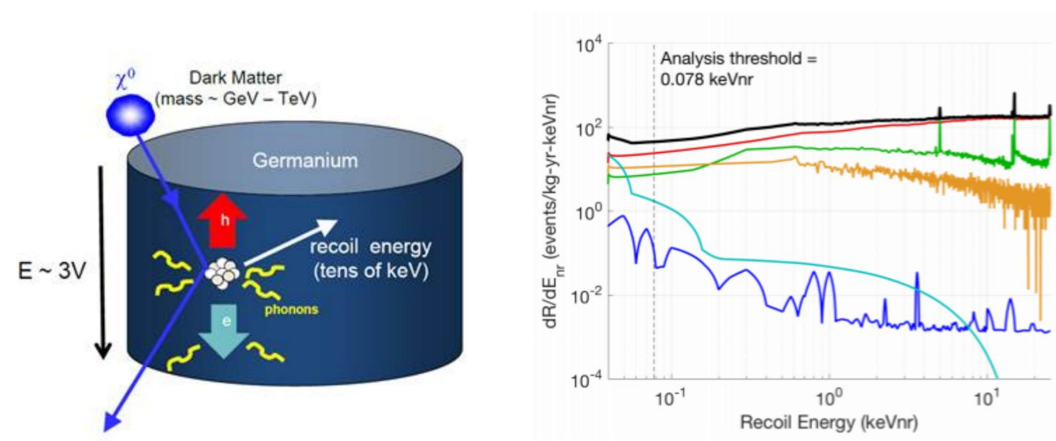

Figure 1: (left) Illustration of detector working principle. (right) A plot showing the different backgrounds of the detector. Thick black line represents total background. Red is detector impurity. Green and orange are surface beta and surface lead. Cyan and blue are neutrino and neutron backgrounds.

CS230: Deep Learning, Winter 2018, Stanford University, CA. (LateX template borrowed from NIPS 2017.)

Since the sensor data is read out as voltage traces and in principle carries all the information about the energy deposition, including any positional information, it should be possible for a machine learning algorithm to reconstruct the position just from looking at the voltage traces. Achieving a fine position resolution and thus an effective fiducial cut can potentially improve the signal-to-background ratio by up to 2-20x depending on the impurity level of the crystal. Positional information can also be used to traceback the origin of high-rate background and further improve the operating environment of the detector. A neural network implementation can also be introduced at trigger level and increase bandwidth for potential dark matter candidates.

The input of our network will be the time series (1000 timesteps) of voltage readings from the 12 detector channels ($12 \times 1000 = 12000$ inputs). We feed the voltages into a fully connected neural network to either z (single) or x, y, z (multi) coordinates of the energy deposit, relative to the center of our detector.

2 Related work

Position reconstruction in cryogenic phonon detectors is not a particularly well-developed area. However we did find a paper on shower reconstruction from time series of detector readout voltages along with other meta-information. There were also internal efforts in the CDMS to do position reconstruction using e.g. template fitting. None of these methods are fully satisfactory both in terms of performance as well as complexity.

3 Dataset and Features

The dataset we use is generated by the Monte-Carlo simulation program SuperSim developed internally by the CDMS collaboration. The simulation recreates the sensor response for energy hits at inputted positions. An example dataset set is shown in figure 2 where we generated hits with varying z-coordinates while keeping the radial position fixed at (0,0). The y-axis is the normalized voltage (ranges from 0 to 1) and the x-axis represents time steps. Each hit is formatted as a single vector with 12000 numbers before inputting into our network.

For this project we generated 15000 hits at random positions in the detector. The train/dev/test split is 75/12.5/12.5. We also added a small ($\sigma = 0.001$) Gaussian noise to the voltage.

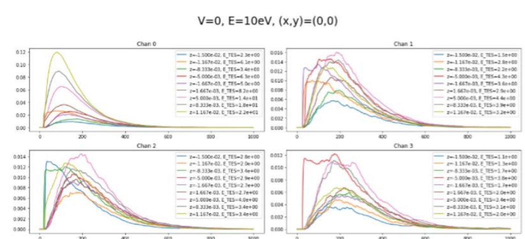

Figure 2: An example set of traces. The different colors represent different z positions. Only 4 out of 12 channels are shown here.

4 Methods

Early rounds of training suggest that a simple fully-connected network might already be sufficient for our purpose. We used Adam optimization to train our network. Originally we used the MSE loss but we noticed that the trained networks have a non-zero bias which is unphysical. Therefore we modified the loss function by adding a bias penalty:

$$L = \sum_i \left| y_i - \hat{y}_i \right|^2 + \lambda \left| \sum_i (y_i - \hat{y}_i) \right|^2 \tag{1}$$

This additional term significantly reduced the bias of the prediction. Note that y_i, $\hat{y}_i$ can either be real numbers (single-coordinate) or 3-vectors (multi-coordinate).

Parameter	Values
#Layers	1-5
#Hidden units	5-25
Learning rate	10^{-4} - 10^{-2}
Decay rate	10^{-4} - 10^{-1}
λ	1-10
Epoch	30 - 200
Mini-batch size	32,64,128,256,512

Table 1: List of hyperparameters used in training.

5 Results

We trained two networks for single-coordinate prediction and multi-coordinate prediction. The hyperparameters we chose are listed in table 1.

For our purpose we are most interested in rejecting events on the top and bottom surfaces, thus the most important figure of merit is the resolution of the reconstructed z position σ_0. This is calculated by first binning $z_i - \hat{z}_i$ into a histogram and fitting the resulting distribution by a Gaussian function with zero mean. The model with the smallest σ_0 is selected during hyperparameter tuning.

The results are summarized in figure 3. In the end we have achieved a final resolution of about 2 mm over a 33 mm tall detector volume, giving us close to 90% fiducialization. This is extremely good given that the collaboration has always been assuming a 10 mm z resolution and no fiducialization.

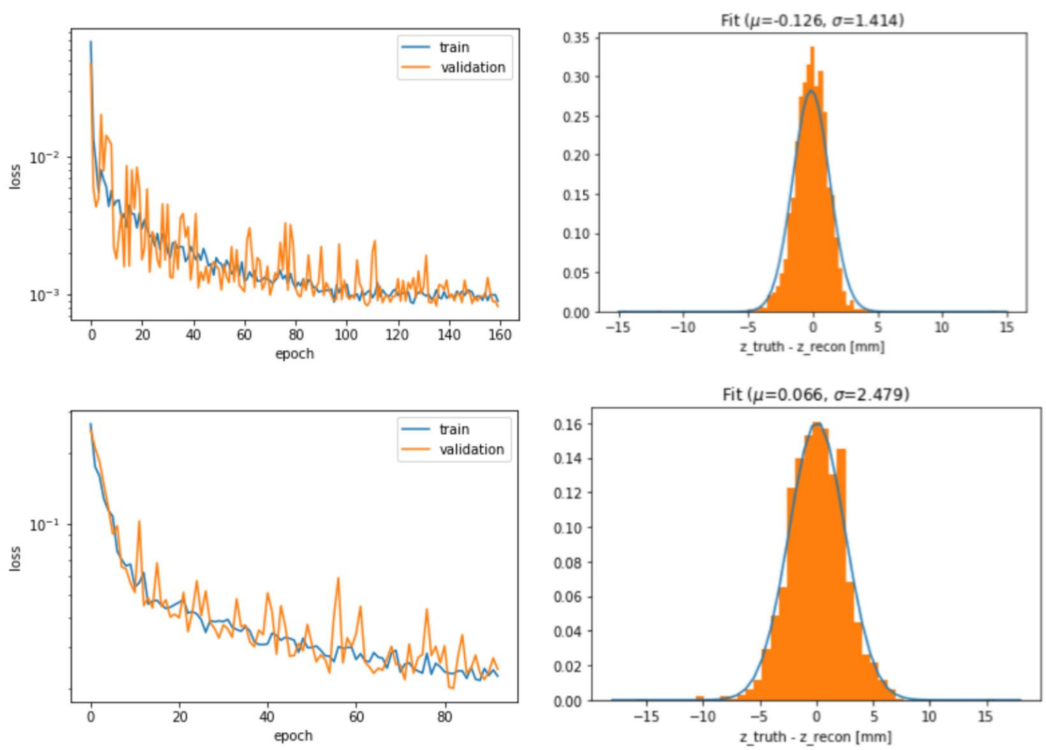

Figure 3: Loss and final fitted resolution for the (top) single-coordinate and (bottom) multi-coordinate networks.

6 Future Work

This work is still far from complete. Our simulated traces do not contain many of the problematic features in the real data such as noise bursts, pile-ups and random time shifts (figure 4). To deal with the complexity of real-life traces, a more complicated architecture (e.g. CNN) has to be used. Unfortunately we have ran out of time before we could generate synthesized traces with pile-ups, etc and experiment with more complicated architectures.

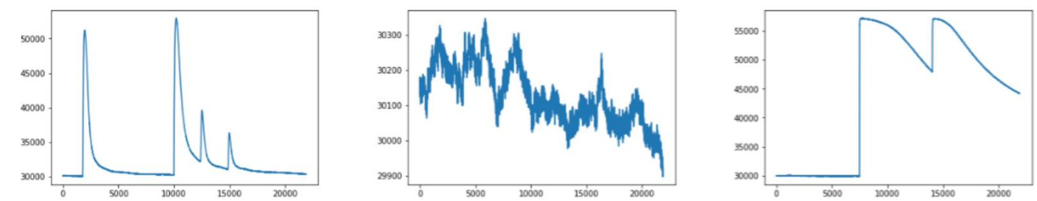

Figure 4: Examples of real traces which contain (left) multiple pulses or pile-ups, (center) noise bursts and (right) a different pulse shape.

7 Contributions

TCY was responsible for data preparation, some early implementation of the algorithm, evaluation of network performance and report writing. SM was responsible for implementation of the algorithm on AWS, training and tuning the models and preparation of the final poster.

References

[1] SuperCDMS Collaboration (Agnese, R. et al.), Projected Sensitivity of the SuperCDMS SNOLAB experiment, Phys.Rev. D95 (2017) no.8, 082002, arXiv:1610.00006

[2] Erdmann, M. et al., A Deep Learning-based Reconstruction of Cosmic Ray-induced Air Showers, Astropart.Phys. 97 (2018) 46-53, arXiv:1708.00647

[3] Abadi, M. et al., TensorFlow: A system for large-scale machine learning, arXiv:1605.08695