Real Time HG DREAM Calorimeter Analysis Using AI

Janna Goodman[1,2], Matteo Cremonesi[1], Peter Meiring[1]

[1]Carnegie Mellon University; [2]Cornell University

Introduction

In particle detectors, calorimeters are used to detect and identify particles through electromagnetic and strong interactions with the calorimeter material. In hadron calorimetry, where energy is degraded and detected through strong interactions, much of the hadron-nuclei interactions are undetectable. This is due to energy lost from nuclear dissociation, delayed particles, and meson decay and absorption. By tuning the ratio of responses from electromagnetic and hadronic signals to be one to one and measuring the electromagnetic energy, we can accurately calculate the hadronic part. DREAM, a dual readout calorimeter, measures the electromagnetic component by looking at two signals, a scintillating signal proportional to energy degradation from ionization and a Cherenkov signal in plastic fibers proportional to the energy carried by relativistic particles.

Figure 1: Bundles of scintillating and plastic fibers in the original DREAM prototype

Recently, the original DREAM detector has been significantly improved by using Silicon Photomultipliers (SiPMs), allowing the original detector to be reconfigured into smaller transverse towers as well as improving segmentation by the timing of particle pulses. This allows for 24 times higher granularity. The improved detector is called the High Granularity Dream (HG DREAM) detector.

Figure 2: Hexagonal transverse towers in the original DREAM detector

Our goal is to see how well a Convolutional Neural Network (CNN) can use data received from an HG DREAM detector to identify and differentiate each particle pulse. We also looked at how well this model can be converted onto a chip to be applied to the front end of future colliders. We want to optimize for efficiency, accuracy, and latency.

Methodology

We used a 1 dimensional CNN with 10 hidden activation layers of 64 and 32 neurons and 1 neuron for the final hidden layer. We trained this model on SiPM pulse data we received from our collaborator Jordan Damgov from Texas Tech University. We then added a variable amount of Gaussian noise to the data.

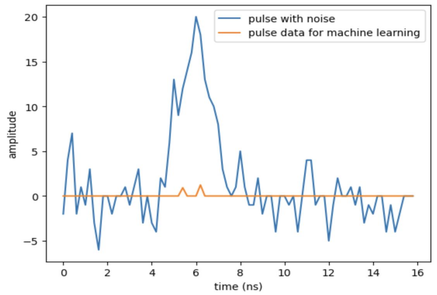

Figure 3: The orange is the model input and the blue is the pulses with added noise. The pulse data is observed over 16 nanoseconds divided into 80 time segments. A similar graph can be produced for each “event”- our model was trained on 1500.

In order to meet the strict latency requirements of less than one microsecond, we plan to synthesize our model onto highly specialized chips called Field Programmable Gate Arrays (FPGA). These circuits will eventually be attached to the next generation of particle accelerators, which will likely be lepton colliders.

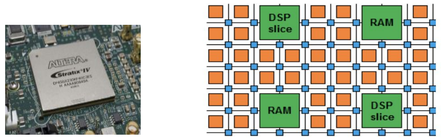

Figure 4: (Left) An image of an integrated circuit like an FPGA. (Right) A schematic drawing of the components of the FPGA. The orange are logic cells which perform logical functions. The DSP slices perform arithmetic and the RAM cells store memory.

To synthesize our CNNs onto FPGAs we use a package called hls4ml. In addition to translating the model onto the FPGA, hls4ml allows us to optimize latency and resource use and increase computational efficiency through tools like pruning and quantization.

Results

To ensure that our model was actually learning, we plotted the loss function over each round of training or epoch. If our model is working, we should see a decaying loss function that eventually plateaus when the model is finished training.

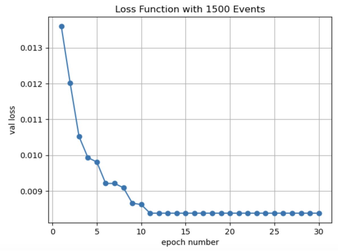

Figure 5: A decay function for our training. The exponential decay shows that our model is successfully learning.

After training our model using the Keras software package, we converted the model to hls4ml. To maximize the agreement between the hls model and keras model, we need to create the hls model with enough bits of precision to cover the range of values in each layer of the neural network. We experimented with manually adjusting the precision of our model.

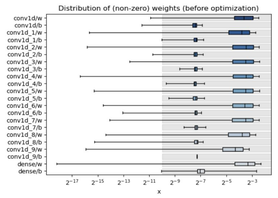

Figure 6: The range of values at each layer of our neural network. Most layers need at least 10 decimal bits of precision to properly train. The gray here represents the range of values that can be described with <16,6> precision (16 total bits, 6 integer bits).

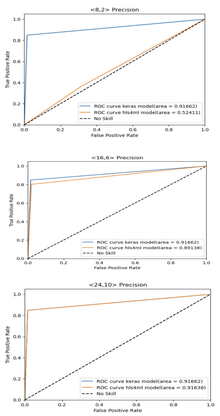

Figure 7: ROC curves plotting the true positive against the false positive for increasing precision. The first number in the brackets represents the total number of bits and the second number is the amount of integer bits. At least 10 decimal bits are needed to get a reasonable model. And at <24,10> precision (14 decimal bits), the models are almost indistinguishable. The ROC curves show an AUC for the Keras model of about 0.92.

Conclusion

Our work is the beginning of a long and exciting project to create fast and accurate models using machine learning techniques for the next generation of particle colliders. Particle detectors handle terabytes of data in seconds, making experimental particle physics a prime example of a field that can be substantially improved through advancements in machine learning.

For next steps, we want to start quantization aware training, which will allow for the best model performance without compromising the latency and resource restrictions of the FPGA. We also want to explore how our model will behave as we adjust the training data to have different sampling frequencies, noise levels, and numbers of pulses. We also hope to soon synthesize our hls4ml model onto an FPGA chip and analyze what the actual resource usage and latency would be.

For the next generation of higher performance colliders, including the Future Circular Collider (FCC), detectors will hopefully be much more equipped to handle large amounts of information accurately and efficiently.

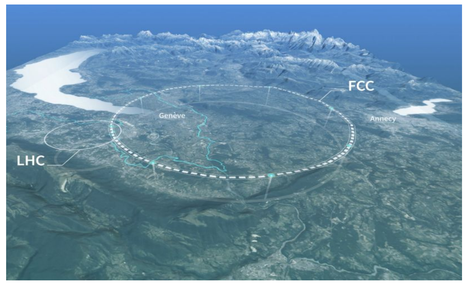

Figure 8: Location and design of the FCC. It will be 100 km in circumference compared to the 27 km LHC. Construction is set to begin in the 2030s.

Acknowledgements

This material is supported by NSF Award #2244348, NSF REU in Physics and AI. I am thankful for the support and encouragement of Professor Matteo Cremonesi and the patience and guidance of Dr. Peter Meiring. Thank you to Spencer Allen who will continue my work during the academic year. I would also like to thank the organizers of the Summer Scholars Program for an enjoyable, well rounded summer experience.