Academic Outcome Report by the Recipient of the Osaka University FrontierLab Summer Program
大阪大学超短期留学生受入れ奨学金受給者学修成果報告書

Name（氏名）： Jordan Xiao

Using Deep Neural Networks to Improve Particle Tracking

Student: Jordan Xiao
Home Institution: University of California, Berkeley
Supervisors: Yoshitaka Kuno, Ting Sam Wong
Kuno Laboratory, Graduate School of Science, Osaka University

Experiment Background:

The Standard Model of particle physics classifies all known elementary particles and three out of four known fundamental forces. Its self-consistency and experimental predictive power have earned it a rightful reputation as one of the crowning achievements of modern science. Nonetheless, the model remains incomplete, and the current paradigm within the field is to search for evidence of phenomena not accounted for within the SM.

The Coherent Muon to Electron Transition (COMET) experiment seeks physics beyond the SM in the form of charged lepton flavor violation (CLFV). In the process of interest, muons captured by atomic nuclei decay into electrons without the creation of neutrinos:

$$\mu^{-} + N(A,Z) \;\rightarrow\; e^{-} + N(A,Z)\,. \tag{1}$$

Though this process is not technically forbidden by the SM, the predicted branching ratio of $O(10^{-54})$ is extremely small and requires sensitivities well beyond the current limit of $O(10^{-13})$, set by the SINDRUM-II experiment in 2006. The 40-order-of-magnitude span between the current limit and SM prediction leaves plenty of room for new theories and experiments to probe. Furthermore, the proposed process has a low event background, meaning the theoretically predicted momentum signature of the electron on the right-hand side of Equation 1 is unique (at 105 MeV/c) and unlikely to be mistaken with different phenomena. Thus, CLFV proves to be an ideal candidate to explore for new physics.

The COMET experiment is split into two phases. Phase I seeks to achieve a 100-fold improvement in sensitivity on the current limit, and Phase II's eventual target is a 10000-fold improvement. A diagram of the experimental apparatus for Phase I is shown below.

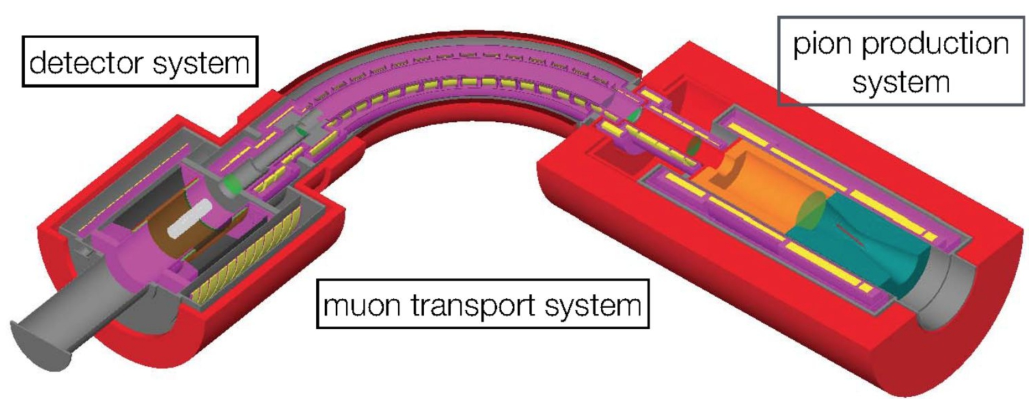

Figure 1: COMET Phase I experiment diagram.

At the right of Figure 1, pions generated from a proton beam decay into muons as they are transported to the detector system (left of Figure 1), where the muons are captured by atoms in an aluminum stopping target. Any muons that decay into electrons by the process of interest are ejected from an unstable orbit and follow a helical trajectory due to the application of a longitudinal magnetic field (Fig. 2a). Phase I of COMET uses a cylindrical ionization drift chamber (Fig. 2b) comprising thousands of sense wires to detect the arcing electron tracks with high spatial resolution in the perpendicular plane. Data is collected in the form of a hit map showing which wires within the CDC are triggered (Fig. 2c), after which an extensive track-fitting software pipeline is applied to determine the parameters of the helix.

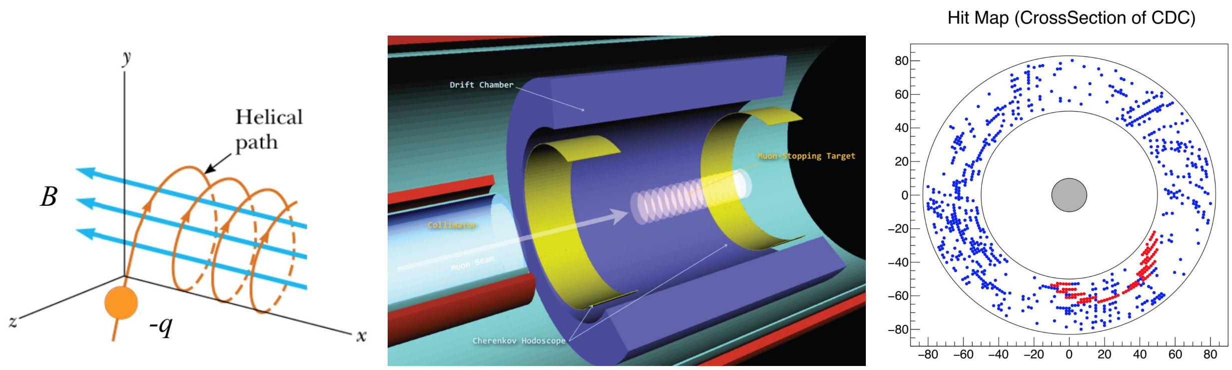

Figure 2: a) Charged particles follow a helical trajectory in a magnetic field. b) 3D model of the CDC, with beam entering at left. c) Simulation of data collected by CDC; blue represents background hits, and red is the desired signal.

Project Motivation:

Since particle energy determines the radius of curvature, precision track-fitting is critical to identifying signal electrons in the correct momentum range as determined by theoretical calculations (105 MeV/c). The final-stage fitting for COMET is performed using GENFIT, a well-established and commonly used software framework for track reconstruction. However, the Kalman filter algorithm used in GENFIT is highly sensitive to the initial guess provided by users; poor starting parameters can lead to nonsensical outputs. In order to optimize this initial guess, we apply machine learning to perform pre-fitting of parameters, where the input is a filtered hit map (background hits removed) and the desired output is a set of nine parameters describing the helical trajectory of a signal electron.

Among machine learning models, deep neural networks show particular promise in this task for their flexibility at processing image-like data relative to traditional pre-fitting algorithms. The primary disadvantage is the black box nature of a trained neural network, but a full discussion of this question is deferred until a fully trained model exists to be compared with the traditional algorithms. Thus, the goal of this project is to implement a robust software framework that allows efficient training and evaluation of deep neural networks tailored to this stage of the analysis pipeline.

Methods:

The current COMET deep neural network (COMET-DNN) is implemented in Python using the open-source Tensorflow machine learning framework, and based on a premade neural network designed for the CIFAR-10 image dataset. Rather than a classification task like CIFAR-10, our track-fitting is a multi-label regression problem, so the code has been extensively adapted accordingly. We use large quantities of Monte Carlo simulation data from Geant4 (another external software platform) in a standard supervised learning process to train the neural network, and design evaluation functions to gauge the predictive power of our models.

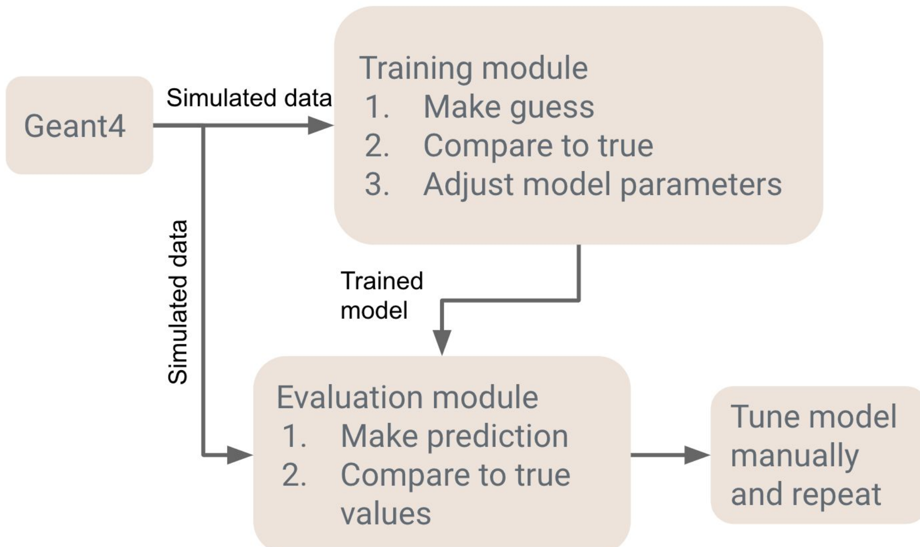

Figure 3: Developmental workflow for COMET-DNN.

In the final step shown in Fig. 3, manual tuning entails adjusting hyperparameters such as network layer types, number of nodes in each layer, learning rate, training batch size, optimization algorithm, loss function, etc. Training is performed on the Taurus supercomputing cluster at the Technische Universität Dresden.

Results and Conclusion:

As of August 2018, COMET-DNN includes working training and evaluation modules that allow for highly tunable model development as well as reliable post-training assessment. A brief summary of changes over the course of the project period (June to August 2018):

- ❖ Training for arbitrary sets of labels now enabled. Previously, modules could only train for labels in a hard-coded order.
- ❖ Reliable model saving and restoration. Previously, trained models were not properly reloaded when passed to the evaluation module.
- ❖ Streamlined plotting of diagnostic histograms and graphs for efficient model assessment.
- ❖ Deprecated functions replaced, and documentation updated in all relevant modules.

Unfortunately, all trained models to date suffer from similar issues when used for inference on new data. In particular, predictions tend to be highly Gaussian when true distributions are skewed (e.g. momenta) or uniform (e.g. vertex coordinates).

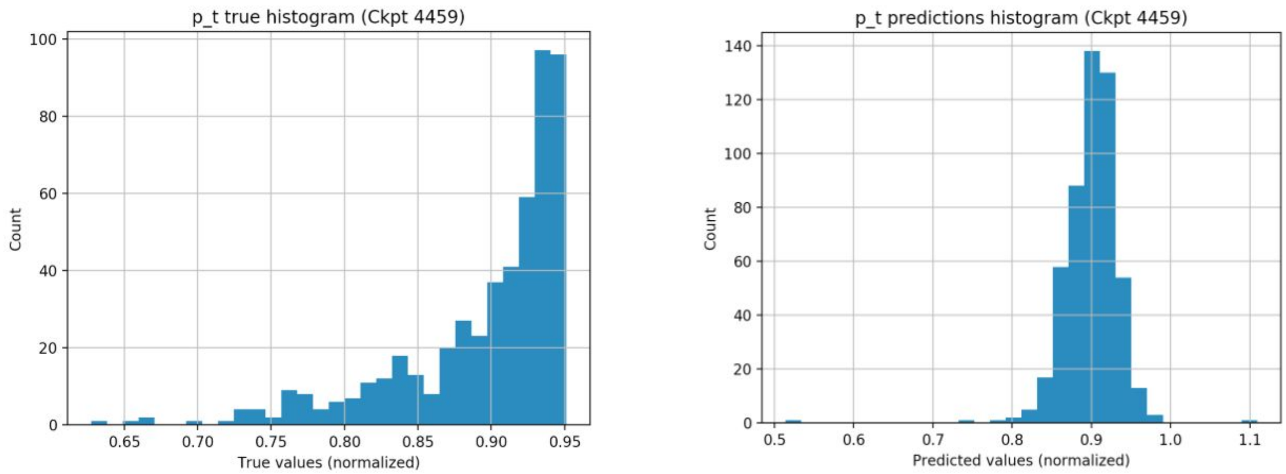

Figure 4: a) Example distribution of true tangential momentum values, clearly left-skewed. b) Corresponding predicted distribution from a trained deep neural network, clearly Gaussian in shape.

Among other attempted solutions, sweeping hyperparameters has not improved model performance, so the underlying causes for these issues are still under exploration. While the project is still in relatively early stages, with more flexible training and robust evaluation now implemented, the groundwork has been laid for future model optimization and debugging on the path to a powerful new tool for the COMET data analysis pipeline.

References:

Yoshitaka Kuno, on behalf of the COMET Collaboration; A search for muon-to-electron conversion at J-PARC: the COMET experiment, *Progress of Theoretical and Experimental Physics*, Volume 2013, Issue 2, 1 February 2013, 022C01, https://doi.org/10.1093/ptep/pts089

Images courtesy of Ting Sam Wong and Chen Wu.