Particle Track Reconstruction with Deep Learning

Steven Farrell, Paolo Calafiura, Mayur Mudigonda, Prabhat
Lawrence Berkeley National Laboratory
{SFarrell,PCalafiura,Mudigonda,Prabhat}@lbl.gov

Dustin Anderson, Josh Bendavid, Maria Spiropoulou, Jean-Roch Vlimant, Stephan Zheng
California Institute of Technology
{dustinanderson111,joshbendavid,maria.spiropulu, jeanroch.vlimant,st.t.zheng}@gmail.com

Giuseppe Cerati, Lindsey Gray, Keshav Kapoor, Jim Kowalkowski, Panagiotis Spentzouris, Aristeidis Tsaris, Daniel Zurawski
Fermi National Accelerator Laboratory
{cerati,lagray,kkapoor,jbk,spentz, atsaris,zurawski}@fnal.gov

Abstract

Particle track reconstruction is a challenging pattern recognition task in high energy physics experiments such as those at the Large Hadron Collider. Traditional algorithmic solutions rely on hand-engineered features and metrics, do not parallelize easily, and scale poorly with detector occupancy. In this paper we present our work to identify and evaluate solutions based on modern machine learning techniques such as deep neural networks. Models have been developed which draw inspiration from computer-vision tasks to identify tracks and estimate track trajectory parameters in image-like detector data. Additional models have been developed which can operate on a continuous distribution of spacepoint measurements to construct tracks in a structured way. We will evaluate these ideas on toy detector data and semi-realistic simulated tracking data and discuss their strengths and limitations for application in tracking applications.

1 Introduction

In high energy physics experiments such as ATLAS [1] and CMS [2] at the Large Hadron Collider [3] (LHC), a challenging but essential aspect of data processing is the measurement of charged particle trajectories in tracking detectors. Highly granular silicon-based sensors collect tens of thousands of position measurements ("spacepoints") from thousands of particles in every proton-proton beam collision event, as illustrated in figure 1. Tracking algorithms partition these spacepoints into disjoint groups ("tracks") and fit parametrized trajectories to extract particle kinematics and locations of production vertices. These results are combined with measurements from additional detector systems to construct a complete physical model of the particles in an event. Large datasets of these reconstructed events are then analyzed to test the fundamental laws of nature.

Traditional tracking algorithms have been used with great success in the experiments thus far but suffer from some limitations that motivate new ways of thinking. The algorithms are inherently serial, rely on linear dynamics models, and scale poorly with detector occupancy. In fact, in the expected conditions of data taking in 2025 (the so-called High Luminosity LHC), tracking algorithm code

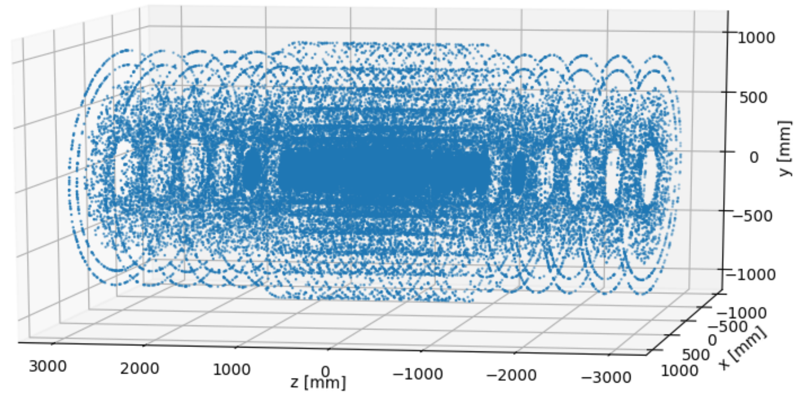

Figure 1: Distribution of particle spacepoints in a particle collision event in a generic simulated HL-LHC tracking detector.

will consume a disproportionate amount of offline computing resources that cannot be supplied with expected computing budgets.

Machine learning methods such as deep neural networks have some promising characteristics that could prove effective for high energy physics tracking. Neural networks are known to be very good at finding patterns and modeling non-linear dependencies in data. They also involve highly regular computation that can run effectively on parallel architectures such as GPUs. While there exists some literature from the 1980s-1990s studying neural network algorithms for tracking [4–6], modern techniques in deep learning have yet to be studied extensively in this regime.

We have explored two categories of approaches for machine learning solutions, image-based and point-based models. In the image-based models, inspiration is drawn from computer vision techniques such as semantic segmentation and image captioning, whereby we treat the detector data as an image and apply convolutional and recurrent neural networks to detect tracks. In point-based models, we use continuously distributed spacepoint measurements and structure them in a list or tree for learning how to group them into track candidates.

2 Image-based approaches

We investigated the applicability of sequence-based and image-based models for the problem of track-building on toy detector data in which spacepoints are binned in a 2D or 3D histogram [7]. An LSTM model was developed which reads the layers of the detector as a sequence of pixel arrays and emits a prediction for the correct location of a target track amidst background spacepoints. A similar model using convolutions was developed which processes the entire detector image and classifies pixels belonging to the target track. Several variations on these models were studied with toy data and semi-realistic simulated track data under varying numbers of background tracks, with the toy data results summarized in figure 2. The models showed good performance on toy datasets and promising results on semi-realistic data that suggest neural networks are effective at recognizing particle track patterns in detector data.

3 Point-based approaches

The discrete models explored thus far map nicely onto well-studied problems in computer vision and sequence modeling. However, they face difficulties when scaling up to the realistic complexity of LHC data, suffering from high dimensionality and sparsity. This motivates development of models that properly utilize the structure of the data as points localized on detector layers. These points can be structured as sequences, trees, or graphs for neural networks to learn representations on.

The first point-based approach utilizes a recurrent neural network as an iterative filter similar to a Kalman Filter. The model is trained to read a sequence of points and predict the position of the point

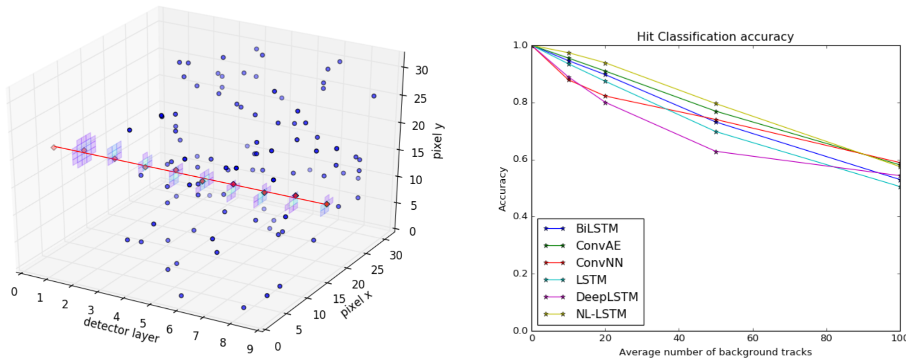

Figure 2: On the left is an example 3D toy data input with a target track shown as the red connected points and an LSTM model prediction shown as colored surfaces. On the right is the spacepoint classification accuracy of a variety of LSTM and convolutional models shown for varying numbers of background tracks [7].

on the next detector layer. It can be used to build tracks by selecting the closest spacepoint to the prediction or by implementing a combinatorial tree-search algorithm which considers plausible points at every layer and searches until a complete track is found. The architecture used here is an LSTM plus fully connected linear layer. An example trajectory and predictions from an ACTS [8] simulated dataset are shown in figure 3 and the prediction errors are shown in figure 4. The RNN model is observed to produce reasonable trajectory predictions after a sequence two to three spacepoint measurements, though some asymmetry and tails in the error distributions show there may be some limitations in the modeling that require further study.

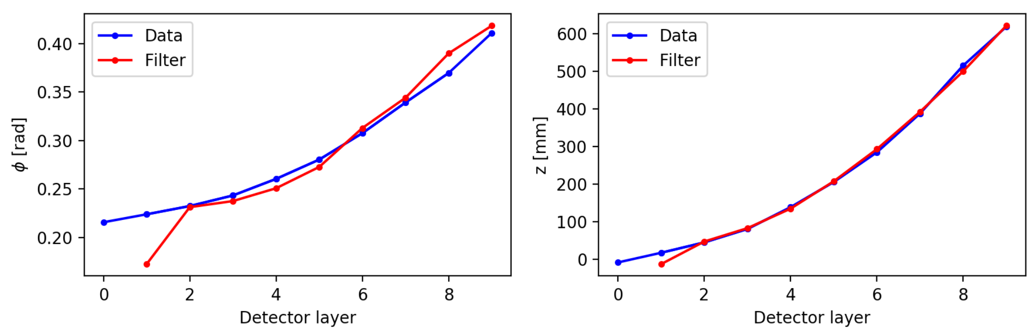

Figure 3: Example track measurements and RNN filter predictions in the ϕ (azimuth) and z coordinates as a function of detector layer. The RNN filter model is shown to make reasonable predictions after a sequence of about two spacepoints.

Another point-based method arranges all of the spacepoints of the detector in a sequence sorted according to the cylindrical coordinates and feeds them into a recurrent network model which outputs for every spacepoint a probability assignment vector of track classes. The target track classes are similarly sorted according to coordinates. The architecture, shown in figure 5, has three layers of bi-directional GRU units followed by a fully-connected layer and softmax activations to normalize the probability predictions for every spacepoint. A toy data sample was used based on a 3D cylindrical version of the TrackML RAMP challenge dataset [9] with random noise hits added. Figure 6 shows the assignment accuracy of the model and its dependence on the detector occupancy with 3D toy cylindrical data. While the model performs well with low occupancy, there is seemingly room for improvement as the accuracy degrades with increasing multiplicity. Still, this study demonstrates that such a model can learn to arrange spacepoints into appropriately sorted candidates under particular conditions. If such a model does not scale to a full event occupancy it may still be powerful in smaller sections of a detector.

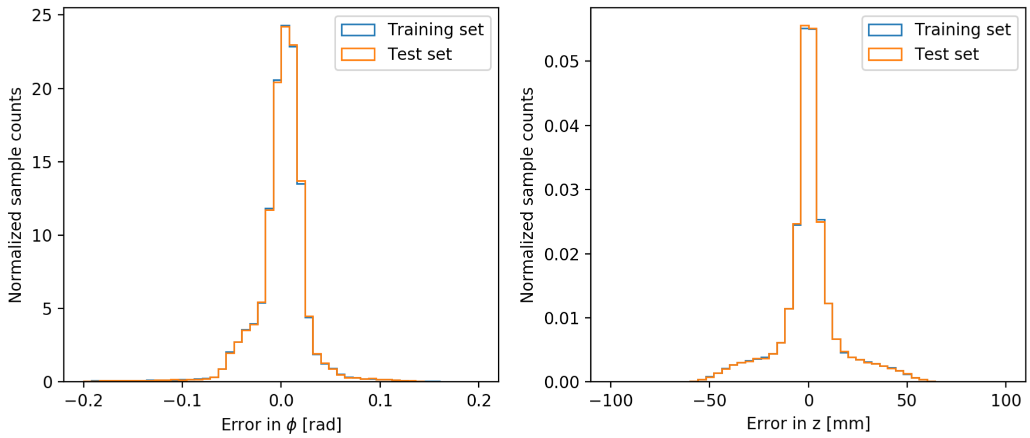

Figure 4: Error in the RNN filter predictions on both training and test datasets. The histograms are normalized to unit area. There is excellent agreement between training and test samples, though some asymmetries and long non-Gaussian tails are observed.

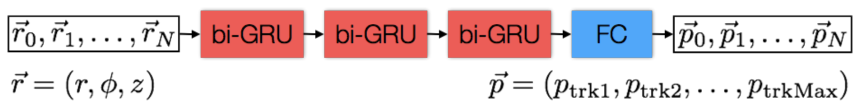

Figure 5: The multi-track spacepoint assignment model is a recurrent neural network that takes as input the full set of spacepoint measurements in the detector and outputs a track probability assignment matrix.

Conclusion

A variety of deep learning approaches for the problem of particle track reconstruction at high energy physics experiments have been studied. Both image-based and point-based approaches show promise in this problem. The point-based approaches seem to be the most suitable for scaling to full HL-LHC data conditions because they exploit the structure of the data while avoiding the sparsity and dimensionality of the image-based approaches.

Ongoing and future work in this area will involve careful evaluation of these methods and comparison with traditional solutions, as well as further explorations into new types of models that exploit the structure of the data. For the track filter model, it may be interesting to train a model to produce distributions via parameterized probability density functions (e.g. Gaussian) for its spacepoint predictions rather than just central values. Such an approach would allow a model to express uncertainty and potentially produce more useful predictions. Next, the filter model needs to be incorporated into a tree-search algorithm in order to demonstrate its capability to find tracks in full collision events. For the multi-track spacepoint assignment model, the accuracy scaling with occupancy should be further studied. Improvements may come from incorporating physics constraints into its track assignment classes or by discovering more efficient and expressive ways to embed the hit data. Graph neural networks [10] may be one powerful approach for modeling local relations between spacepoints.

References

[1] ATLAS Collaboration (2008) The ATLAS Experiment at the CERN Large Hadron Collider. JINST 3, S08003.

[2] CMS Collaboration (2008) The CMS Experiment at the CERN LHC. JINST 3, S08004.

[3] Evans, L., Bryant, P. (2008) LHC Machine. JINST 3, S08001.

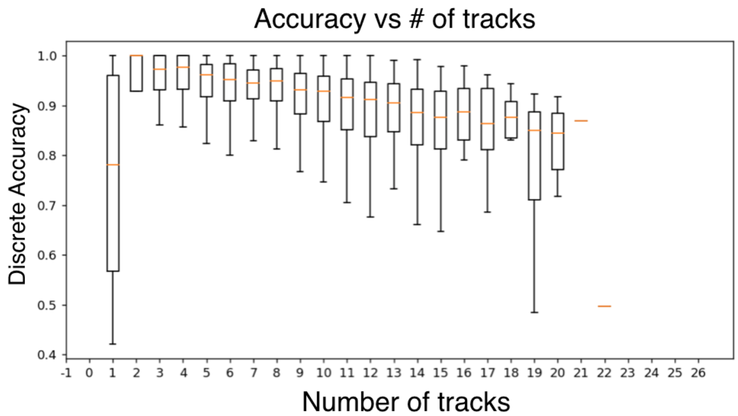

Figure 6: Accuracies of the spacepoint sequence hit assignment model as the number of tracks is increased. For 21 and 22 tracks no box interval is shown is because there is only a single sample.

[4] Denby, B. (1988) Neural networks and cellular automata in experimental high energy physics, *Computer Physics Communications* 49 (3) 429–448.

[5] Peterson, C. (1989) Track finding with neural networks, *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 279 (3) 537-545.

[6] Stimpfl-Abele, G., Garrido, L. (1991) Fast track finding with neural networks, *Computer Physics Communications* 64 (1) 46-56.

[7] Farrell, S., Anderson, D., Calafiura, P., Cerati, G., Gray, L., Kowalkowski, J., Mudigonda, M., Prabhat, Spentzouris, P., Spiropoulou, M., Tsaris, A., Vlimant, J.R., Zheng, S. (2017) The HEP.TrkX Project: deep neural networks for HL-LHC online and offline tracking. *EPJ Web Conf.* 150 00003.

[8] A Common Tracking Software Project, http://acts.web.cern.ch/ACTS/index.php

[9] Amrouche, S., *et al.* (2017) Track reconstruction at LHC as a collaborative data challenge use case with RAMP, *EPJ Web Conf.* 150 00015.

[10] Bronstein M., Bruna J., LeCun Y., Szlam A., Vandergheynst P. (2016) Geometric deep learning: going beyond Euclidean data, arxiv:1611.08097 [cs.CV].