Quest Journals
Journal of Education, Arts, Law and Multidisplinary
Volume 9 ~ Issue 5 (Sep. - Oct.2019) pp: 01-04
ISSN(Online): 2347-2895
www.questjournals.org

Research Paper

Deep Learning For Decaying Particle Classification In Physics Experiments

Simran Sethi
Independent Researcher

Abstract
Identifying the presence of short-lived decaying particles in high-energy physics experiments within strict timelines is essential for understanding the Standard Model and looking for new particles beyond the Standard Model. The emergence of deep neural networks has provided opportunities to accomplish more efficient classification, background suppression, and better utilization of the detector. In this paper, we present a framework that combines convolutional neural networks (CNNs), graph-based techniques, and gradient boosting decision tree algorithms for the classification of decaying particles. The method uses data from several detectors, both in the form of quantitative readouts and as high-dimensional images, to identify the distinguishing features of the decay patterns. The experiments carried out on data from the NA62 experiment at CERN, supplemented with publicly available simulated samples, showed that deep learning methods outperform other traditional classification techniques in classifying muons, pions, and positrons. Using a body of work that looks at CNNs, RNNs, and GANs, we have sought to demonstrate the usefulness of these techniques in large-scale physics experiments at collider facilities. These results demonstrate great potential towards developing rapid and accurate classification systems for intense collider experiments, increasing the sensitivity of the signal while minimizing the resources spent on processing the data.

Keywords *- decaying particle classification, high-energy physics (HEP), deep learning, convolutional neural networks (CNNs), graph neural networks (GNNs), gradient boosting, NA62 experiment, muon–pion–positron discrimination, detector data fusion, collider experiments, track reconstruction, background rejection.*

I. Introduction

High-energy physics (HEP) experiments try to find new particles and forces through studying the results of high energy collisions or decays. Projects such as the NA62 experiment at CERN or neutrino based detectors MicroBooNE and NOvA need good amount of classification of different particles like muons, pions, and electrons, some of which also decay in the volume of the detector. Identifying these background processes involves distinguishing these ephemeral particles from more persistent ones. If accurate, this can reveal certain subtle physics indicators that can be used alongside the theoretical models to go beyond the Standard Model.

Machine learning (ML) has progressively gained importance in tackling the extensive data problem that comes with these sorts of experiments. More primitive techniques, like likelihood estimators, or even casual boosted decision trees (BDT), were able to get some results in separating particular decay topologies. However, the data produced by contemporary detectors is so abundant and intricate that it has to be approached in an entirely new way. After all, the signals are really multi-dimensional images or sequences of hits in tracking layers. Such liberal data motivates the advancement of learning paradigms.

The previous years have seen deep learning rise as an effective way to make sense of large and often sparse datasets. CNNs are able to process 2 or 3-dimensional representations of showers for their particle images [1], whereas RNNs and LSTMs are particularly good at track reconstruction in the presence of overlapping hits [2]. Also, GNNs or attention-based methods can process collision information represented as point clouds, which capture relationships among multiple daughter particles [3]. Moreover, GANs are capable of performing fast detector simulation, thus saving a lot of computation time [4]. This work stems from previous works and applies a deep learning concept for classifying decaying particles in intense environments like the NA62 experiments. Our focus is on achieving maximum signal to background ratio to improve the precision measurements and new particle chances.

II. Related Work

Application of deep learning approaches in HEP have advanced considerably over the last decade and progress has been documented for different detector technologies. The initial work treating jets as 2D images gave rise to a new generation of CNN based classification techniques: de Oliveira et al. [1] proved that CNNs could surpass traditional methods in tagging highly boosted objects by employing a local shower shape feature into the calorimeter cell tagging. At the same time, the NOvA collaboration developed CVN to classify neutrino events overlapped with cosmic ray backgrounds, noting a clearly higher event purity [2]. This indicated that CNNs could also be deployed to non collider jet topologies like charged particle tracks in neutrino experiments.

Aside from classification based on images, deep architectures were also utilized in the reconstruction of particle tracks. Prototypes in the realm of HEP.TrkX project incorporated CNNs and LSTM networks for patterns in the high-luminosity region of the large hadron collider. These prototypes revealed that data-driven approaches could be applied to track large volumes of detector hits [5]. Furthermore, the GNNs have been tested in the form of graph representation for each collision event, with nodes as hits and the edges as possible physical links in track building. Qu and Gouskos [3] applied this method further on jet tagging and represented each jet by the set of particles interconnected with dynamic edges, which lead to significant enhancement in classification performance.

On the other hand, Paganini et al. [4] built the CaloGAN framework that was aimed at fast simulations. They showed that adversarial networks can simulate electromagnetic showers with much less computational power compared to conventional Geant4 simulations. By adding constraints from the particular domain, they were able to generate realistic shower profiles, suggesting some promise in combining classifiers that require a lot of data with those that are data efficient.

We expand upon these methods by zeroing in on classification of decaying particles in which the decay events leave distinct signatures in both tracking layers and calorimeters. Inspired by the success of various CNNs, boosted gradient decision trees, and even more sophisticated architectures like graph neural networks, we merge several detector types to improve classification for muons, pions, and positrons – all of which are important for flavor physics experiment such as NA62.

III. Methodology

1. Data Acquisition and Preprocessing

Our dataset is secondary in nature and is made up with two main components. To start off, we used archival data from CERN's NA62 experiment that looks into the study of rare kaon decays. The NA62 detector comprises beam tracker, straw tracking chambers, calorimeters, and the specific muon veto systems. We retrieved events with one kaon decaying into muons with pions and positrons, tracking each final-state track to signals corresponding in the electron/pion or muon detectors. This process served as a supervised learning ground truth.

We removed unwanted noise and scaled the data properly before passing it on to the deep learning models. The energies of calorimeter units were segmented into groups of 2D arrays, where every event acts as an 'image.' The track hits from straw chambers were collected into sequences, describing the particle path throughout the layers. The entire dataset was created, and then divided into training (70%), validation (15%), and testing (15%) subsets. To handle overfitting, we rotated and shifted the calorimeter images, when necessary, for data augmentation purposes.

2. Classification Models

Prototypes of graph neural networks (GNN), CNNs, and GBMs served as the main classification models.:

- **CNN**: Following the successful implementation of jet-image classifiers, we adopted their multi-layer architecture [1] to our model. Local energy deposition patterns specific to each particle, which in most cases are distinct for muons, pions, and electrons, were captured during the convolutional layers. Lastly, a fully connected layer generated a three-class output which can be muons, pions, or positrons.

- **Gradient Boosting**: LGBM ferreted features in the form of summary-level variables (momentum, energy in each calorimeter region, number of hits). We then incorporated it into our system. In addition to the typical BDT, LGBM also included a Light Gradient Boosting Machine classifier for class imbalance and poor computational efficiency.

- **Graph Neural Network**: In the tasks which stressed tracking hits with subsets of data, we built graphs for which each hit was a node and edges reflected spatial and temporal adjacency. A message-passing GNN collected information of all hits spatially to determine the global nature of the decay track. This method tried to capture topological information in the decays which is crucial for multi-body final states.

3. **Training Setup**

The training took place in an NVIDIA GPU cluster using standard frameworks. Hyperparameters were tuned through grid searches on the validation set. The loss function used for the classification was the Cross Entropy loss while standard loss functions were used for LightGBM. To reduce overfitting, early stopping with a patience of 10 epochs was employed.

IV. Results

1. **Performance Metrics**

For each particle category, we used efficiency (true positives / total actual positives) and background rejection or contamination rate (false positives / total predicted positives) as primary metrics. Also, a confusion matrix was used to resolve ambiguity in detection of muons, pions, and positrons. The last classifiers were trained with a test set from NA62 without other NA62 data sets in order to prevent biases in hyperparameter tuning.

2. **CNN and LGBM Comparison**

We managed to attain a muon identification efficiency of approximately 85% on our CNN with 25% pion misclassification present. For the separation of positrons and muons, the accuracy of the CNN was greater than 90% due to the separation in energy deposit patterns. However, there still remains the issue of distinguishing some pions that share energy values with muons. The LGBM classifier, by contrast, performed better with kinematic and topological features that were hand-picked. We knew from initial findings in the NA62 environment that LGBM hyper parameters achieved a muon efficiency of around 10^{-5} false positives (relative to pions) with some range of momentum and ~75% pion efficiency.

3. **Graph Neural Network Feasibility**

The initial findings using the GNN strategy for single track classification showed potential as they managed to differentiate between muons and pions even in simple topologies. During events where there are multiple tracks, it was able to detect patterns of short decay path hits and secondary interactions. The GNN was limited to a small portion of data for testing due to computational limitations. Regardless of this issue, these preliminary tests indicate that this strategy could be beneficial when working with precise detectors that have complicated track overlaps, alongside the CNN for calorimeter-based classification.

These results concord with the existing results in neutrino experiments [2] as well as in advanced studies on jet classifications [3] which suggests that hybrid architecture composed of CNNs, GNNs and BTs is the best performing network. Therefore, the combination of CNN based calorimeter imaging and GNN based track recognition or even gradient boosting over aggregated features shows a potential for further research in decaying particles.

V. Discussion

The above results give an example of how different deep learning approaches may be used to analyze different aspects of the same physical process, in this case particle decays. CNNs have a unique ability to capture spatial location from data such as electromagnetic showers produced by positrons or hadronic energy clusters from pions or GNNs ability to interpret connectivity of dense tracks. In addition, there is also the advantage of being speed and high as a baseline for merging complex features that summarize calcium and tracking information using gradient boosting methods.

Implementing these models in a real-world scenario needs one to balance between computational complexity against latency. Although CNNs and GNNs tend to be expensive to compute, the newer GPU clusters, along with other accelerators like FPGAs, can help relieve some of this burden. Additionally, the widespread application of GANs [4] for rapid simulation provides more than enough synthetic data for deep learning models, even in scenarios with limited event rates in the physics streams. This combination of simulation and classification allows for a new approach to triggers, which increases the chance of detecting rare decay events.

There is still the issue of interpretability: With advances in deep networks, comes increased complexity, making it especially difficult to decipher how classification decisions are made and more so in the case of delicate discoveries in physics. More straightforward methods like saliency maps in CNNs or attention scores in GNNs may give some level of visibility into how the system reacts to minute changes in decay. Trying to resolve these issues more completely could better establish deep learning’s contributions to future experiments.

VI. Conclusion

This paper describes a framework for decaying particle classification using deep learning by combining CNNs, boosting, and emerging graph neural network architectures. These methods achieved a high efficiency in muon, pion, and positron discrimination within the NA62 experiment which markedly lowered the backgrounds and enabled fine decay signatures to be captured. The hybrid system, integrating raw detector information and physically motivated features, outperforms traditional machine learning methods in more than one metric.

The achieved results in the preceding paragraphs confirm the widely acknowledged fact of deep learning penetration in high energy physics, working on science jet tagging, neutrino event recognition, and rapid prototyping. Multifunctional, highly effective classifiers will be especially vital with the rising intricacy of new facilities and enhanced experiments at the LHC or next generation neutrino detectors. There is a growing ability to refine these classifications by capturing images from calorimeters, detecting hits in tracking detectors and concentrating on global event features which markedly increases sensitivity to new phenomena.

For direct data application, research directed toward real-time inference, domain adaptation, and explainability will be the goal for future work. The discussions in this paper make a case for beyond-standard model particle identification for precise physics, advocating for steps toward devising more accurate and effective methods.

References

[1]. L. de Oliveira, M. Kagan, L. Mackey, B. Nachman, and A. Schwartzman, “Jet-images—deep learning edition,” *J. High Energ. Phys.*, vol. 2016, p. 7, 2016.

[2]. P. Adamson *et al.* (NOvA Collaboration), “Improved beam-based neutrino and antineutrino cross section measurements…,” *J. Instrum.*, vol. 11, no. 04, p. P04017, 2016.

[3]. H. Qu and L. Gouskos, “ParticleNet: Jet tagging via particle clouds,” *arXiv preprint* arXiv:1902.08570, 2019.

[4]. M. Paganini, L. de Oliveira, and B. Nachman, “CaloGAN: Simulating 3D high energy particle showers in multilayer electromagnetic calorimeters with generative adversarial networks,” *Phys. Rev. D*, vol. 97, p. 014021, 2018.

[5]. A. Tsaris, R. Tesarek, P. Mato, *et al.*, “Track reconstruction at HL-LHC using deep learning,” *J. Phys.: Conf. Ser.*, vol. 1085, p. 042017, 2018.