GA
TN
MI
BW
ML
CI
BI

MACHINE LEARNING FOR HIGH-ENERGY PHYSICS: PARTICLE IDENTIFICATION AND REGRESSION USING DEEP NEURAL NETWORKS

Allocation: Illinois/125 Knh
PI: Benjamin Hooberman[1]
Co-PI: Amir Farbin[2]
Collaborators: Ryan Reece[3], Matt Zhang[1]

[1]University of Illinois at Urbana-Champaign
[2]University of Texas at Arlington
[3]University of California at Santa Cruz

EXECUTIVE SUMMARY

The Large Hadron Collider (LHC) at CERN, in Switzerland, is the world's most powerful particle accelerator. The LHC recreates the conditions of the universe one tenth of a nanosecond after the Big Bang by colliding together protons traveling at 0.99999997 times the speed of light 40 million times every second. Each proton–proton collision creates up to several hundred particles that pass through one of four detectors situated at the LHC interaction points. Reconstructing the collisions requires identifying these particles using their signatures in the detector. Recent advances in machine learning and artificial intelligence, which are also known as deep learning, have made it possible to apply learning networks to many kinds of problems. In particular, identifying particles from their energy deposition in calorimeter cells bears a strong resemblance to problems in machine vision, in which objects are reconstructed from intensity values in pixel arrays. We exploit deep learning techniques to identify and measure particles produced at colliders and find that they provide improvements in performance with respect to conventional methods.

RESEARCH CHALLENGE

As previously mentioned, the Large Hadron Collider (LHC) at CERN in Switzerland recreates the conditions of the universe a tenth of a second after the Big Bang by colliding together high-energy protons. In 2012, the Higgs boson was discovered in LHC data, completing the standard model of particle physics and leading to the Nobel Prize in Physics in 2013. This discovery transformed our understanding of the building blocks of matter and the fundamental forces by explaining the origin of the masses of subatomic particles. However, the standard model is not capable of resolving key open questions and thus cannot be the final theory of nature. In particular, it cannot explain the origin of dark matter, which comprises about five times as much total mass in the universe as visible matter but whose nature is not understood. Various beyond-the-standard-model scenarios, including supersymmetry and extra dimensions of spacetime, have been posited to resolve these problems. These scenarios generically predict the existence of exotic new particles, which may be produced at particle colliders such as the LHC. Searching for these particles to understand the nature of physics beyond the standard model is now the highest priority of the LHC physics program and the focus of this project.

Analyzing LHC data to search for physics beyond the standard model requires identifying and measuring the particles produced in proton–proton collisions. Particles produced in collisions traverse detectors, depositing their energy in calorimeters consisting of a granular array of detecting elements ("pixels"). The resulting "image" can be analyzed to distinguish among the six species of stable particles (electrons, photons, charged hadrons, neutral hadrons, and muons) and infer their energies. Electrons and photons are expected signatures of a wide variety of interesting new physics scenarios, but may be mimicked by charged and neutral hadrons, which are produced at rates that are higher by several orders of magnitude. Since each collision typically contains thousands of particles, discriminating signals from electrons and photons from hadronic backgrounds is complicated by the presence of additional overlapping particles. Identifying and measuring electrons and photons, especially those with low energy, is thus a major challenge of high-energy physics.

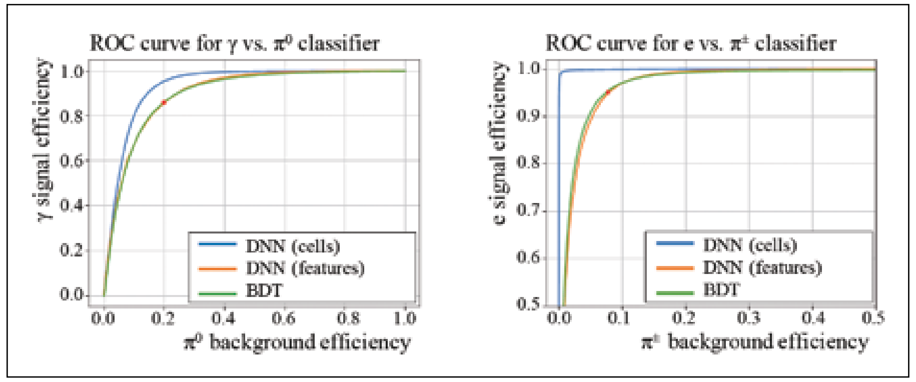

Figure 1: Signal vs. background efficiency Receiver Operating Characteristic (ROC) curves for (left) photon vs. neutral pion and (right) electron vs. charged pion discrimination.

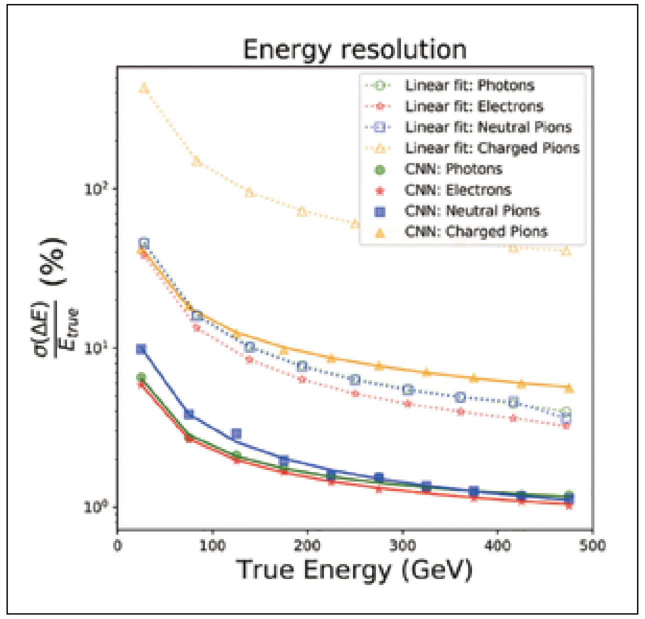

Figure 2: The relative energy resolution of the four particle types versus the true particle energy.

METHODS & CODES

Recent advances in machine learning and artificial intelligence, known as deep learning, have made it possible to apply learning networks to many kinds of problems. These techniques are driven by the emergence of large data sets, powerful Graphical Processing Unit (GPU) processors, and new techniques to train billion-neutron multi-layer artificial neural networks (NN). In computer vision, Deeply-connected Neural Networks (DNN) and Convolutional Neural Networks (CNN) have provided dramatic improvements in performance and speed with respect to conventional algorithms and require minimal engineering.

We employ DNNs and CNNs to distinguish among signals from electrons and photons and hadronic backgrounds and measure particle energies. We simulate samples of individual electron, photon, charged hadron, and neutral hadron images in a simple high-granularity calorimeter detector implemented with the Geant4 simulation toolkit. These images are used to train NNs using PyTorch, which distinguishes among electrons versus charged hadrons and photons versus neutral hadrons and measures the energies of the four particle species. To optimize the network architectures, we vary the NN "hyperparameters," including the number of NN layers ("depth"), number of neurons per layer ("width"), and the learning and dropout rates.

RESULTS & IMPACT

We have evaluated the performance of DNNs and CNNs trained on particle images, and compared the results to the current state-of-the-art algorithms widely used in particle physics. These algorithms employ NNs and Boosted Decision Trees (BDTs) to analyze a precomputed set of particle "features" such as the calorimeter shower depth and width. For both classification and energy measurement using regression, we find that the deep NNs provide significant improvements compared to the conventional methods. These results serve as a first step toward implementing deep learning for particle identification and measurement at the LHC.

WHY BLUE WATERS

Optimizing the network performance using hyperparameter scans requires retraining NNs hundreds or thousands of times, which is especially challenging for memory-intensive networks such as GoogLeNet or ResNet. The 4,228 GPU-enabled XK nodes with 25 TB of GPU accelerator memory available on Blue Waters enable training and optimization of neural nets beyond what has previously been achieved, allowing for detailed investigations of their behavior for both particle physics and general applications.

PUBLICATIONS & DATA SETS

Hooberman, B., et al., Calorimetry with Deep Learning: Particle Classification, Energy Regression, and Simulation for High-Energy Physics. *Proceedings of the Deep Learning for Physical Sciences Workshop at the 31st Conference on Neural Information Process* (Long Beach, Calif., December 8, 2017).