Calorimetry with Deep Learning: Particle Classification, Energy Regression, and Simulation for High-Energy Physics

Federico Carminati, Gulrukh Khattak, Maurizio Pierini
CERN

Amir Farbin
Univ. of Texas Arlington

Benjamin Hooberman, Wei Wei, and Matt Zhang
Univ. of Illinois at Urbana-Champaign

Vitória Barin Pacela
Univ. of Helsinki
California Institute of Technology

Sofia Vallecorsafac
Gangneung-Wonju National Univ.

Maria Spiropulu and Jean-Roch Vlimant
California Institute of Technology

Abstract

We present studies of the application of Deep Neural Networks and Convolutional Neural Networks for the classification, energy regression, and simulation of particles produced in high-energy particle collisions.We train cell-based Neural Nets that provide significant improvement in performance for particle classification and energy regression compared to feature-based Neural Nets and Boosted Decision Trees, and Generative Adversarial Networks that provide reasonable modeling of several but not all shower features.

1 Overview

In High Energy Physics (HEP) experiments, detectors serve as cameras that take pictures of the products of particle collisions. One of the key components of such detectors are calorimeters that image the energy depositions of the showers of secondary particles produced by high energy particles from these collisions interacting with dense detector material. The resulting patterns of depositions, which are characteristic of the particle type, are observed in "cells" analogous to voxels (possibly with irregular shapes) in three-dimensional (3D) images. Physicists, as a first step towards discovering or studying interesting phenomena or new particles, typically use features extracted by sophisticated reconstruction algorithms to identify the type and estimate the energy of particles in large samples of collision events. Machine Learning (ML) techniques are well suited for such tasks, and indeed ML has long played an essential role in HEP, including the 2012 Nobel Prize-winning discovery of the Higgs boson [1, 2] at the ATLAS [3] and CMS [4] experiments at the Large Hadron Collider (LHC).

In the next decade, the High Luminosity Large Hadron Collider (HL-LHC) upgrade of the current LHC will enhance the sensitivity to new physics by increasing the proton-proton collision rate. In addition, many next generation detectors, such as the sampling calorimeters proposed for the ILC [5] and CLIC [6], will improve the ability to identify and characterize particles produced in collisions using highly granular 3D arrays of pixels. These upgrades and future accelerators will lead to higher data volumes and a variety of technological challenges, e.g. real-time particle reconstruction and fast detector simulation. In addition, physics measurements typically require extremely detailed and precise simulation, relying on the well understood micro-physics governing the interaction of particles with matter coded into software packages, the most notable being Geant4 [7]. These simulations are

Workshop on Deep Learning for Physical Sciences (DLPS 2017), NIPS 2017, Long Beach, CA, USA.

generally very CPU intensive and in some cases, such as in the ATLAS experiment, currently require roughly half of the experiment's computing resources and are expected to be significantly more for the HL-LHC. These challenges require novel computational and algorithmic techniques, which has recently prompted efforts in HEP to apply modern ML to calorimetry [8, 9, 10, 11]. This prior work has focused on extracting properties of sprays of particles called "jets," while our studies focus instead on electrons and photons, which may be produced in the decays of a variety of interesting particles including Higgs bosons and hypothetical Supersymmetric particles.

This paper aims to demonstrate the improvement in performance for calorimetric classification, regression, and simulation tasks gained from applying Deep Neural Networks (DNN) and Convolutional Neural Networks (CNN) to calorimeter cells compared to the current ML techniques, such as Boosted Decision Trees (BDTs), applied to features. We train separate DNNs to discriminate electrons (e) from charged pions ($\pi^{\pm}$), discriminate photons (γ) from neutral pions (π^0), and estimate their energies. We observe significant improvement from using cell-based DNNs with respect to feature-based Neutral Networks (NNs) and BDTs. We also demonstrate an application of Generative Adversarial Networks (GANs) [12] to reproduce CPU-intensive simulations of particle showers.

The study is based on pseudo-data simulated with GEANT4 [7] in the proposed Linear Collider Detector (LCD) for the CLIC accelerator [13], which consists of a regular grid of 3D cells with cell sizes of 5.1 mm^3 and inner calorimeter radius of 1.5 m. Individual electron, photon, charged pion, and neutral pion particles are shot into the central part of the detector, in a region with regular sensor geometry and orthogonally to the calorimeter surface. The particle energy is set to 60 GeV for the classification task, and varied between 10 and 510 GeV for the regression and between 100 and 500 GeV for GAN studies. For each event, a 25x25x25 cell slice of the electromagnetic calorimeter (ECAL) and the corresponding 5x5x60 cell slice of the hadronic calorimeter (HCAL) are stored as two 3D arrays of deposited energy in each cell. The slices are centered around the barycenter of each ECAL energy deposit. All DNN models were implemented and trained using KERAS [14] with Tensorflow [15] or Theano [16] back-ends. We used AdaBoost [17] for the BDTs.

2 Classification: Particle Identification

Since the LHC collides protons, the collisions are dominated by strong interactions that produce quarks and gluons that turn into "jets" of mostly pions. These highly copious jets are the primary background to the identification of photons and leptons, which are much rarer. As a result, calorimeters are required to only mis-identify roughly 1 in 10000 jets as a photon or electron. Initial studies of DNN-based classification of the four particles types in our simulated samples yielded extremely good results, with Area Under Curve (AUC) of Receiver-Operator Curves (ROC) of near 1. In order to approximate the particle identification challenge at the LHC, we therefore select subsets of particles most likely to be misidentified as photons or electrons. We considered two interesting cases, each corresponding to a challenging classification task where we could compare BDTs trained on features to DNNs trained on cells or features. The first is charged pions ($\pi^{\pm}$), which are mis-identified as electrons ($e^{\pm}$) due to early showering in the ECAL, that we selected by requiring the ratio of total energies deposited in HCAL to ECAL be less than 0.025. The second is neutral pions (π^0), which nearly always decay to two photons (γ) but are commonly mis-indentified as single photons when the two photons overlap, that we selected by requiring an opening angle of less than 0.01 radians between the two photons (computed from ground truth), corresponding to approximately three cell widths at the inner calorimeter face.

Each dataset was split into 400,000 training and 100,000 testing events. Preliminary studies indicated that simple densely connected multi-layer DNNs were as effective as CNNs, so we chose to focus on simple DNNs due to the smaller number of hyperparameters. Since the ECAL and HCAL input tensors have different shapes, they were separately flattened and fed into the cell-based DNNs with identical architectures and merged before a softmax layer was used to compute categorical cross-entropy loss. Studies of data normalization strategies yielded best performance when the cell energies were passed through a non-linear transformation $f(E) = \frac{1}{2}\tanh(\log E + 1)$. We considered approximately 1000 models in a hyperparameter scan and found little dependence of the classifier performance on hyperparameters. For all DNNs (features, ECAL branch, and HCAL branch) we selected a simple DNN consisting of 4 hidden layers with 256 neurons each, ReLU activation, and dropout of 0.5, trained with Adam [18] optimizer with learning rate of 0.001. A similar

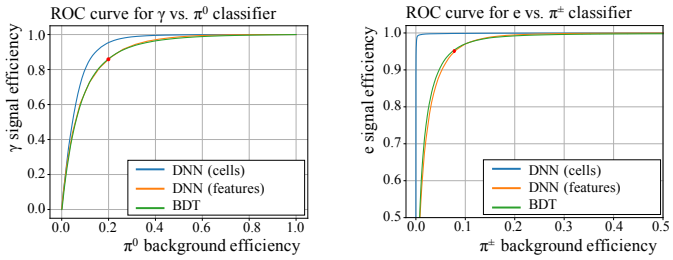

Figure 1: Signal vs. background efficiency ROC curves for the (left) γ vs. π^0 and (right) e vs. π classifier. The red dots mark the chosen BDT working point.

1000 model BDT hyperparameter scan yielded best performance with 400 estimators, maximum depth of 5, and learning rate of 0.5.

The features we computed are commonly used in calorimetry to characterize the particle shower shape and energy deposit. These features are: total energy deposited in ECAL, total number of hits in ECAL, the ratio of energy in ECAL first layer over energy in second layer, the ratio of energy in ECAL first layer over all ECAL energy, second through sixth moments in the detector local x, y, and z of ECAL energy deposits, all equivalent features for HCAL, ratio of HCAL to ECAL energy, and ratio of number of hits in HCAL to ECAL. In our studies, we found that the most powerful features are the second x and y moments that measure the lateral shower width.

	γ **vs.** π^0				e **vs.** π			
Model	**acc.**	**AUC**	$\Delta\epsilon_{\text{sig}}$	ΔR_{bkg}	**acc.**	**AUC**	$\Delta\epsilon_{\text{sig}}$	ΔR_{bkg}
BDT	83.1%	89.8%	-	-	93.8%	98.0%	-	-
DNN (features)	82.8%	90.2%	0.9%	0.95	93.6%	98.0%	-0.1%	0.95
DNN (cells)	87.2%	93.5%	9.4%	1.63	99.4%	99.9%	4.9%	151

Table 1: Performance parameters for BDT and DNN classifiers.

Figure 1 shows the ROC curves for the three classifiers and Table 1 quantifies the performance. The areas under curve (AUC) and accuracies (acc.) for the cell-based DNNs are significantly better than the feature-based DNNs and BDTs, which have similar performance. We also quantify the achievable improvements in signal and background efficiency from the DNNs with respect to the chosen "working point" on the BDT ROC curve indicated in Figure 1. For the γ vs. π^0 (e vs. $\pi^{\pm}$) classifier, the cell-based DNN may be used to either increase the signal efficiency by $\Delta\epsilon_{\text{sig}} = \epsilon_{\text{sig}}^{\text{DNN}} - \epsilon_{\text{sig}}^{\text{BDT}} = 9.4\%$ (4.9%) for fixed background efficiency, or decrease the background efficiency by a factor $\Delta R_{\text{bkg}} = \epsilon_{\text{bkg}}^{\text{BDT}}/\epsilon_{\text{bkg}}^{\text{DNN}} = 1.6$ (151) for fixed signal efficiency.

3 Regression: Energy Reconstruction

We trained a separate dedicated DNN to estimate particle energies from their calorimeter deposits. This DNN is composed of two CNNs for ECAL and HCAL, followed by a flattening and concatenation layer, with a final densely connected layer. The ECAL branch uses a 3-feature convolutional layer with a $4 \times 4 \times 4$ window and stride of 1 in each direction, followed by a $2 \times 2 \times 2$ max pooling layer with a stride of 2. The HCAL branch has a 10-feature layer with a $2 \times 2 \times 6$ window and stride of 1, followed by a $2 \times 2 \times 2$ max pooling layer with a stride of 2. All convolutional layers have ReLU activation. The output of both branches are linearized and merged, followed by a fully connected layer with 1000 neurons. The final neuron has a linear activation function and the mean-squared error (MSE) is used as the loss function. The data sample was split into 40,000 events for training, 10,000 events for validation, and 30,000 events for testing.

As a baseline measure of the energy, we use a simple bi-linear regression of the summed energy in ECAL and HCAL to the true energy. Figure 2 compares the energy dependence of the calorimeter resolution for each particle type and for both the neural net and the simple linear regression models. Table 2 quantifies the results by fitting this dependence to the expected form. We observe significantly

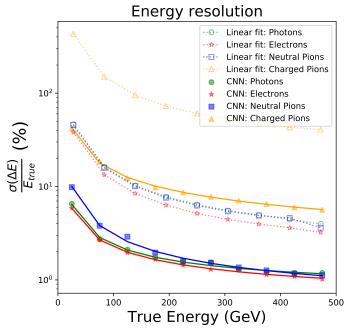

Figure 2: Energy resolution for photons, electrons, neutral and charged pions compared for the CNN vs. linear model.

Simple Linear Model

Particle Type	a	b	c
Photons	55.5	1.85	1245
Electrons	42.3	1.51	1037
Neutral pions	55.3	1.71	1222
Charged pions	442	25	11706

CNN Model

Particle Type	a	b	c
Photons	18.3	0.75	131
Electrons	18.7	0.574	111
Neutral pions	19.3	0.45	231
Charged pions	114	1.02	893

Table 2: Calorimeter resolution parameters from equation $\frac{\sigma(\Delta E)}{E_{\text{true}}} = \frac{a}{\sqrt{E_{\text{true}}}} \oplus b \oplus \frac{c}{E_{\text{true}}}$ for the resolution curves in Fig. 2.

better performance from the DNN as compared to the simple model, with resolution enhancement of a factor of 3.5–7 at low energies and 2–4 at high energies, for all four particle types.

4 Generative Model: Particle Simulation

We use the sample of ECAL 3D energy arrays to demonstrate the ability to simulate particles at given energies using GANs, as a proof of concept for a much larger plan to integrate a generic deep-learning tool for fast simulation into the GeantV detector simulation library [19].

Both the GAN generator and discriminator models consist of four 3D convolution layers with leaky ReLU activation functions. The number and sizes of filters were tuned to optimize the description of the transverse and longitudinal shower shapes. The discriminator models take the calorimeter image as input and produce two outputs: classification of the images as real or generated and regression of the energy, in the manner described in the previous section. The generator takes as input the desired particle energy and a latent noise vector initialized to a uniform probability distribution, and outputs a $25 \times 25 \times 25$ ECAL image. The results of GAN-simulated particles are shown in Fig 3, in comparison with the particles generated via GEANT4 [7]. The GAN provides reasonable modeling of the longitudinal shower width but further tuning is required to model the transverse shower width.

5 Conclusion and Future Work

This paper shows how deep learning techniques could outperform traditional and resource-consuming techniques in tasks typical of physics experiments at particle colliders, such as particle identification, energy measurement, and detector simulation. To continue this work, we will push forward particle classification and energy regression into new areas, using multi-particle events with overlapping

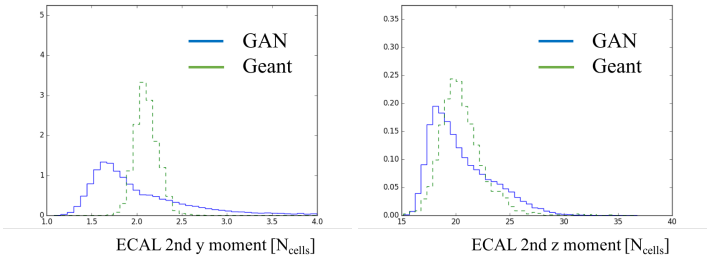

Figure 3: Comparison of (left) transverse shower width and (right) longitudinal shower width for GAN vs. Geant simulation of electrons with energies of 200-300 GeV.

showers, and including particles which impact the calorimeter at variable angles. We will also continue to refine GAN simulation, with the goal of producing a tool for fast simulation.

6 Acknowledgemnts

This project is partially supported by the United States Department of Energy, Office of High Energy Physics Research under Caltech Contract No. DE-SC0011925 and makes use of the Blue Waters Supercomputer at the National Center for Supercomputing Applications. JR is partially supported by the Office of High Energy Physics HEP-Computation. MS is grateful to Caltech and the Kavli Foundation for their support of undergraduate student research in cross-cutting areas of machine learning and domain sciences.

References

[1] Georges Aad et al. Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC. *Phys. Lett.*, B716:1–29, 2012.

[2] Serguei Chatrchyan et al. Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC. *Phys. Lett.*, B716:30–61, 2012.

[3] G. Aad et al. The ATLAS Experiment at the CERN Large Hadron Collider. *JINST*, 3:S08003, 2008.

[4] S. Chatrchyan et al. The CMS Experiment at the CERN LHC. *JINST*, 3:S08004, 2008.

[5] Ties Behnke, James E. Brau, Brian Foster, Juan Fuster, Mike Harrison, James McEwan Paterson, Michael Peskin, Marcel Stanitzki, Nicholas Walker, and Hitoshi Yamamoto. The International Linear Collider Technical Design Report - Volume 1: Executive Summary. 2013.

[6] L. Linssen, A. Miyamoto, M. Stanitzki, and H. Weerts. Physics and Detectors at CLIC: CLIC Conceptual Design Report. *ArXiv e-prints*, February 2012.

[7] S. Agostinelli et al. GEANT4: A Simulation toolkit. *Nucl. Instrum. Meth.*, A506:250–303, 2003.

[8] Luke de Oliveira, Michael Kagan, Lester Mackey, Benjamin Nachman, and Ariel Schwartzman. Jet-images — deep learning edition. *JHEP*, 07:069, 2016.

[9] Luke de Oliveira, Michela Paganini, and Benjamin Nachman. Learning Particle Physics by Example: Location-Aware Generative Adversarial Networks for Physics Synthesis. *Comput. Softw. Big Sci.*, 1(1):4, 2017.

[10] Michela Paganini, Luke de Oliveira, and Benjamin Nachman. CaloGAN: Simulating 3D High Energy Particle Showers in Multi-Layer Electromagnetic Calorimeters with Generative Adversarial Networks. 2017.

[11] Josh Cogan, Michael Kagan, Emanuel Strauss, and Ariel Schwartzman. Jet-Images: Computer Vision Inspired Techniques for Jet Tagging. *JHEP*, 02:118, 2015.

[12] I. J. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. Courville, and Y. Bengio. Generative Adversarial Networks. *ArXiv e-prints*, June 2014.

[13] P. Lebrun, L. Linssen, A. Lucaci-Timoce, D. Schulte, F. Simon, S. Stapnes, N. Toge, H. Weerts, and J. Wells. The CLIC Programme: Towards a Staged e+e- Linear Collider Exploring the Terascale : CLIC Conceptual Design Report. 2012.

[14] François Chollet et al. Keras. `https://github.com/fchollet/keras`, 2015.

[15] Martín Abadi et al. TensorFlow: Large-scale machine learning on heterogeneous systems, 2015. Software available from tensorflow.org.

[16] Rami Al-Rfou et al. Theano: A Python framework for fast computation of mathematical expressions. 2016.

[17] Yoav Freund and Robert E. Schapire. Experiments with a new boosting algorithm. In *Proceedings of the Thirteenth International Conference on International Conference on Machine Learning*, ICML'96, pages 148–156, San Francisco, CA, USA, 1996. Morgan Kaufmann Publishers Inc.

[18] D. P. Kingma and J. Ba. Adam: A Method for Stochastic Optimization. *ArXiv e-prints*, December 2014.

[19] G Amadio, A Ananya, J Apostolakis, A Arora, M Bandieramonte, A Bhattacharyya, C Bianchini, R Brun, P Canal, F Carminati, L Duhem, D Elvira, A Gheata, M Gheata, I Goulas, R Iope, S Jun, G Lima, A Mohanty, T Nikitina, M Novak, W Pokorski, A Ribon, R Sehgal, O Shadura, S Vallecorsa, S Wenzel, and Y Zhang. Geantv: from cpu to accelerators. *Journal of Physics: Conference Series*, 762(1):012019, 2016.