Classification of Electron vs Photon Showers Using Deep Learning Models

Manami Kanemura, Dr. Michael Andrews, Abhirami Harilal, Professor Manfred Paulini

Northeastern University, Carnegie Mellon University

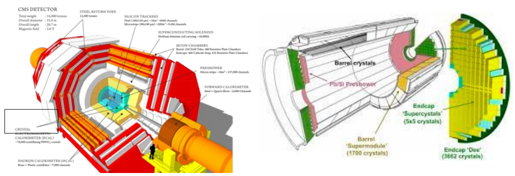
Introduction

Electromagnetic showers, or simply showers, are created by electrons or photons interacting with matter. In the vicinity of the electric field of a nucleus, an electron emits a photon losing energy while a photon creates an electron-positron pair at high enough energies. By repeating these process with the resulting particles, a cascade or a so-called shower is created (see below).

The Compact Muon Solenoid (CMS), one of the detectors at CERN's Large Hadron Collider (LHC), observes the scattered particles emerging from the collisions of two protons accelerated to almost the speed of light. CMS consists of multiple cylindrical sub-detectors, with one of the inner sub-detectors, the electromagnetic Calorimeter (ECAL), specializing in the detection and measurement of the energy of electrons and photons.

This project aims to compare the model performance of electron versus photon shower classification in the CMS ECAL using several deep learning methods.

CMS & ECAL



The ECAL, a sub-detector of the CMS experiment, consists of crystals in which the electromagnetic showers evolve, together with photodetectors that are installed at the back of each crystal to measure the energy of each shower. The ECAL is immersed in a constant magnetic field of 3.6 T created by the CMS solenoid. Showers created by electrons and photons in the ECAL look very much alike, making it difficult to distinguish both particle types with traditional approaches. To address the problem, deep learning methods are employed as a new method of classifying the particles.

Properties of Photons and Electrons

Photon

A photon is a particle with no electric charge. While there are several possible interactions of photons with matter, pair production into a positron and an electron is the most dominant interaction of photons with matters in high-energy situations at particle energies from collision events at CMS.

Electron

An electron is a particle with an electric charge. In ECAL, electrons lose energy when entering the crystal by emitting a photon in the electric field of a nucleus resulting in a photon (Bremsstrahlung) with the electron continuing its path but with reduced energy.

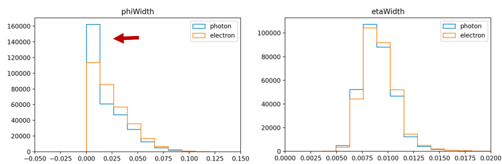


Figure 1: Histogram of ECAL shower width in the phi direction (left) and in the eta direction (right) with normalized data.

The CMS magnetic field allows us to discriminate particles with or without electric charge. In the case of photons and electrons, electrons are bent in the phi direction by the magnetic field (applied in the z-direction) before reaching the ECAL while photons straightly travel to the ECAL. The bend of the electron path causes the radiation of photons, which broadens the shower in the phi direction as shown in Figure 1. The shower width in the eta direction where no bending occurs is shown as a comparison indicating no difference between electron and photon showers (Figure 1).

Methodology

In the first project, we feed an image dataset of ECAL energy deposits to Artificial Neural Networks (ANN), Convolutional Neural Networks (CNN), and ResNet models. We then compare the performance of each network in order to validate the models. In the second project, we convert the image dataset into nodes featuring a matrix and feed the converted dataset to an EdgeConv model, a kind of Graph Neural Network (GNN). During the training of the GNN, the connections (edges) among the nodes are created by the each EdgeConv layer. Finally, the performance of the ANN, CNN, ResNet, and GNN models are compared.

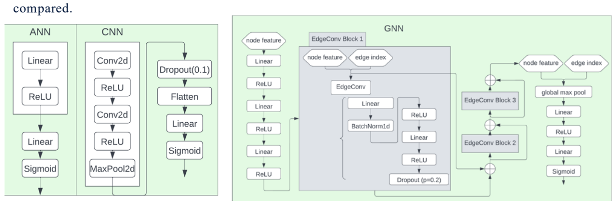


Figure 2: ANN and CNN architecture (left) as well as GNN architecture (right)

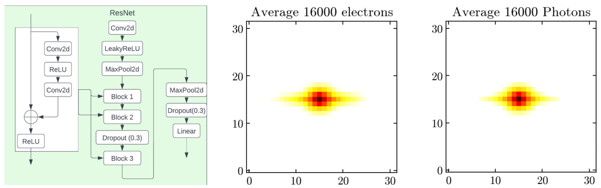


Figure 3: Left: ResNet architecture. Center: Average of energy deposits of electrons. Right: Average of energy deposits of photons.

Results

1. ANN vs CNN vs ResNet

We trained ANN, CNN, and ResNet models on the image dataset shown in Fig. 3 center and right. Since ANN takes 1D data as input, the 2D images were flattened while the image data were directly input to the CNN and ResNet models. Fig. 4 center shows the performance score from the Receiver Operating Characteristic (ROC) curve in form of the Area Under the ROC curve (AUC) score resulting in a value of 0.742 for ANN and 0.741 for CNN (Fig. 4 left). Additionally, Fig. 4 right illustrates the AUC score for the ResNet model, which reaches 0.791. Among the three models, the ResNet classifier obtained the highest AUC performance score.

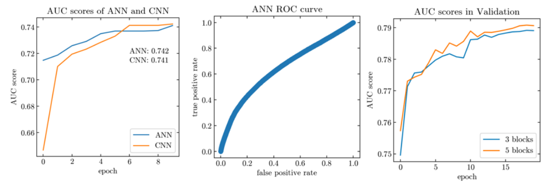


Figure 4: Left: AUC scores of ANN and CNN models. Center: ROC curves of ANN model. Right: AUC trend of ResNet model.

In general, a CNN model is more powerful for image classification than an ANN model as it learns spatial features during training while ANN learns only pixel features. However, the AUC score of CNN is almost identical to that of ANN, which could be caused by insufficient training time.

2. GNN

2-1 GNN & EdgeConv

CNN and ResNet are expected to be superior to ANN for image classification since they learn the spatial features of images. However, they fail to classify an image if it is rotated. Conversely, GNN is a graph-based algorithm that ignores Cartesian coordinates. It frees the model from 2D grids and captures data as nodes and edges. EdgeConv, a type of GNN, learns nodes and creates edges in every layer, which makes the EdgeConv GNN more power than the other GNN models.

2-2 GNN

After converting the image dataset into a node feature matrix, we input it to the GNN model. Figures 5 and 6 exemplify how GNN outputs image data. The graphed photon (Fig. 5 left) is denser than the graphed electron on the right-hand side because electrons are bent in the phi-direction due to the CMS magnetic field, leading to more spread of the shower in the phi-direction. Interestingly, Figure 6 left shows an example of pair production from a photon, where one group is a positron and the other group is an electron. Figure 6 right shows an example of a Bremsstrahlung process.

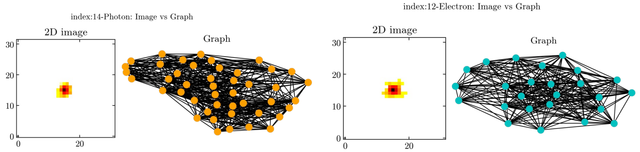


Figure 5: Left: Image and Graphed photon. Right: Image and graphed electron.

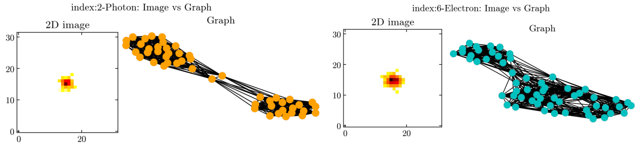


Figure 6: An example of pair production (left) and of Bremsstrahlung (right).

Finally, we feed the node features to the GNN model and obtain an AUC score of 0.766. Note, Fig 5-6 were generated from the trained model.

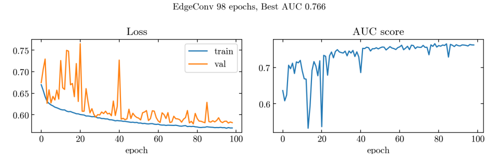


Figure 7: Training and validation loss and AUC scores of GNN model

Discussion

In this project, performance comparisons were studied for ANN, CNN, ResNet, and GNN models in an attempt to classify electron versus photon showers in the CMS ECAL. As shown in Table 1, the ResNet model yielded the highest AUC score of 0.791, followed by GNN with a 0.766 score. While ResNet achieved the best AUC scores due to the deep network, GNN produced better performance than CNN as it does not rely on a Cartesian coordinate system. The next step of this project is to study different clustering algorithms and coordinate information as well as explore different graph architectures.

AUC scores

ANN	0.742
CNN	0.741
ResNet	0.791
GNN	0.766

Table 1: Summary of AUC scores

Acknowledgement

This project was conducted during Summer Undergraduate Research Program and supported by by the Mellon College of Science at Carnegie Mellon University. I greatly appreciate my mentors, Dr. Michael Andrews and Abhirami Harilal, and my faculty advisor, Professor Manfred Paulini who provided insight and expertise that greatly assisted this project.

Reference

[1]L. Neukermans et al., The CMS experiment at the CERN LHC, vol. 3, no. 8. IOP Publishing, 2008, pp. S08004–S08004.

[2] Source code for models trained on image datasets - https://github.com/Manami-16/Intro_to_Deep_Learning_for_Particle_Physics.git