Classify Quark and Gluon Initiated Jets using CNNs

Kamil Laurent

September 4, 2023

Physics of the Problem

Quark vs Gluon Initiated Jets

In particle accelerators, the goal is to study the fundamental particles of the standard model. The fundamental particles include quarks and gluons, but they are not visible. Therefore, we deduce their properties using the characteristics of the jets initiated by these particles.

Since we only observe the final jets of particles, we need techniques that can classify quark-initiated and gluon-initiated jets. In this project, we aim to do that using a convolutional neural network (CNN).

Physics of the Problem

Physics of the Detector

The detectors of particle accelerators measure the transverse momentum (p_T), energies, and angular positions (θ, ϕ) of different final-state particles:

- Charged hadrons: p, $\pi^{\pm}$, $K^{\pm}$
- Charged leptons: $e^{\pm}$, $\mu^{\pm}$
- Neutral particles: γ, neutral hadrons

The detected particles are then grouped into jets using algorithms (e.g., anti-k_t). Once each jet has been reconstructed, we want to determine whether each jet has been generated from a quark or from a gluon. To apply image recognition to this problem, we need to construct an image of each jet and then apply the CNN to jet images.

Physics of the Problem

Quark Jets vs Gluon Jets

The observable difference between the jets, are the fact that quark-initiated jets are narrower, and each particle carries more energy, while gluon-initiated jets are wider, and each particle carries less energy. This is due to the color charges of quarks (one color) and gluons (both color and anticolor).

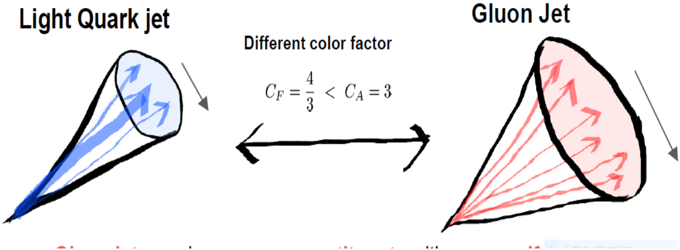


Figure: Shape of quark and gluon initiated jets from *Discriminating quark/gluon jets with deep learning* n.d.

Dataset Generation

Processes Involved

We generate a dataset of 5000 images for each class (quark/gluon) using the event generator Pythia8 and the algorithm Anti-k_t from FastJet to reconstruct the jet images. We simulate pp collision events at $\sqrt{s} = 13$ TeV, We use a trashhold value of $|\eta| < 2.5$. Only final-state visible particles are taken as input for FastJet, the particles are clustered with a jet radius $R = 0.4$.

Dataset Generation

Processes Involved

To label the reconstructed jets correctly, we run two separated simulations. For producing only quark-initiated jets, we turn on only the QCD hard interactions:

$$q\bar{q} \rightarrow q\bar{q} \tag{1}$$

$$gg \rightarrow qq \tag{2}$$

$$qq \rightarrow qq \tag{3}$$

For producing gluon-initiated jets, we only turn on the hard processes:

$$q\bar{q} \rightarrow gg \tag{4}$$

$$gg \rightarrow gg \tag{5}$$

Dataset Generation

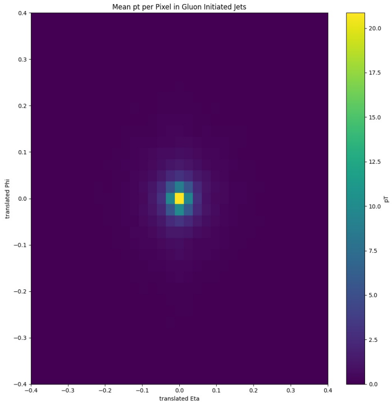
Rescaling and Crop

We create a two-dimensional histogram for each jet. The coordinates are the azimuthal angle ϕ and the pseudorapidity $\eta = -\ln\left[\tan\left(\frac{\theta}{2}\right)\right]$. We set the histogram to have a 33x33 pixel region centered on the jet's p_T centroid, with dimensions $\Delta\eta = \Delta\phi = 0.8$. In Figure 2, we can see the mean histogram for quark-initiated jets (b) and for gluon-initiated jets (a).

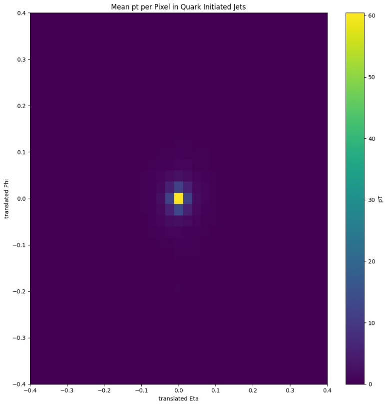
We generate a first set of 5000 labeled images per category by selecting jets with measured p_T in the range of $200 - 220$ GeV.

Dataset Generation

Rescaling and Crop



(a) Mean p_T of Gluon-Initiated Jets



(b) Mean p_T of Quark-Initiated Jets

Figure: Histogram of Average Quark and Gluon-Initiated Jets

CNN Architecture

First Attempt

We try two different model arcihtecture, the first is the architecture used in Patrick T. Komiske and Schwartz 2018, that consist of:

- 3 convolutional layer with a ReLU activation and a maxpooling layer
- a dense layer with a ReLU activation
- an output layer of one unit with sigmoid activation

for the convolutional layers we use 64 features detectors, with a 8x8 dimension and a dropout of 0.5 for the first layer and a dimension of 4x4 and dropout of 0.25 for the last two conv layers.

CNN Architecture

Simplified Architecture

Since the previous architecture was used in Patrick T. Komiske and Schwartz 2018 to train a model using 100000 images and we have only 5000 images, we try a simplified architecture, with less feature detectors and with no dropout:

- 3 convolutional layer with a ReLU activation and a maxpooling layer
- a dense layer with a ReLU activation
- an output layer of one unit with sigmoid activation

for the convolutional layers we use 32 features detectors, with a 3x3 dimension and no dropout. This structure perform a little better and the training require way less time.

Results

Test of the Model

We train our models using the original set of jet images and we obtain an accuracy of 0.74 and a minimum loss on validation set of 0.54.

We also experiment to tain the model with a new set of images of jets including only final state charged particles, but the result is slightly worse (accuracy on validation set $= 0.72$, loss on validation $= 0.55$).

In image 2 one can see the ROC curves for the two models, one for the model trained on the "complete" jets (orange) and one for the model trained on the "charged particles" jets (green)

Results

Test of the Model

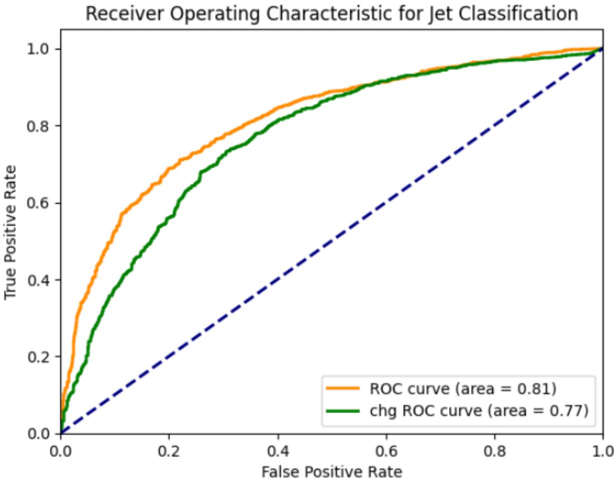


Figure: ROC curves for trained model on complete jet

Results

Adding More Images

The red curve is the resulting ROC of a test prediction using the model trained with 9000 images per set (best accuracy = 0.76, minimum loss = 0.52)

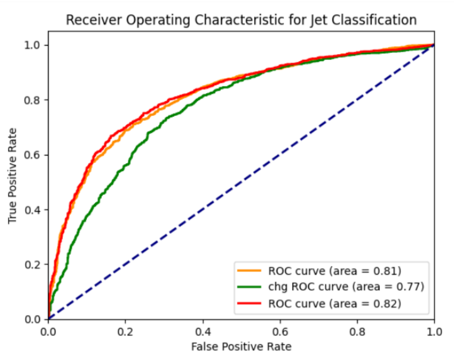


Figure: ROC curves for trained model on 9000 standard jets per set

To reach even better results one can try different approaches:

- add even more images
- use different data preprocesing tecniques (e.g. pixel standardization)
- use the k-folding technique
- divide the image into 3 channel (charged, neutral and charge multiplicity)

Bibliography

Discriminating quark/gluon jets with deep learning (n.d.).

Patrick T. Komiske, Eric M. Metodiev and Matthew D. Schwartz (2018). "Deep learning in color: towards automated quark/gluon jet discrimination". In: DOI: arXiv:1612.01551v3[hep-ph].