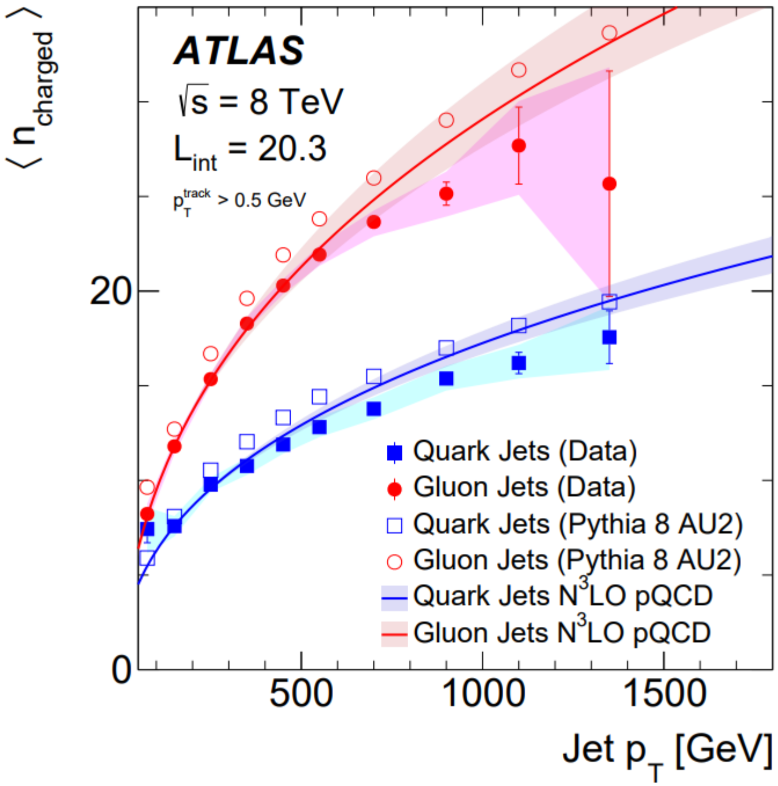
Jet Image and q/g Tagging with Convolutional Neural Networks

ATL-PHYS-PUB-2017-017

Zihao Jiang
Stanford University

Outline

- q/g tagging is an important topic for VBF analyses
- Conventional q/g tagging uses high level quantities of jets such as track multiplicity, jet width, etc.
- We investigate the possibility of a different representation of detector data, i.e. jet image, and seek to apply computer vision techniques for q/g tagging



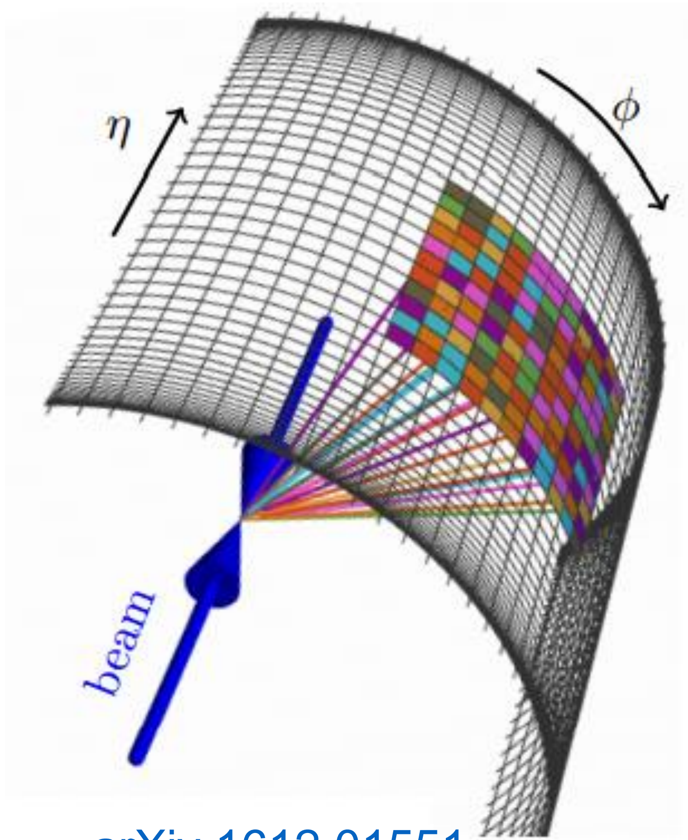
STDM-2015-12

Jet Image

- Represent jets as images; Treat calorimeter cells as pixel

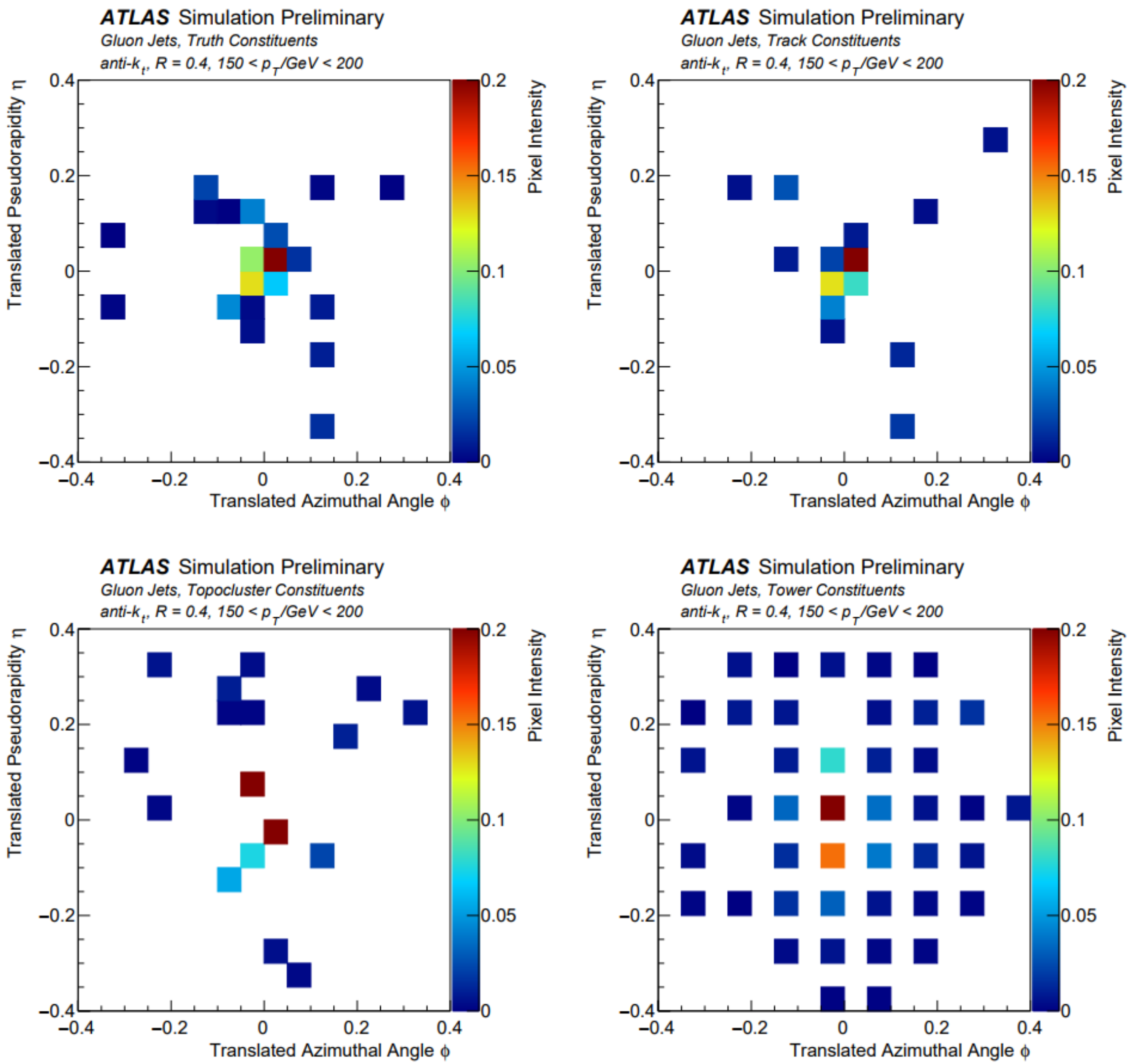
  JHEP 02 (2015) 118
  JHEP 07 (2016) 069
  arXiv 1612.01551

- Direct use of computer vision techniques
- All constituents (truth particle/ track/ topo clusters/ tower) in jets are boosted and rotated such that η(jet) = 0 and Φ(jet) =0
- Crop jet constituents into 16x16 grid (Δηx ΔΦ = 0.05x0.05);
- The intensity of a pixel is the sum of p_T of the constituents in that pixel
- Each image normalized such that $\sum_i \ I = 1$



arXiv 1612.01551

Jet Image Example

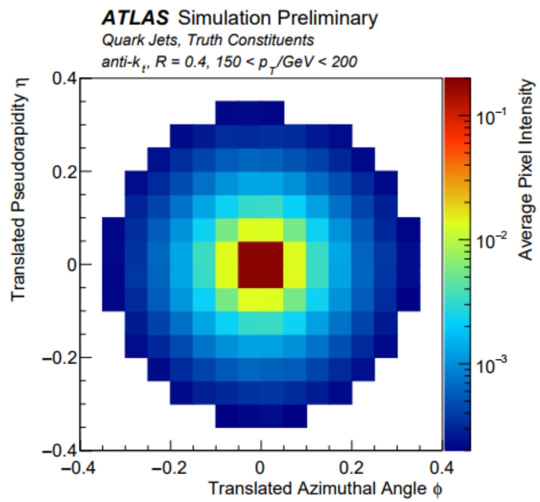


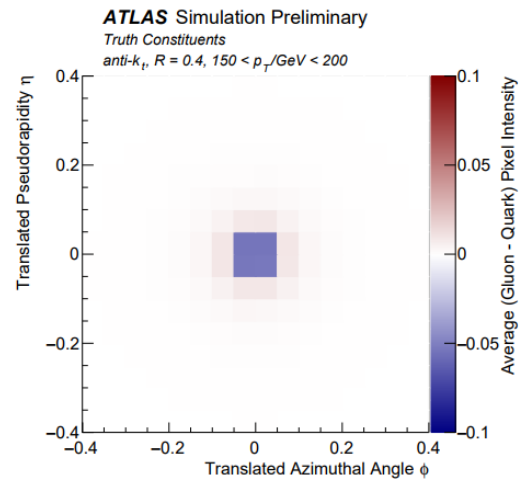
The stable particle (top left), track (top right), topocluster (bottom left), and tower (bottom right) images for a generic gluon jet image.

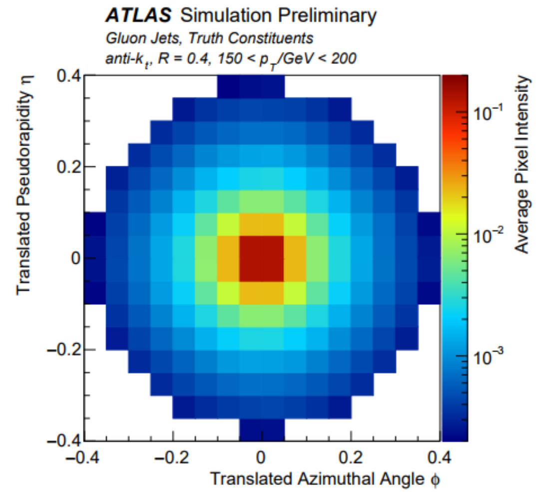
The tower image has gaps between hit pixels because the 0.1x0.1 towers are projected onto 0.05x0.05 jet image.

Average Jet Image

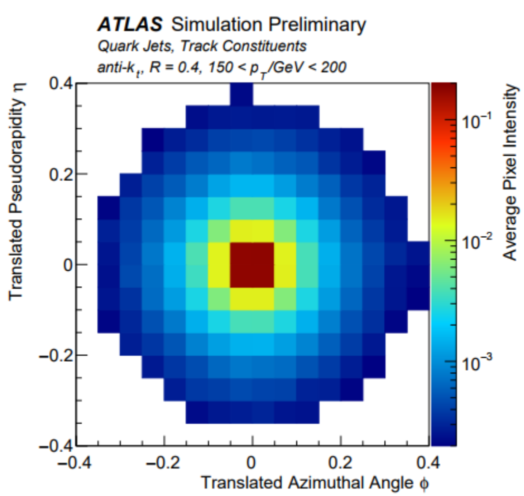
Truth images

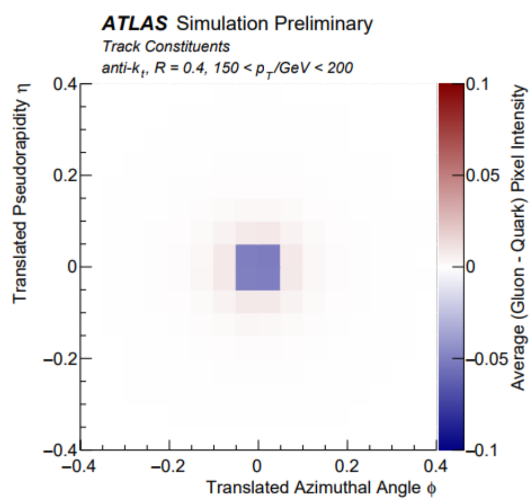


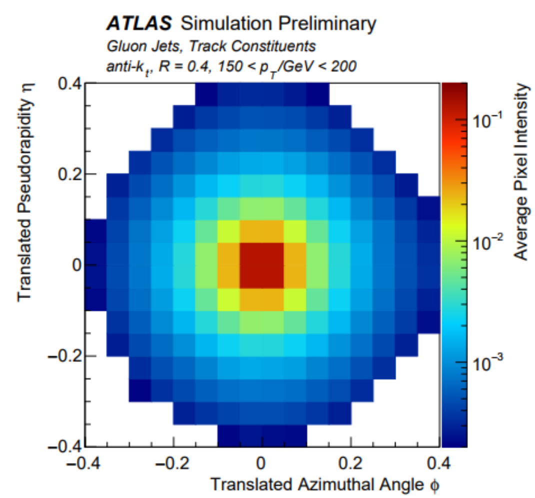




Track images

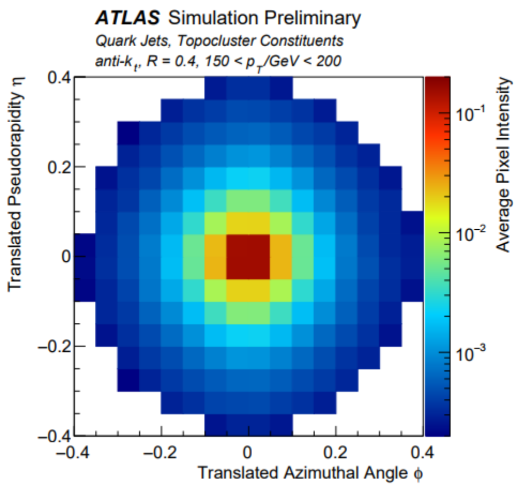


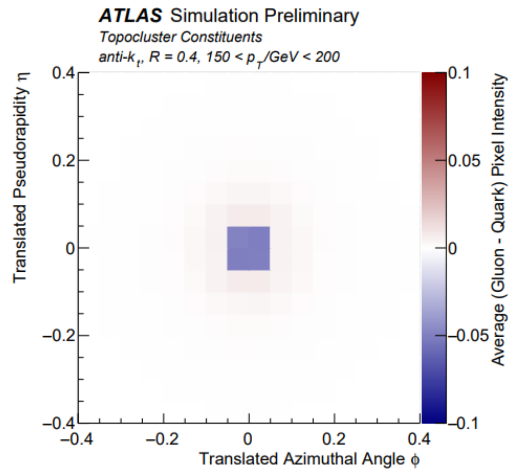


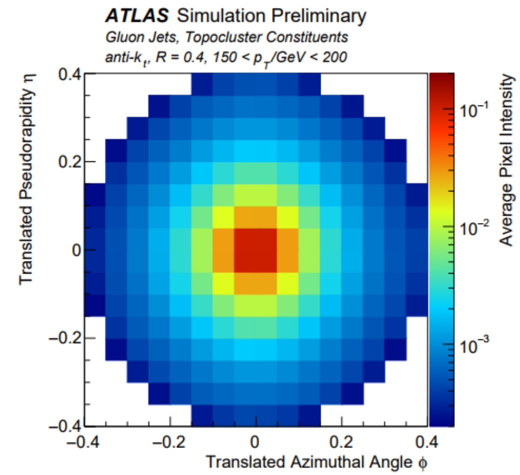


Average Jet Image

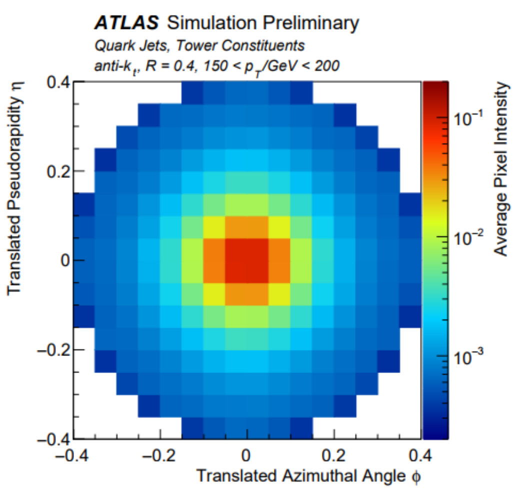
Topocluster images

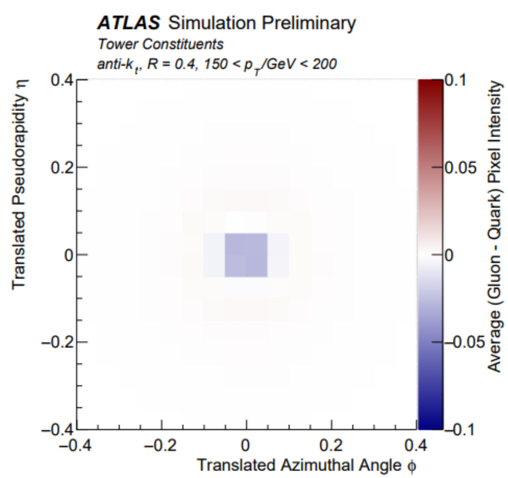


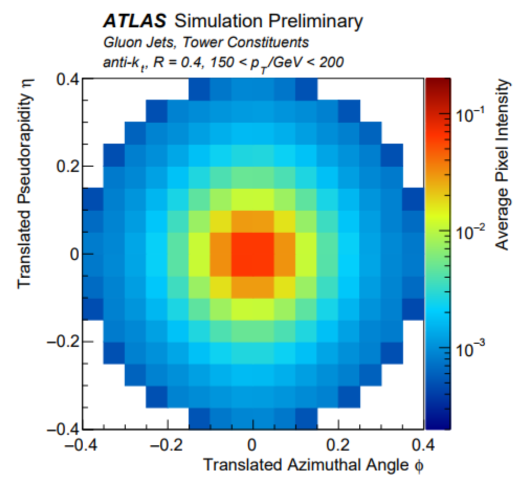




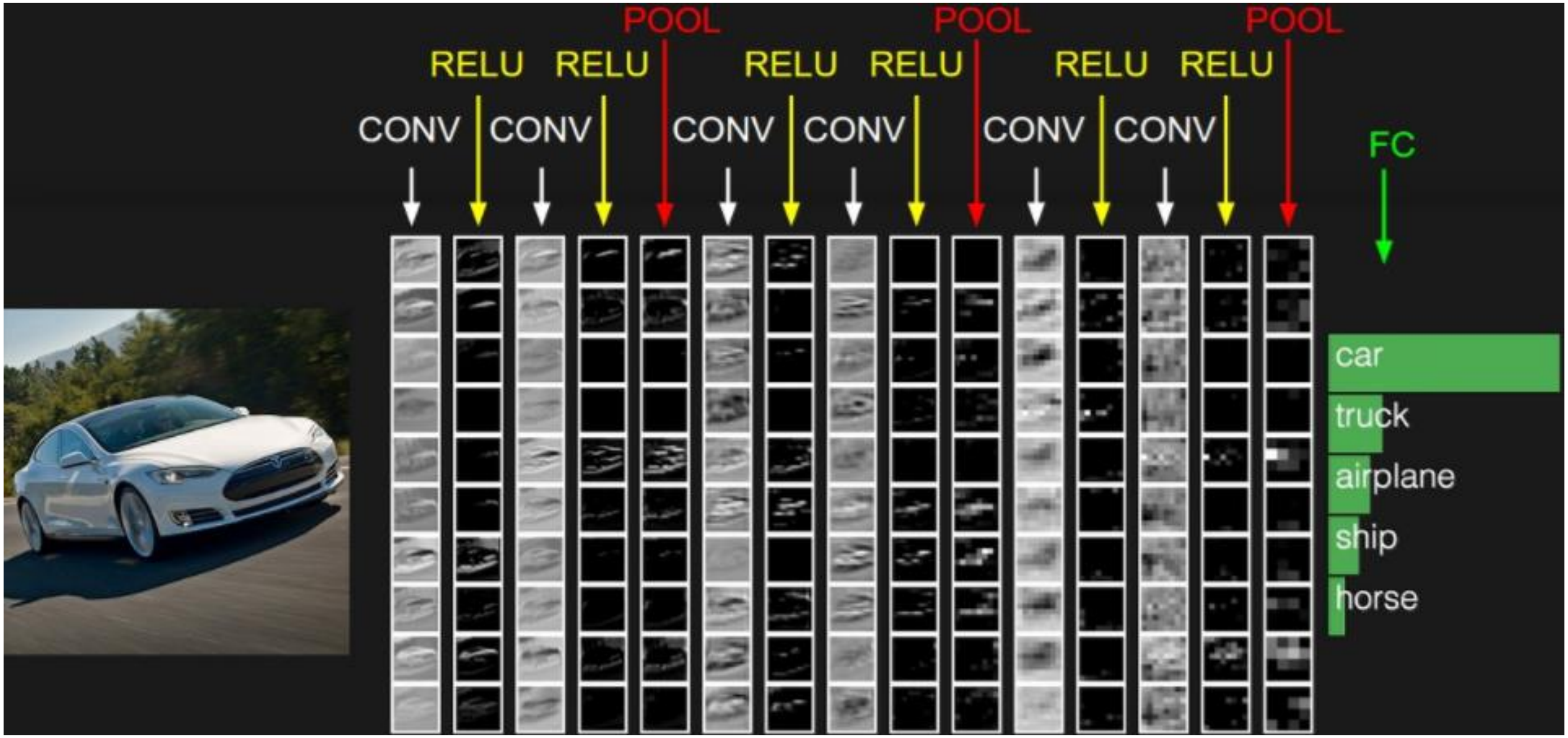
Tower images





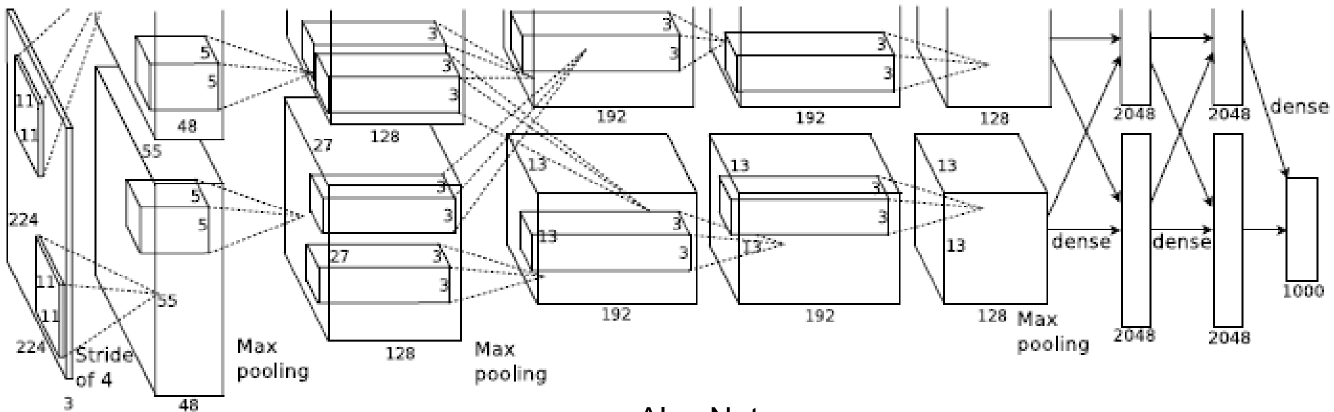


Deep CNN and Image Recognition

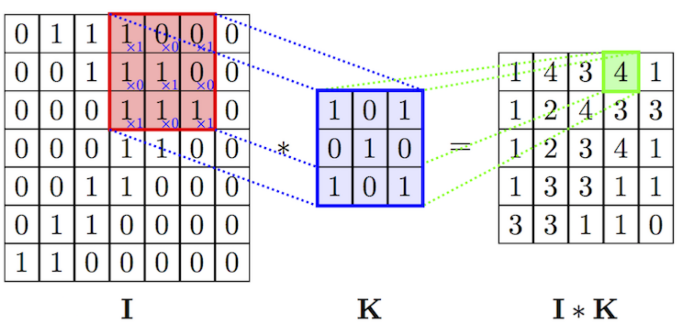


http://cs231n.stanford.edu/

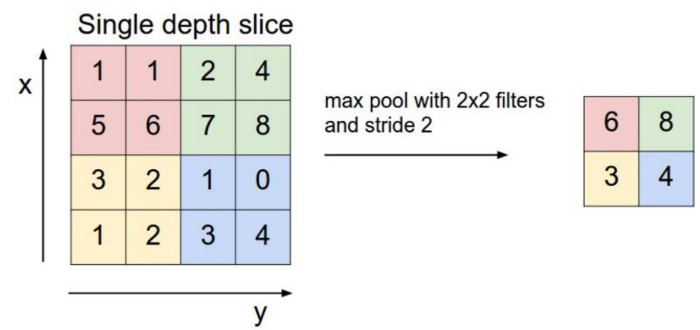
Deep CNN and Image Recognition



Alex Net
[A. Krizhevsky et al.](#)



Convolution Layer



Pooling Layer

CNN for q/g tagging

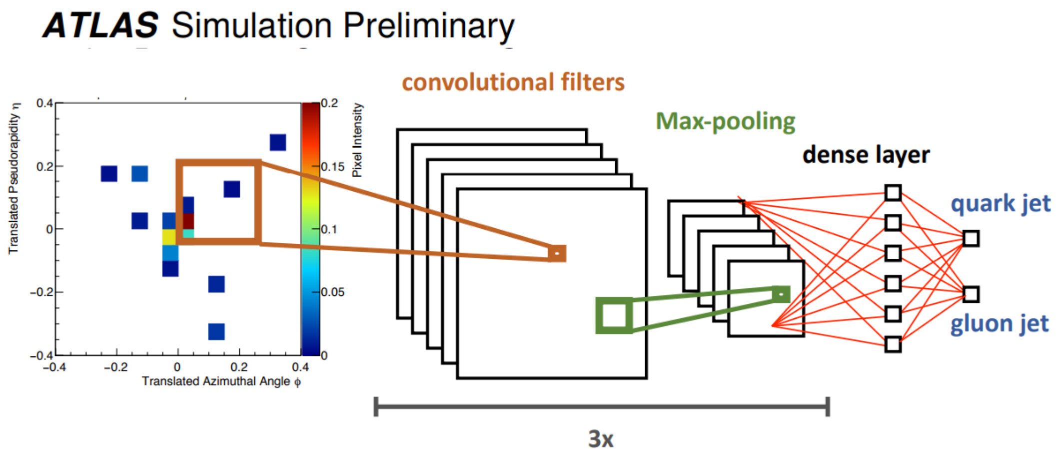
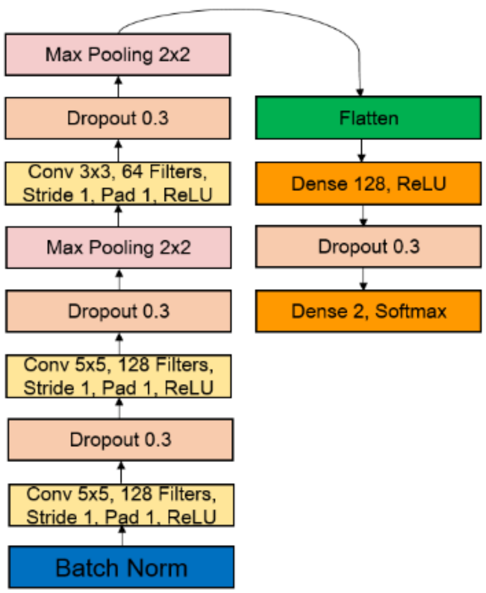


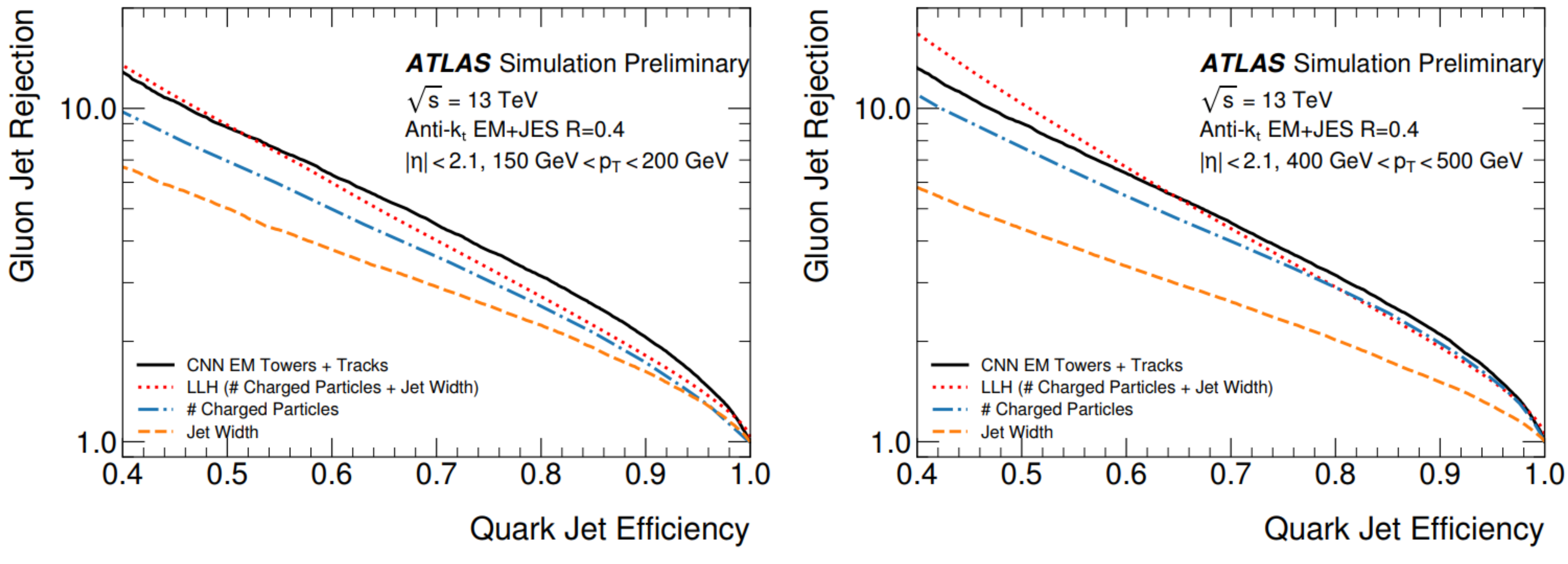
Illustration of CNN q/g tagger



Detailed breakdown of the network structure

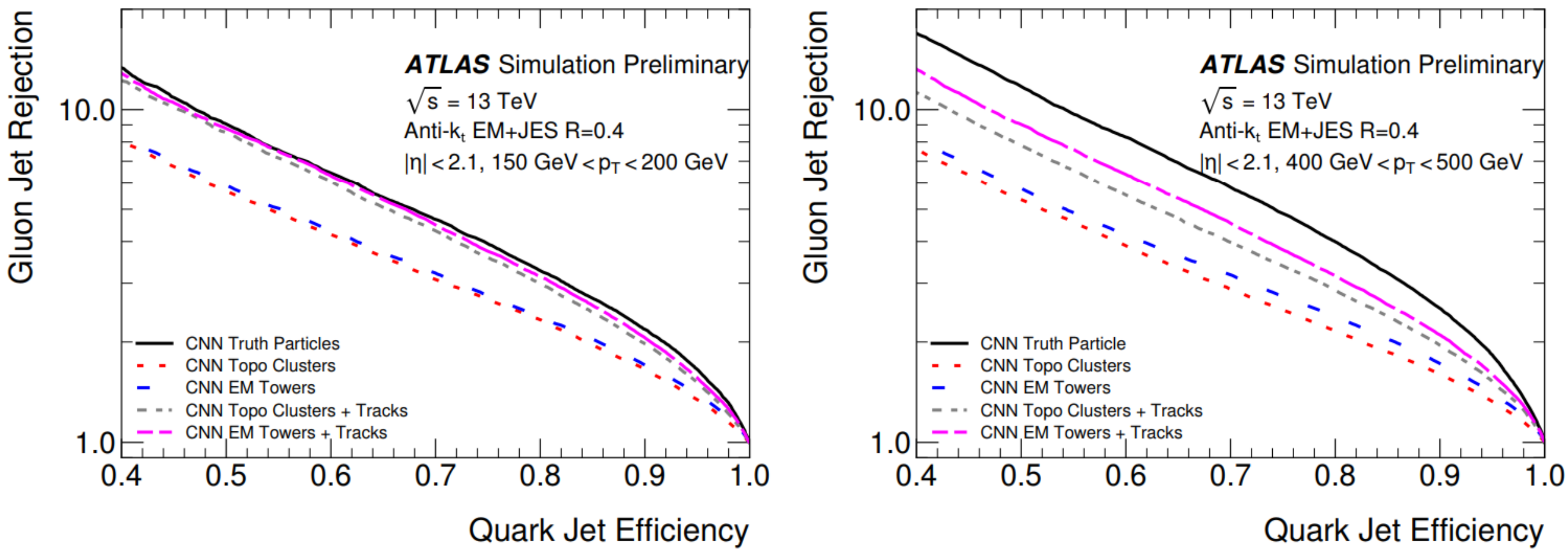
- Once jet images are formed we apply convolutional neural network (CNN) for tagging
- The image can have multiple "colors" (channel), e.g. we can simultaneously feed the network with track+tower images
- The training of the network utilizes NVidia Tesla K80 GPU with 224000 jets as training sample and 56000 jet images as testing sample
- The network is trained with 50 epochs with ~80s/ epoch

CNN: Performance



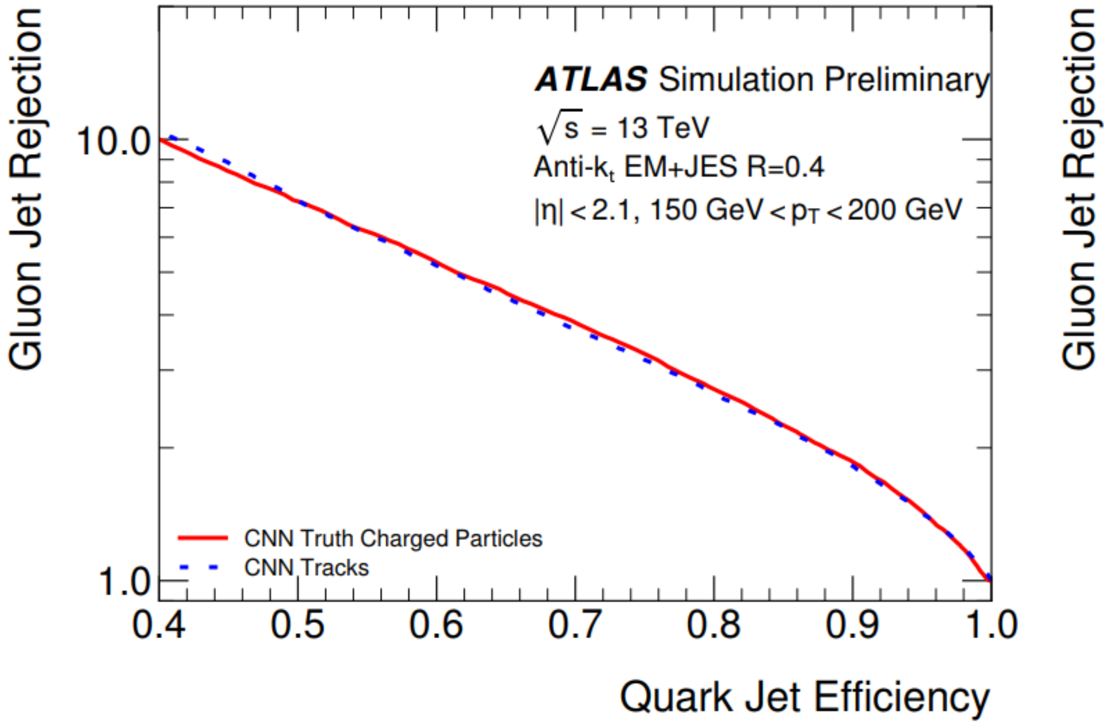
- Gluon jet rejection as a function of the quark jet efficiency using physics motivated observables and jet image discriminants for jets with 150<pT<200 GeV (left) and 400<pT<500 GeV (right)
- The LLH is a tagger constructed from the optimal (likelihood) combination of N(track) and jet width
- The CNN tagger outperforms jet width and track multiplicity and is better or comparable to the combination of them

CNN: Inputs



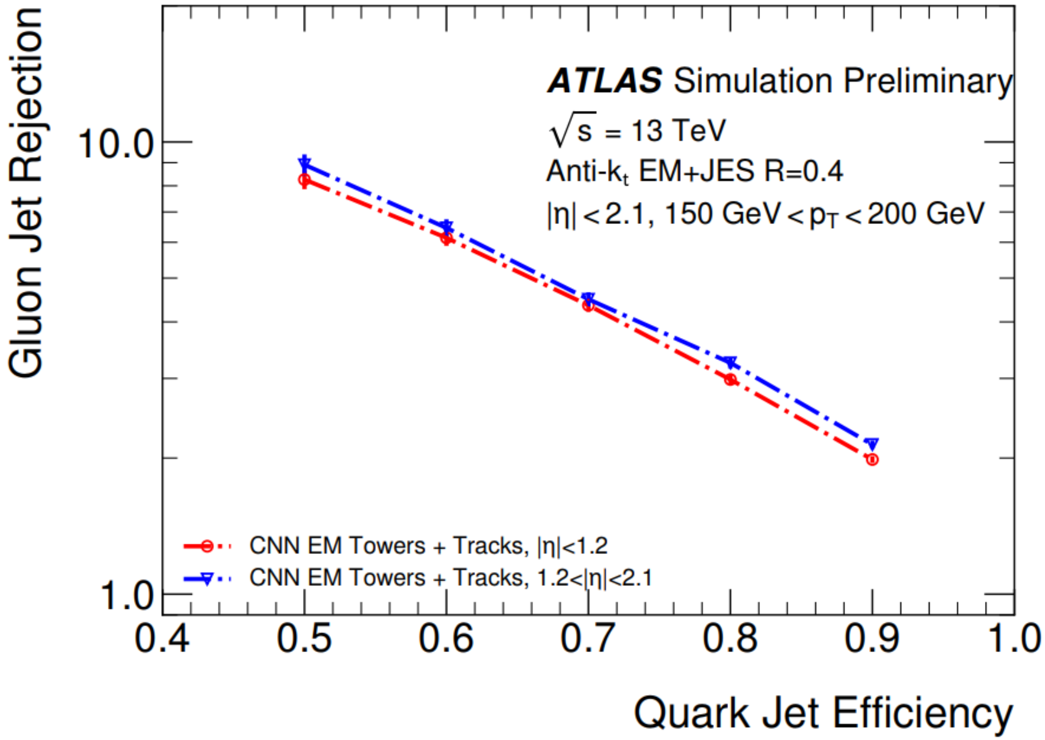
- Gluon jet rejection as a function of the quark jet efficiency using the CNN tagger with different inputs for jets 150<p_T<200 GeV (a) and 400<p_T<500 GeV (b).
- Best discriminating CNN is built with using 2-channel image combining towers and tracks. In general CNN with using tower as inputs has more discriminative power than topo clusters.

CNN: Truth Comparison



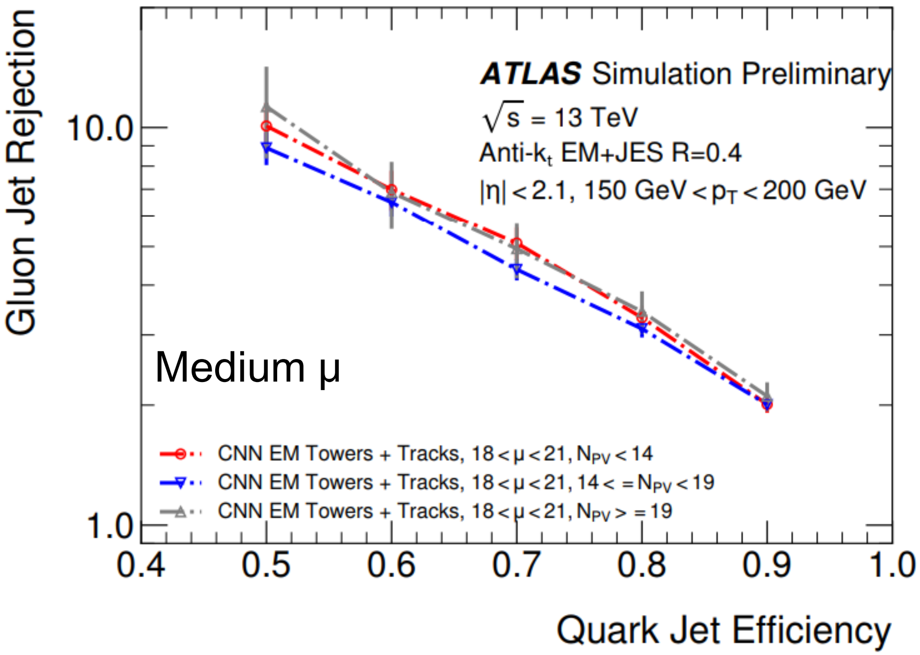
- Comparison between using track images and truth charge particle images as input
- As might be expected, due to the precise measurement of charged-particle trajectories, there is little difference between the particle- and detector-levels. Degradation due to efficiency and resolution effects are only expected at much lower and higher transverse momenta

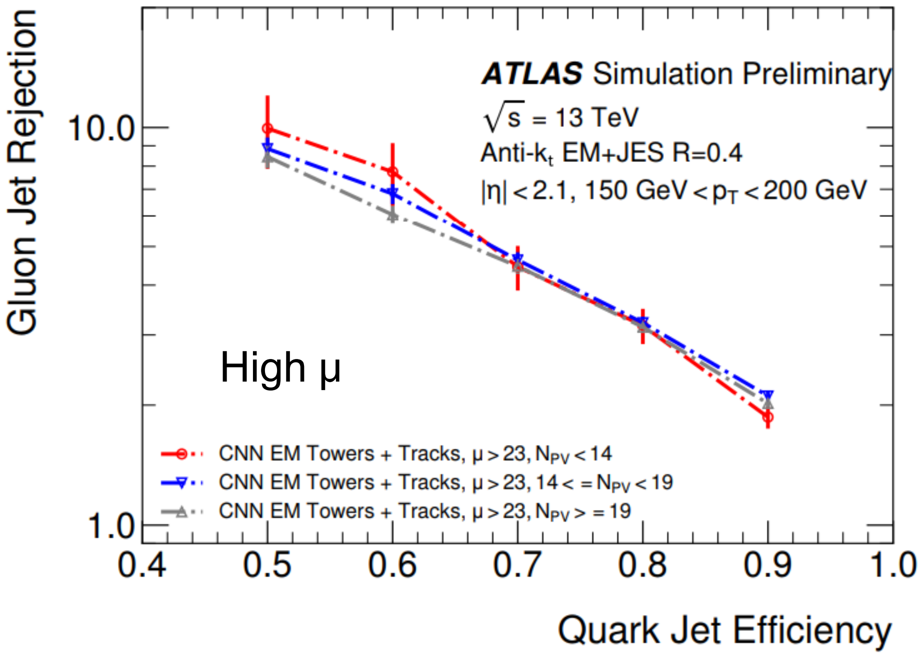
CNN: η Dependence



- Comparison between different |η| ranges. The full |η| range (|η| < 2.1) is used for training
- Similar performance is achieved when testing the tagger on jets that are predominately in the barrel (|η| < 1.2) to those that are in the transition region between the barrel and endcap calorimeters (1.2 < |η| < 2.1)

CNN: Pile-up Dependence

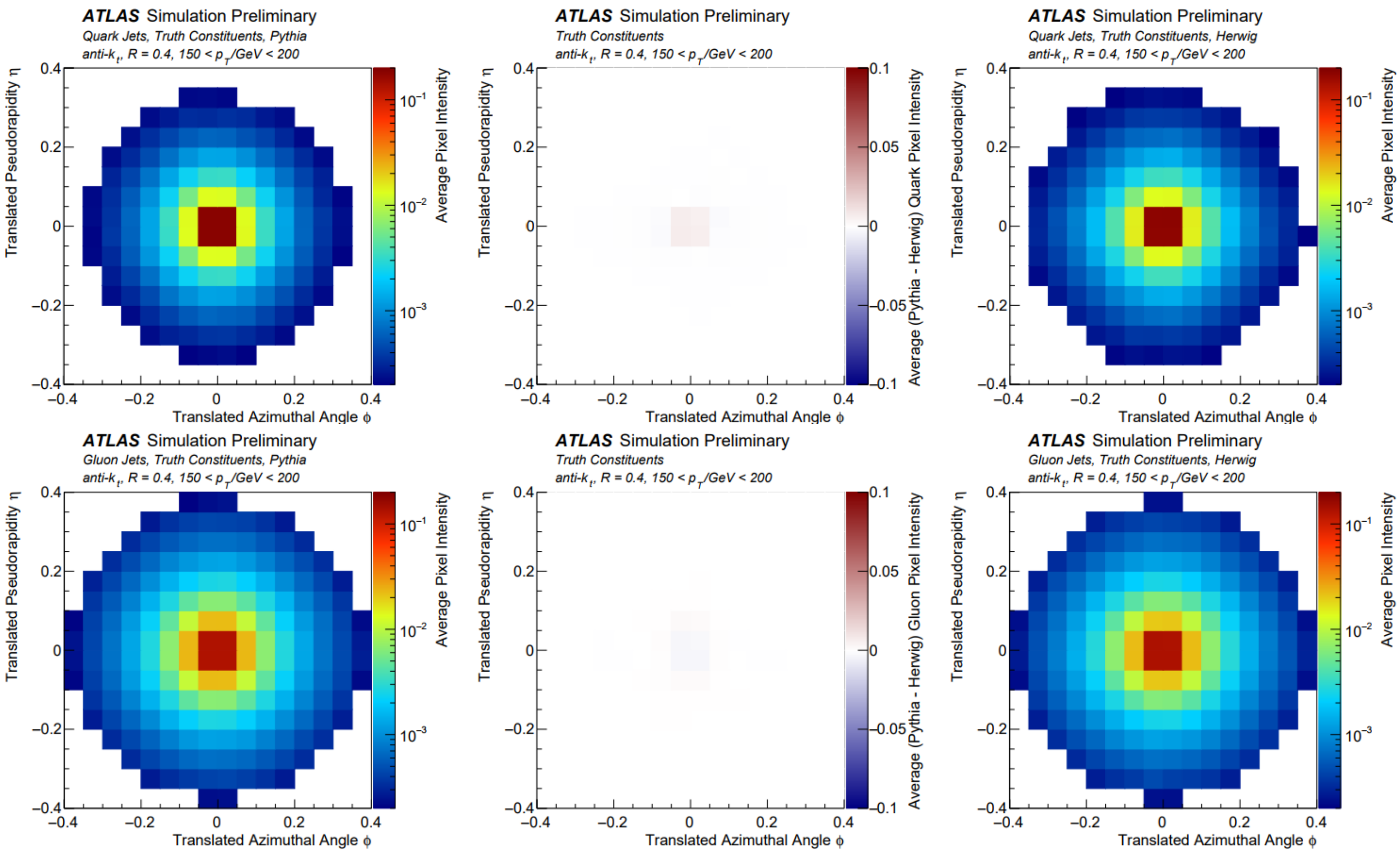




- Comparison the tagger performance in two different regimes of the average number of collisions per bunch crossing (μ) regimes corresponding to the out-of-time pileup representative of LHC Run 2 conditions

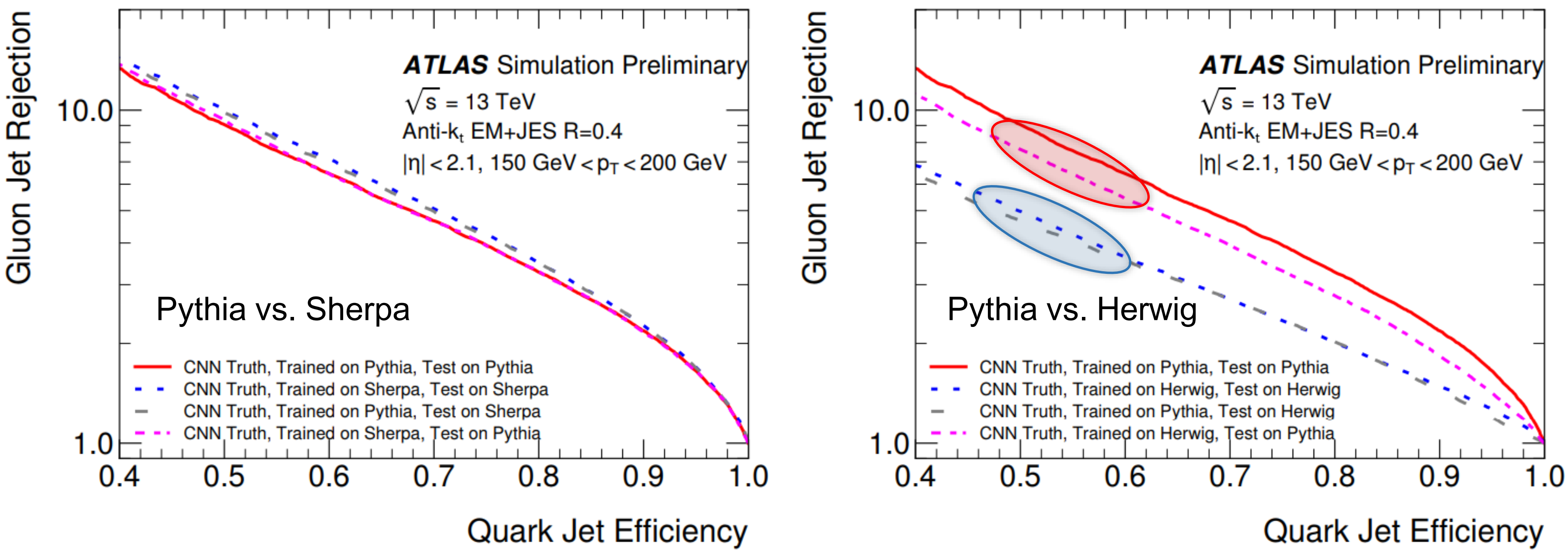
- The distributions of the pixel intensities do vary with pileup, but the performance of the CNN tagger is found to be robust

CNN: Generator Dependence



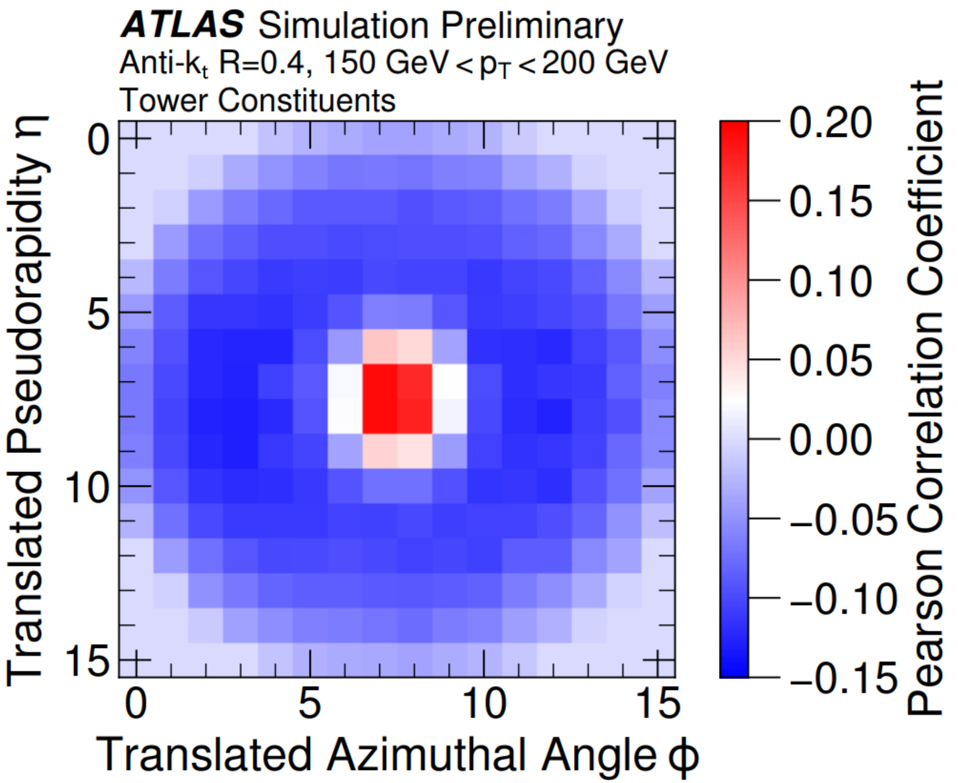
- The radiation pattern inside gluon jets is similar between the Pythia and Herwig, whereas there are larger differences for quark jets
- For Herwig, the quark jets and gluon jets look more similar

CNN: Generator Dependence



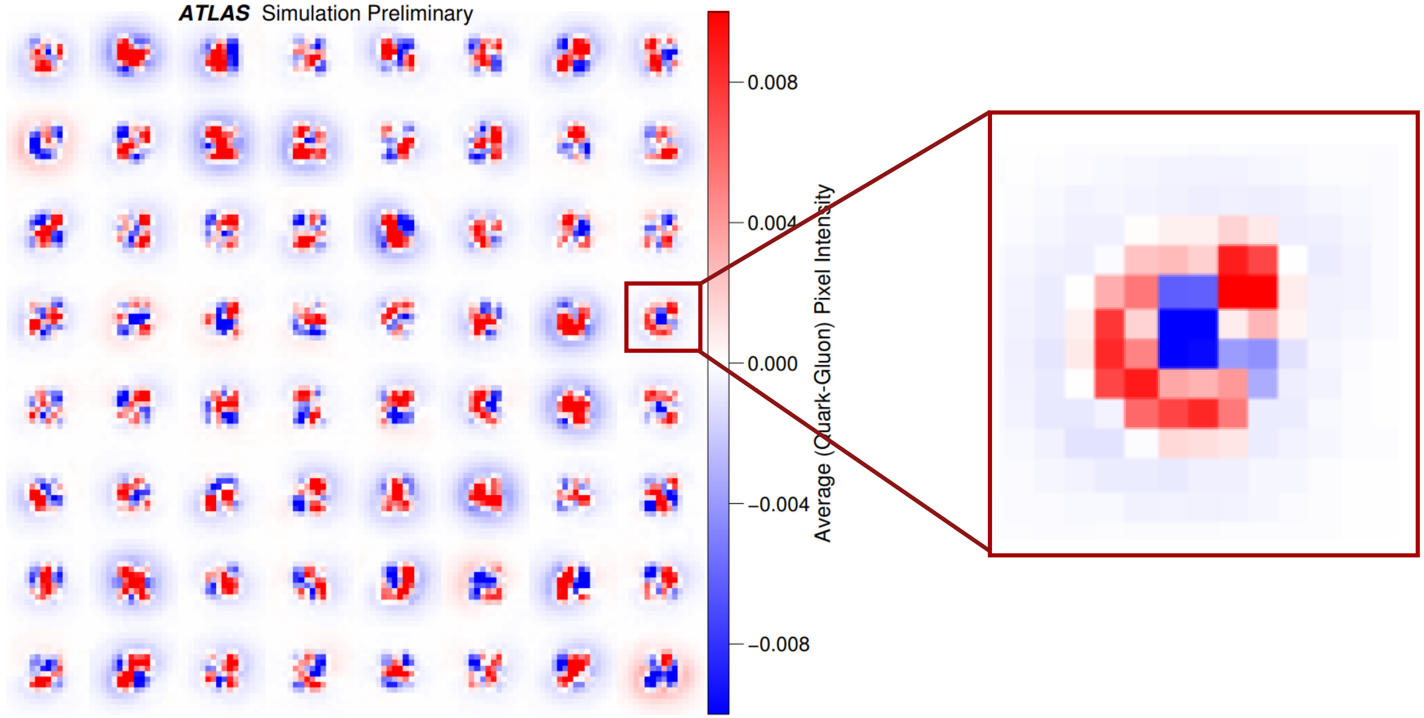
- When generators produce different images, the CNN returns a different performance when training and testing with images from various simulators
- However, if the same network is used for testing and only the training sample is varied, the gap in performance is mostly removed
- One explanation is that the network is learning robust features for quark versus gluon tagging, but the degree to which the features are expressed in the radiation patterns varies between generators

CNN: Visualization



- Per-pixel correlation with CNN tagger output
- The pixel intensity in the image core and the CNN score are highly correlated and thus important for quark-tagging. The outer pixels and CNN score are anti-correlated and thus important for gluon-tagging

CNN: Visualization



- Average convolved filter differences for jet images. The filters of the first convolutional layer are considered
- The conv filters extract the raw features of the images; Circular blobs wraps around the center
- Some filters are rotational copies of each other, indicating the network learned rotational invariance.

Conclusion

- Jet image is a novel representation of detector data
- Computer vision techniques can be applied to directly for classification problems
- We present the first results of CNN application in q/g tagging at ATLAS detector
- CNN q/g tagger yields better (or comparable) performance as taggers combining high level jet quantities
- CNN q/g tagger is robust against η and pile-up variations
- CNN q/g tagger learns features of jets which are consistent with our physics intuition