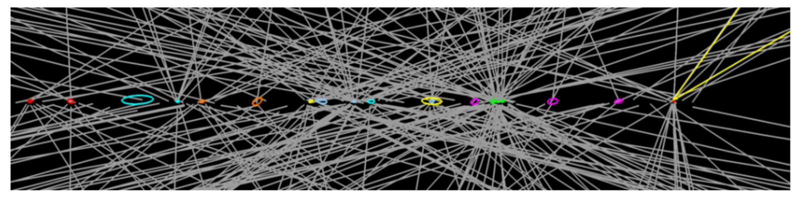
Convolutional Neural Networks for Pile Up ID in ATLAS

Murtaza Safdari (murtazas@stanford.edu)

Motivation

- ATLAS is a physics detector on the LHC looking at proton-proton collisions.
- It sees collimated streams of particles, called jets, in its equipment.
- Jets are crucial to studying any particle physics process.
- The detector records many fake jets, called Pile Up (PU) jets[1] due to particles crossing over from different interaction points.
- Goal: To develop a classifier that discriminates between real (HS) and PU jets better than the current standard[2] using CNNs.



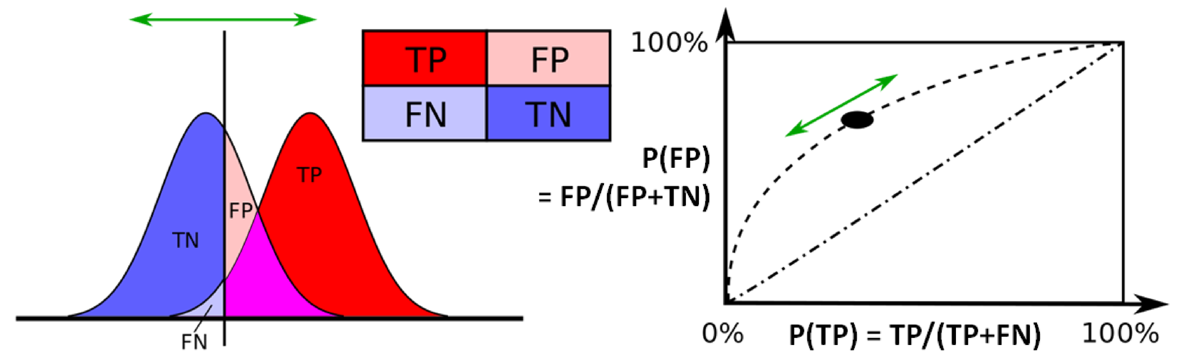
Baseline

The ATLAS standard for discriminating between HS and PU jets in the central region is using the Jet Vertex Tagger (JVT)[2]. The jet Rpt variable serves as a good proxy for the JVT, and shall serve as the baseline against which network performance will be measured.

In addition to jet Rpt, a baseline Neural Network has also been trained using jet Rpt and p_T as input features. This is theoretically a more challenging baseline to work with, as it uses p_T information to improve predictions.

Evaluation Metric and Loss

In addition to accuracy as a metric used to gauge the performance of a discriminator in ATLAS, we also use Receiver Operating Characteristic (ROC) curves.



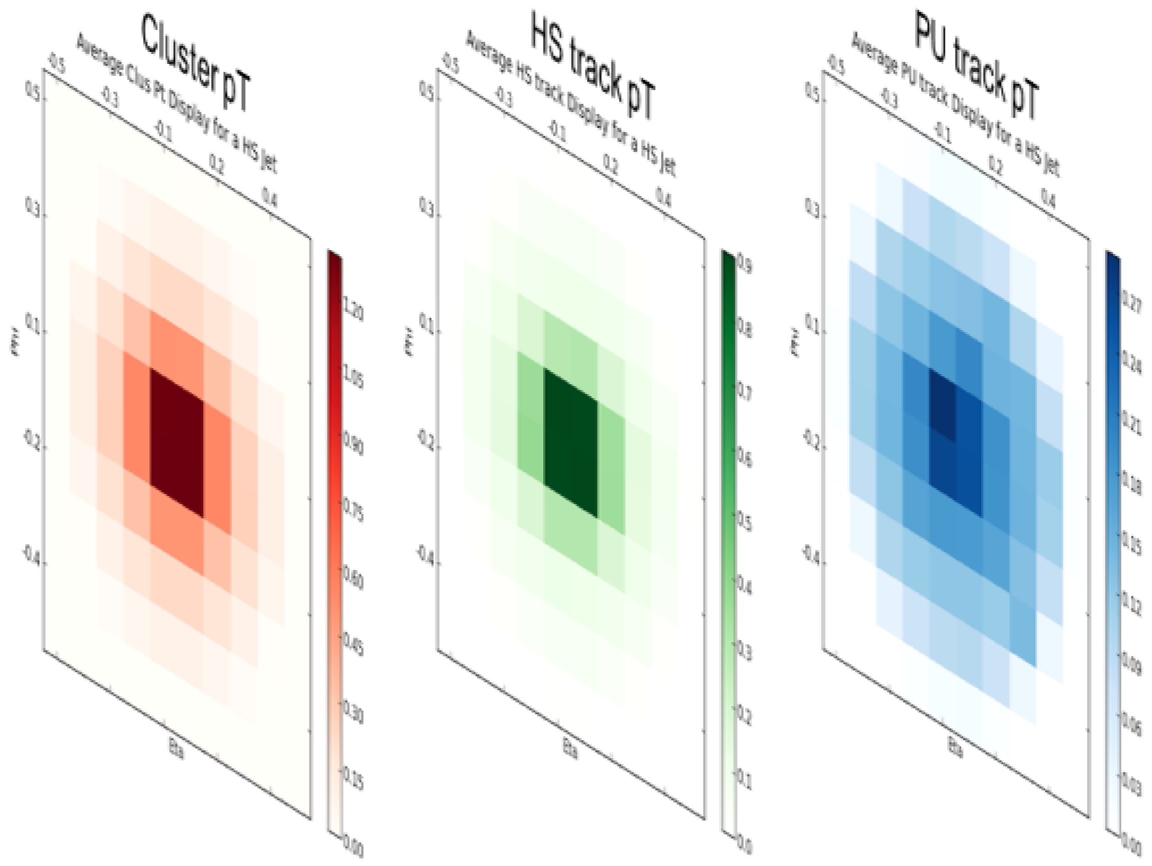
Cross Entropy Loss is used as it tries to accumulate the probability distribution on the true labels, making the output of the network a good discriminator, as opposed to margin losses which settle once a margin is achieved.

Dataset

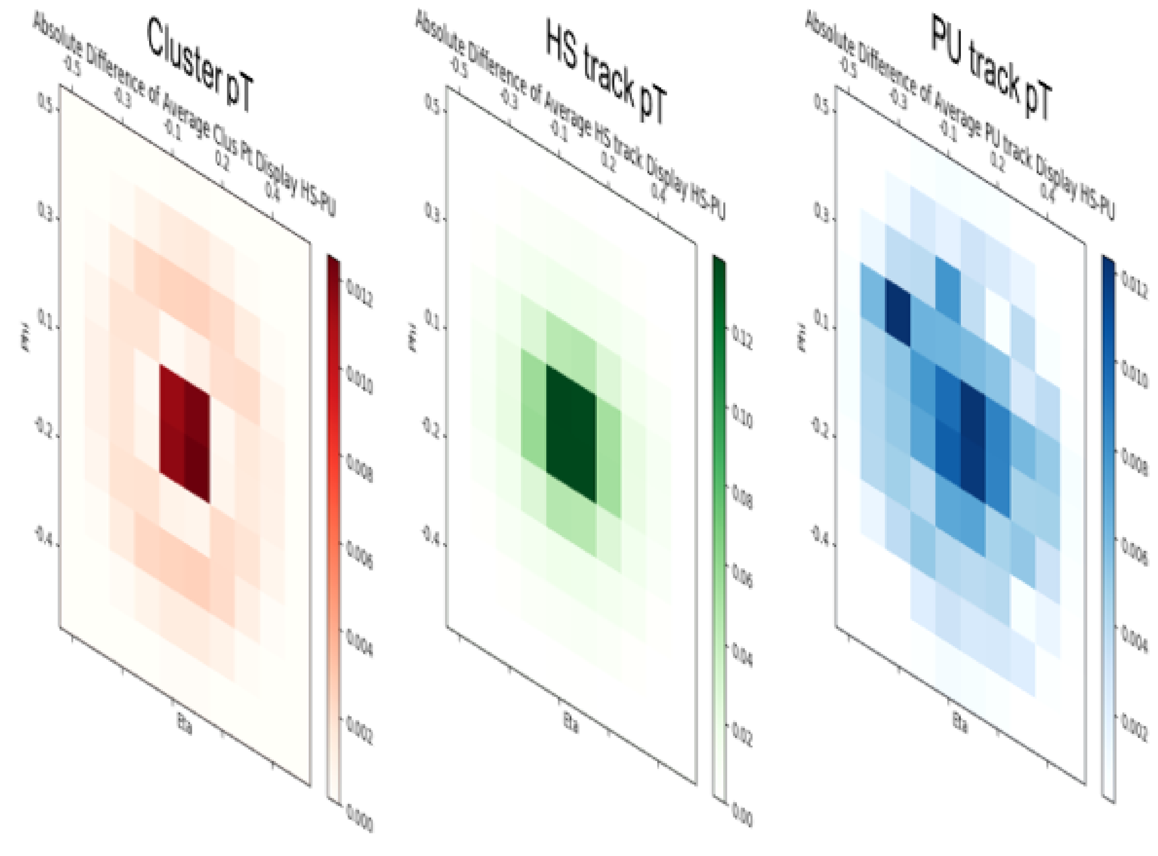
The dataset consists of $\sim 4.10^5$ detector level jets which contain:

- The true jet p_T, (η, ϕ) coordinates
- The p_T, (η, ϕ) coords for the clusters in a jet
- The p_T, (η, ϕ) coords for tracks leading into a jet, separately for HS and PU tracks
- The jet Rpt; the sum of the p_T of tracks from the PV divided by the pT of the jet

Data split: 80% training, 10% CV, 10% test. Only central jets with $|\eta| < 0.8$ are taken for uniform detector response, and with $p_T \in [20, 30]$ GeV are considered to wash out any p_T dependence. Images are formed using the cluster p_Ts, HS track p_Ts, and PU track p_Ts binned in the $\eta - \phi$ plane.



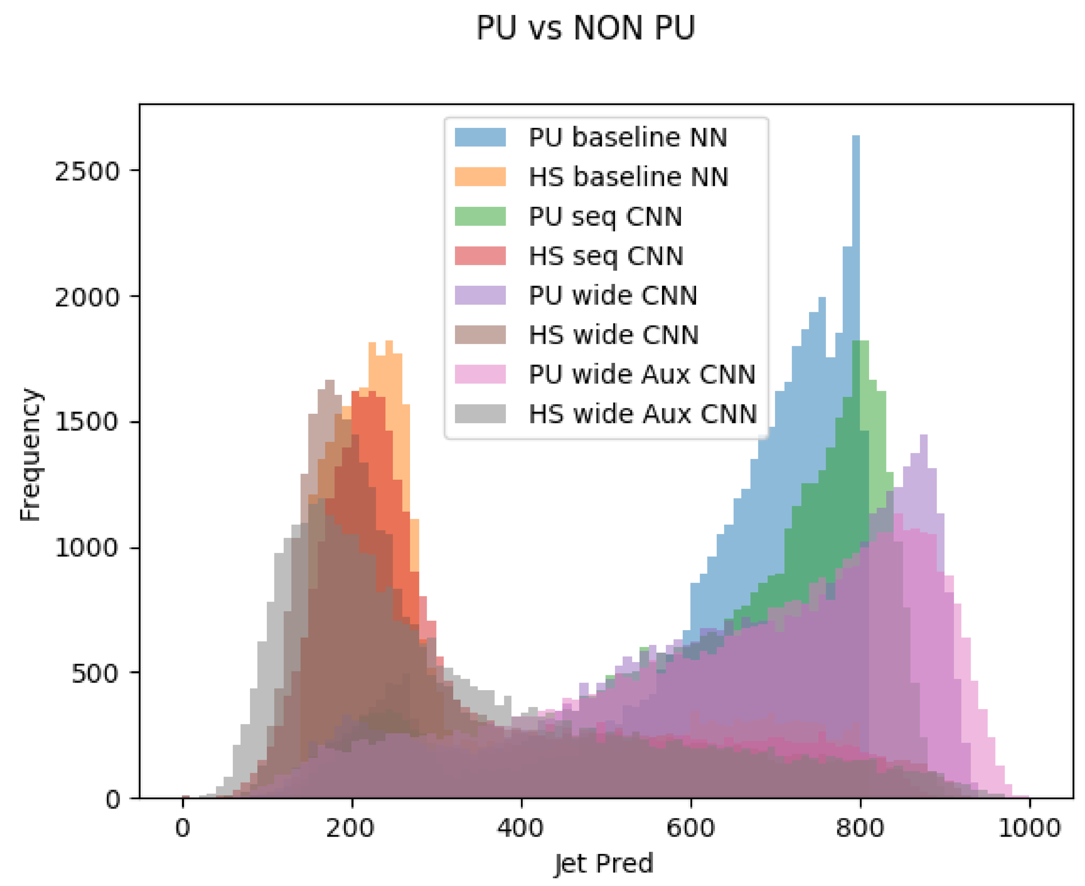
Averaged image of HS jets in the (η, ϕ) plane.

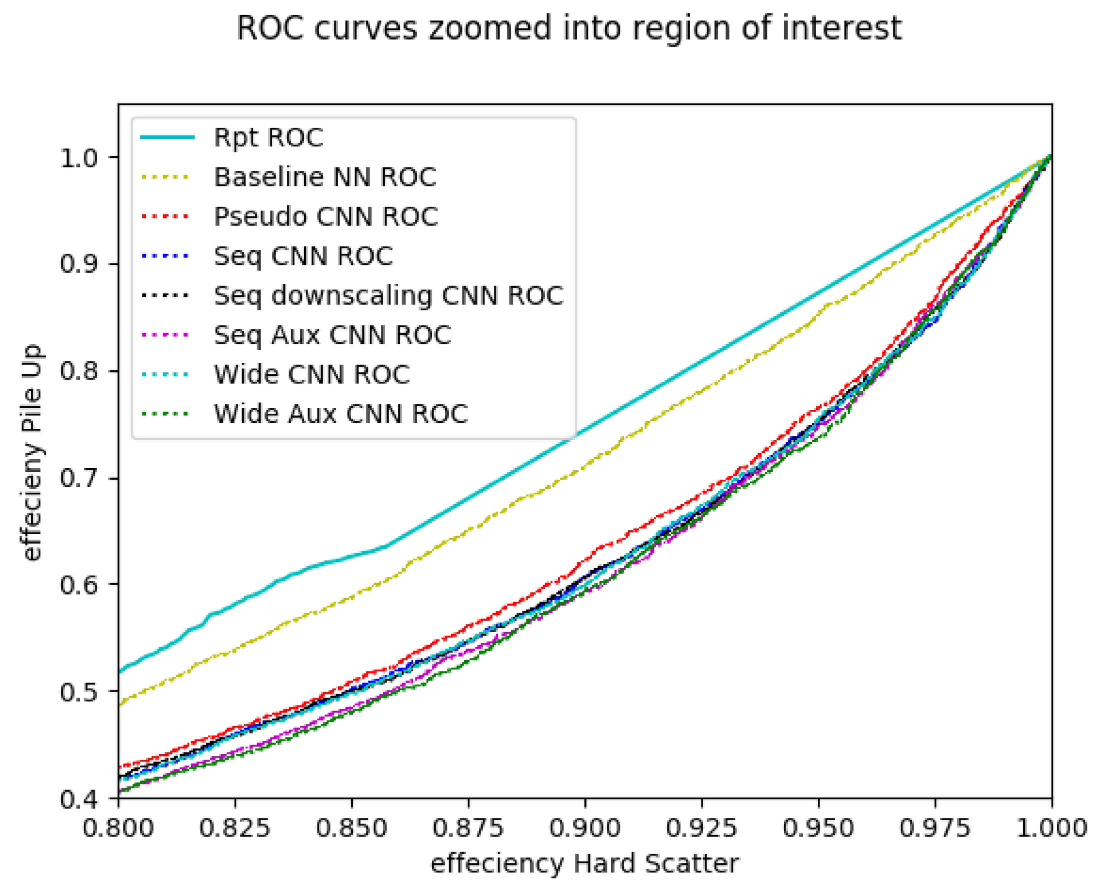


Absolute difference in the averaged HS & PU jets.

Results

Several different approaches were taken to network modeling. Sequential models were made with a couple of convolutional layers. Wide inception-inspired models were also made to combine the convolutional capacities of different kernels. Models with the jet p_T passed as an Auxiliary input were also experimented with. The results are presented below:





The model accuracies

Models	Accuracy
ATLAS standard proxy, Rpt	**0.5005**
Baseline NN using jet Rpt and p_T	**0.6994**
Pseudo CNN with full sized kernels and angular regularization	0.7013
Sequential CNN with 3x3 "Same" Conv2D followed by full sized Conv2D	0.7025
Sequential CNN with downscaling to 5x5 image followed by full sized Conv2D	0.7029
CNN with parallel convolutions of 3x3, 5x5, 10x10 filters	**0.7036**
CNN with parallel convolutions and Auxiliary Input of jet p_T	0.7072
Sequential CNN with Auxiliary Input of jet p_T	0.7073

Discussion and Future Steps

- **Trained CNNs outperform Baseline Rpt discriminant by 20 - 30 % in PU efficiency.**
- **Much of the physics analysis at ATLAS happens in the central region; These results have the potential to massively impact ATLAS Pile Up ID procedures.**
- Interest to note effectiveness of CNNs at a classification job intractable by human eyes alone.
- Accuracies suggest models with jet p_T passed as auxiliary inputs perform best.
- However learning on jet p_T makes the trained models sensitive to the p_T scale of the data, rendering it non generalizable.
- **Consequently best network: Wide Inception inspired model**, learned from different convolutions. This makes physical sense given the sparse nature of the input images.
- Detailed study of learned weights is required to understand how and why these networks outperform the current standard.
- Formal proposal to ATLAS needs to be made following a more thorough analysis.

References

[1] Menke, Sven. "Pile-Up in Jets in ATLAS " Talk given at the BOOST 2013, Flagstaff, AZ.

[2] The ATLAS Collaboration. "Tagging and suppression of pileup jets with the ATLAS detector" ATLAS-CONF-2014-018