Discriminating quark/gluon jets with deep learning

Seungjin Yang

University of Soul

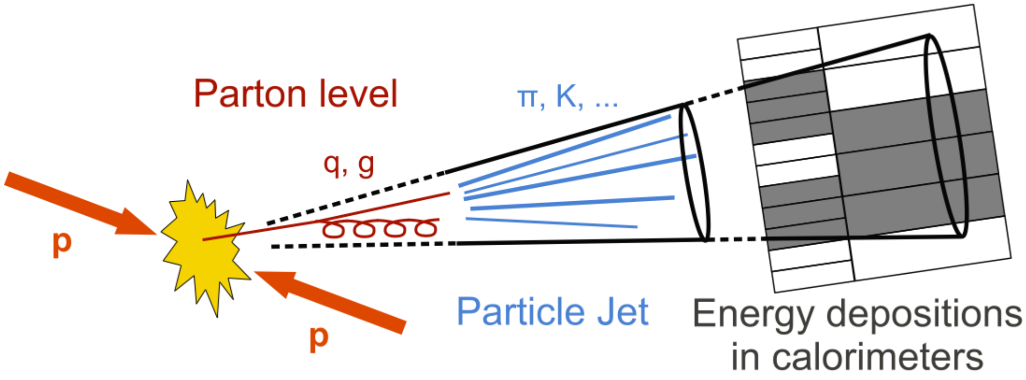
Contents

1. Introduction to jet discrimination

1.1 Jets in particle physics

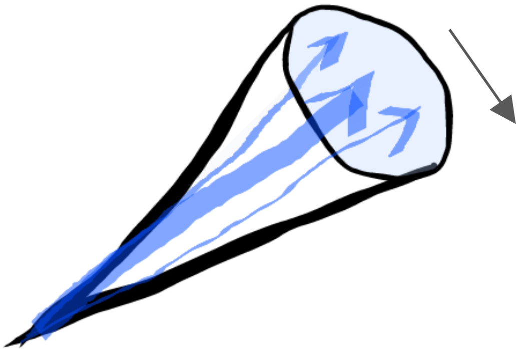
- ➢ Narrow cone of hadrons and other particles produced by the hadronization of a quark or gluon in a particle physics or heavy ion experiment.
  (from https://twiki.cern.ch/twiki/bin/view/CMSPublic/WorkBookGlossary)

- ➢ In many searches for new physics signals at the LHC, jets are initiated by light-flavour quarks, while the jets in Standard Model background processes are initiated by gluons.



1.2 Difference between quark and gluon jets

Light Quark jet



Different color factor

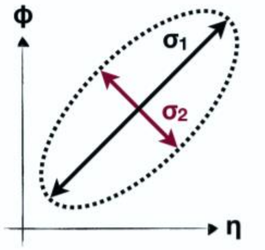
$$C_F = \frac{4}{3} < C_A = 3$$

Gluon Jet

- ➢ **Gluon jets** produces **more constituents** with **more uniform energy** fragmentation. Therefore those are **wider**.
- ➢ **Quark jets** produces **fewer constituents** and hard constituents carry a **significant fraction of the energy** fragmentation. So those are **narrow**.

➢ **Angular spread** (Minor axis)

1.3 Discriminating variables

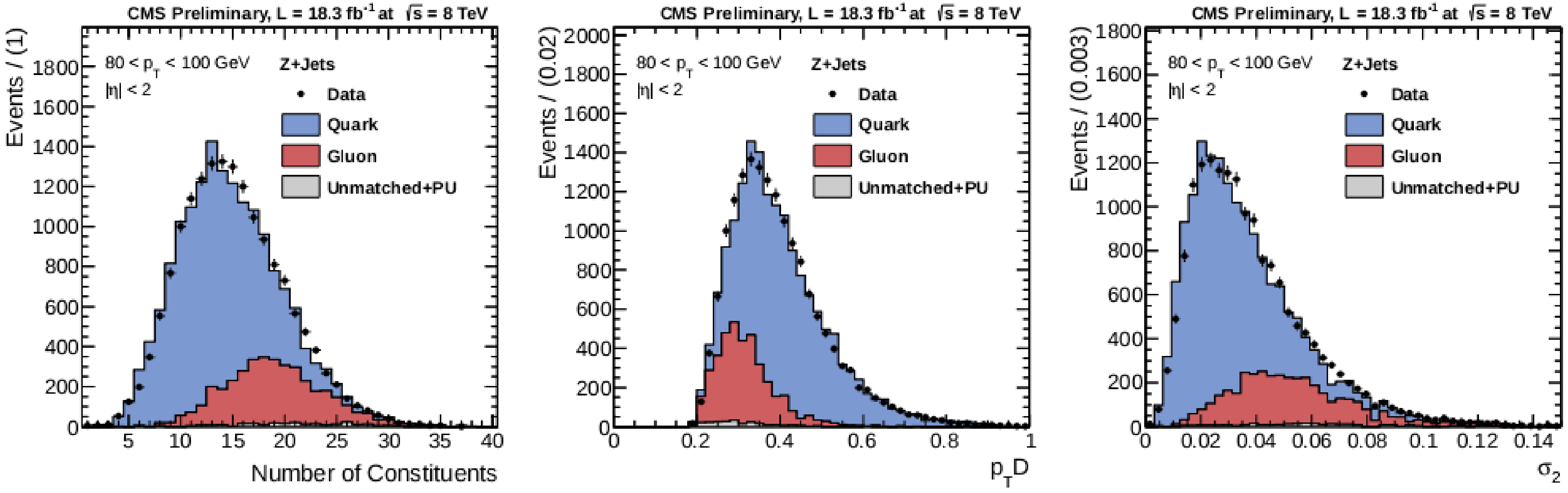


➢ **Multiplicity**

- The number of particles constituting the jet

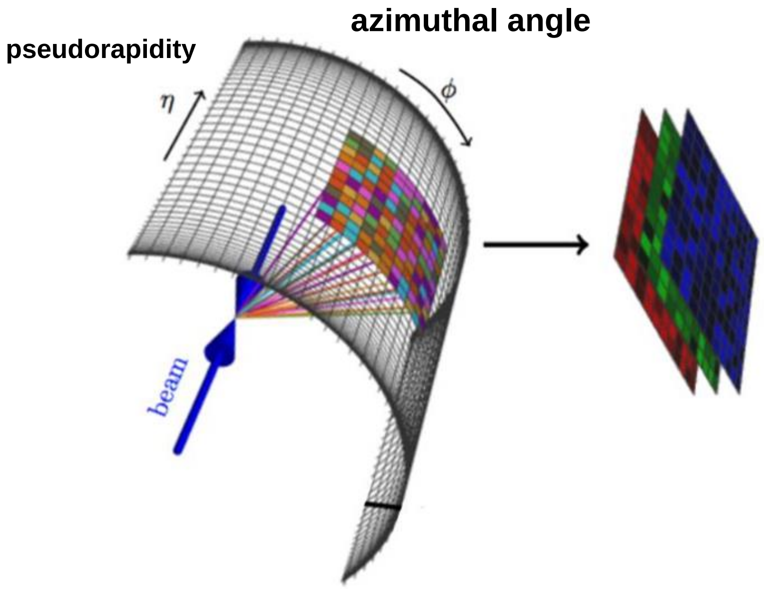
➢ **Jet energy sharing variable**

$$p_T D = \frac{\sqrt{\sum_i \; p_{T,\;i}^2}}{\sum_i \; p_{T,\;i}}$$



(ref. Tom Cornelis, for the CMS Collaboration. Quark-gluon Jet Discrimination At CMS), arXiv:1409.3072 [hep-ex]

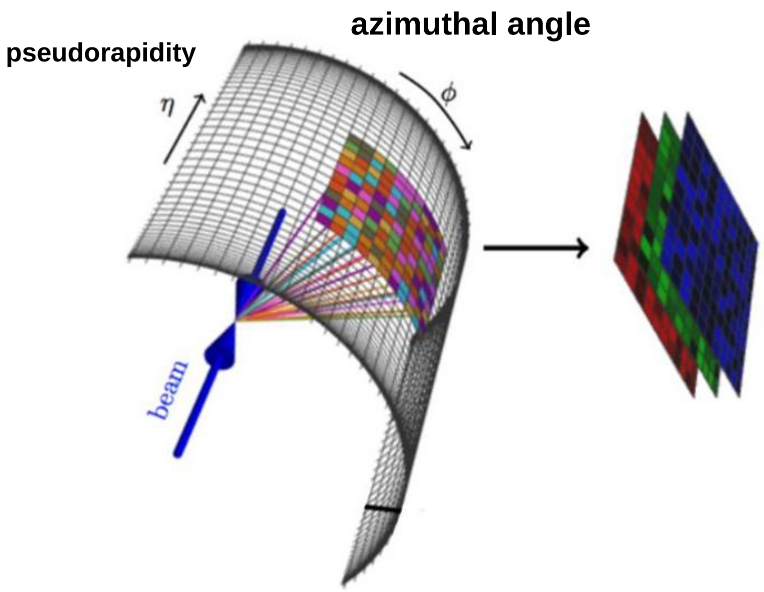
1.3 Jet images



- The idea behind jet images is to treat the energy deposits in a calorimeter as intensities in a 2D image. (ref. Patrick T. Komiske, Eric M. Metodiev, and Matthew D. Schwartz, *Deep learning in color: towards automated quark/gluon jet discrimination,* arXiv:1612.01551 [hep-ph])

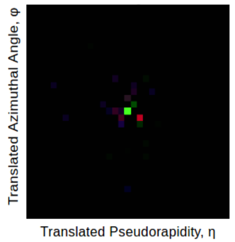
- We introduce a novel approach to jet tagging and classification through the use of techniques **inspired by computer vision.** (ref. Josh Cogan, Michael Kagan, Emanuel Strauss, Ariel Schwartzman, *Jet-Images: Computer Vision Inspired Techniques for Jet Tagging* arXiv:1407.5675 [hep-ph])

1.3 Jet images

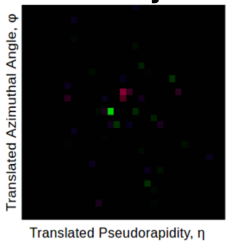


➢ Jet images

Down quark jet



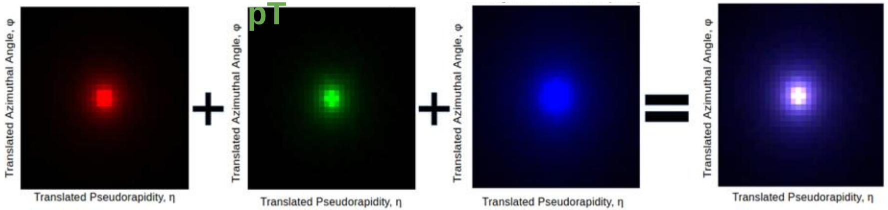
Gluon jet



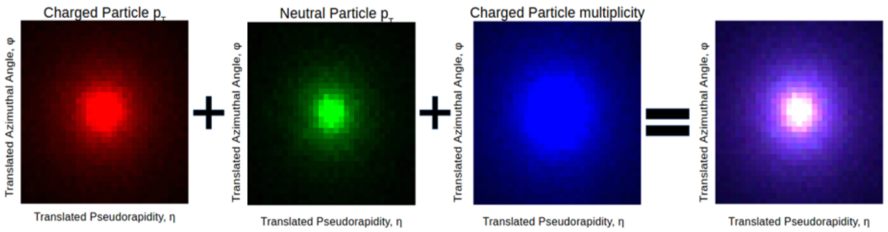
➢ Summed jet images (220K)

- Quark jets

Charged pT **Neutral pT** **Charged multiplicity**

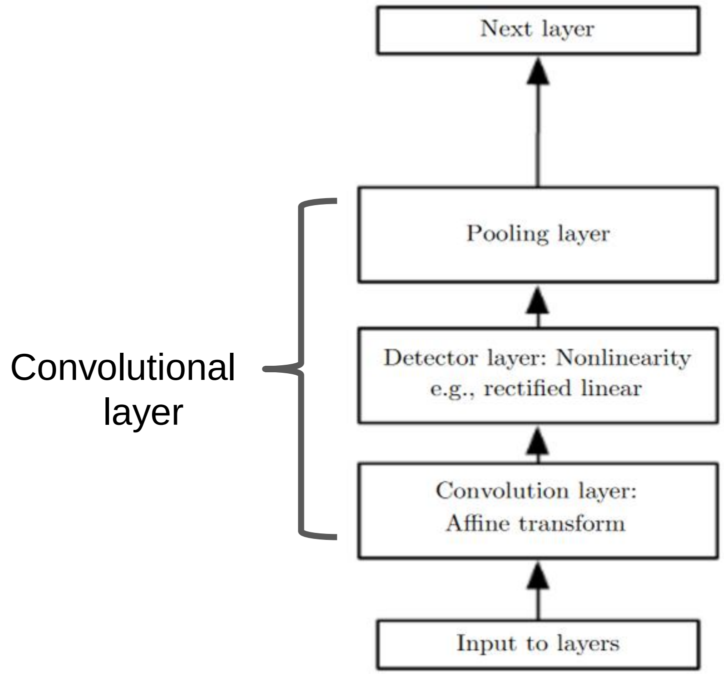


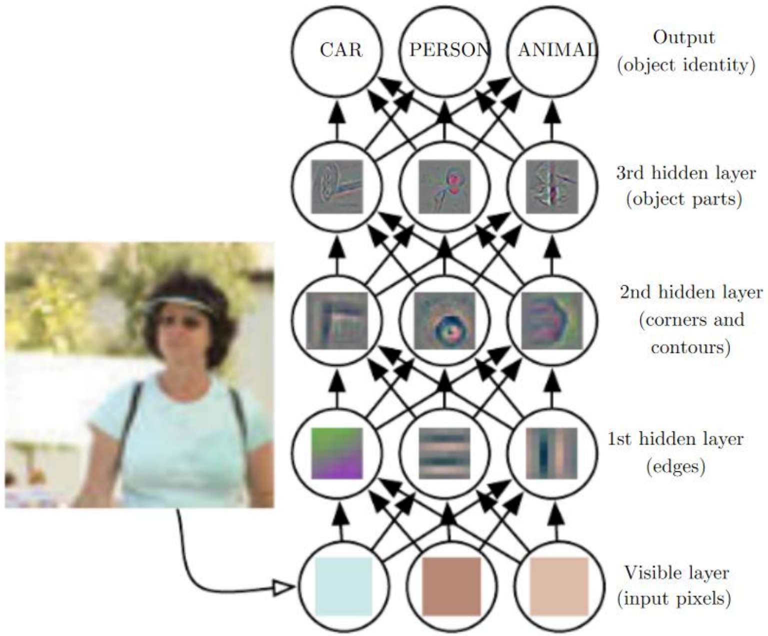
- Gluon jets



2. Image Classification with Deep learning

2.1 Convolutional Neural Networks



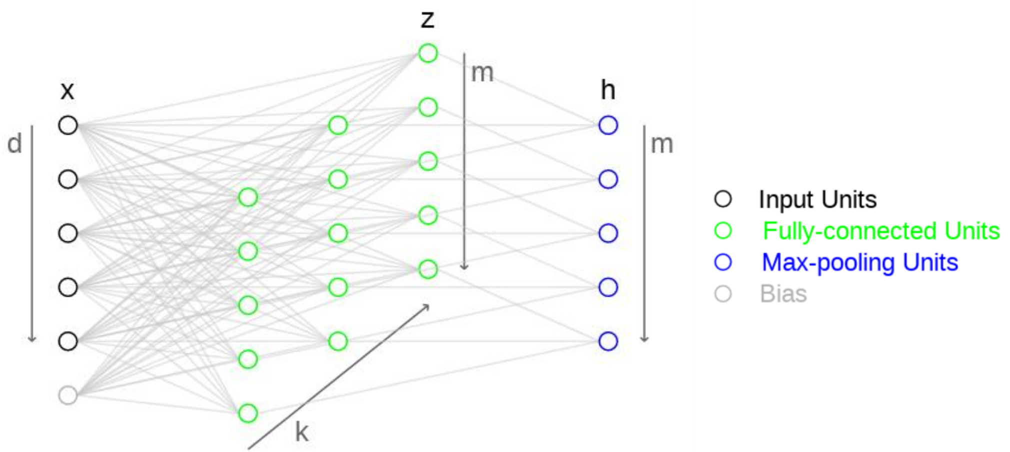


ref. Ian Goodfellow and Yoshua Bengio and Aaron Courvill, *Deep learning*

2.2 Maxout unit

➢ $z_{ij} = x^T W_{\dots ij} + b_{ij}$

$$h_i(x) = \max_{j \in [1,k]} z_{ij}$$

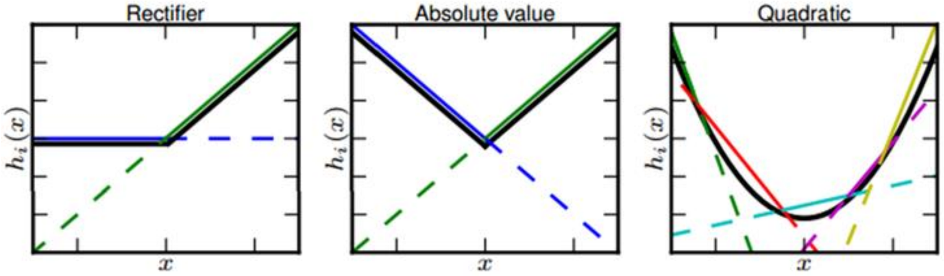


(ref. http://www.simon-hohberg.de/2015/07/19/maxout.html)

➢ Generalization of rectified linear units.

➢ **Learning the activation function**

➢ With large enough k, a maxout unit can learn to approximate any convex function with arbitrary fidelity.



ref. Ian Goodfellow and Yoshua Bengio and Aaron Courvill, *Deep learning*

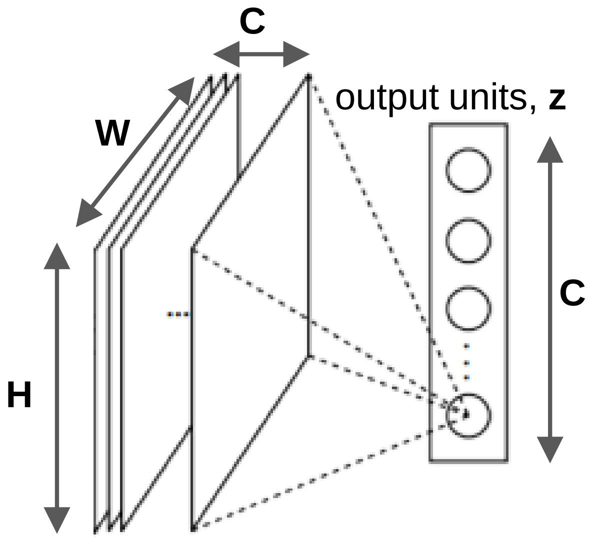
2.3 Global Average Pooling

- ➢ $z_i = \sum_{j=1}^{H} \sum_{k=1}^{W} x_{ijk}$

- ➢ Global average pooling operation reports the average sum within all parameters in a feature map.

- ➢ Introduced to replace the traditional fully connected layers in CNN

- ➢ **Global average pooling is itself a structural regularizer**.

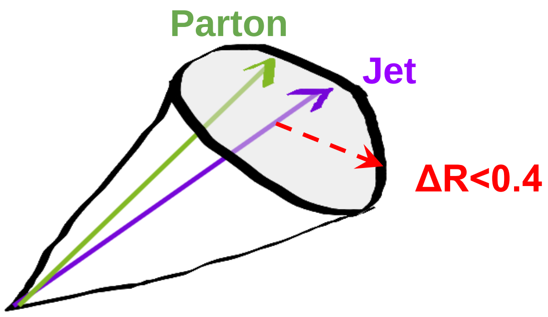


ref. Min Lin, Qiang Chen, Shuicheng Yan, *Network In Network,* arXiv:1312.4400 [cs.NE]

3. Experimental Setup

3.1 Dataset

- ➢ Samples
  - QCD event generated by Pythia8 at LO using the CMSSW for the detector simulation.
  - jet eta cut: |η| < 2.4, jet pT < 20 GeV
- ➢ Jet image
  - 3 channels: **charged particle pT**, **neutral particle pT** and **charged particle multiplicity**
  - η, $\varphi \in$ (-0.4, 0.4), where (η, φ)=(0,0) is the jet center
  - 33 x 33 pixels.
  - NCHW format (cuDNN default)
- ➢ Jet matching
  - dR(jet, hard parton) < 0.4



3.2 Architecture of the model

Layer	Output size	#weight
Conv3-64 x2 Max Pool	**[64, 17, 17]**	**1152** **(=3x3x64x2)**
Conv3-128 x2 Max Pool	**[128, 9, 9]**	**2304**
Conv3-256 x2 Max Pool	**[256, 5, 5]**	**4608**
Conv3-512 x2	**[512, 3, 3]**	**9216**
Global Average Pool	**[512]**	**-**
Maxout (k=10) Dropout (p=0.5)	**[128]**	**656,640** **(=513x128*10)**
MaxOut (k=10) Dropout (p=0.5)	**[2]**	**2580**

➢ **How to read**

- Conv<kernel size>-<# output channels>
- Output size: [CHW] or [D]

➢ **The number of weight**

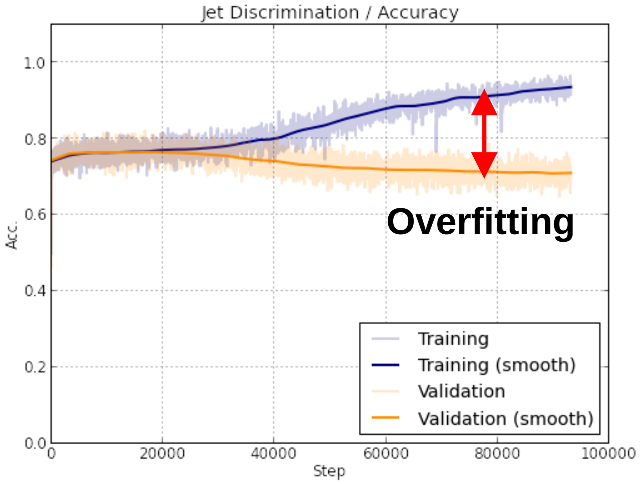
- 676K weight in 10 weight layer
- Most weights are in 1st maxout unit… (97%)

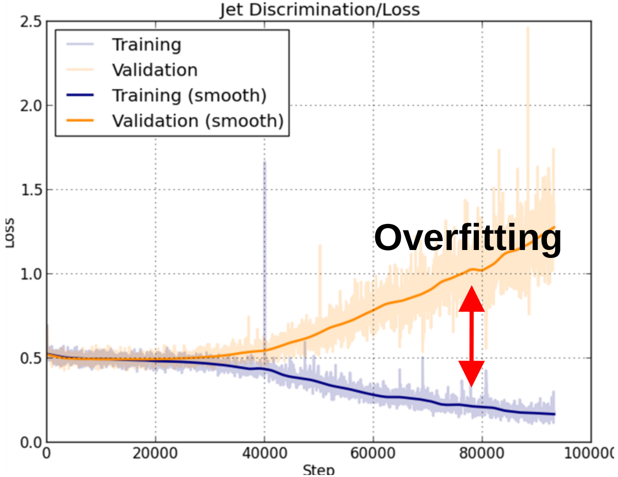
3.3 Setup details

Dataset	**Training data 775K** + Validation data 258K + Test data 258K
Loss	(binary) **cross entropy**
Optimization algorithm	**Adam** learning rate = 0.0015 1st / 2rd momentum decay = 0.9 / 0.999
Weight Initialization	**Xavier initialization**
Batch size	**500 example**
#Epochs	**100** (1 epoch ~ 1550 training step)
Framework	**TensorFlow** (API r1.3)

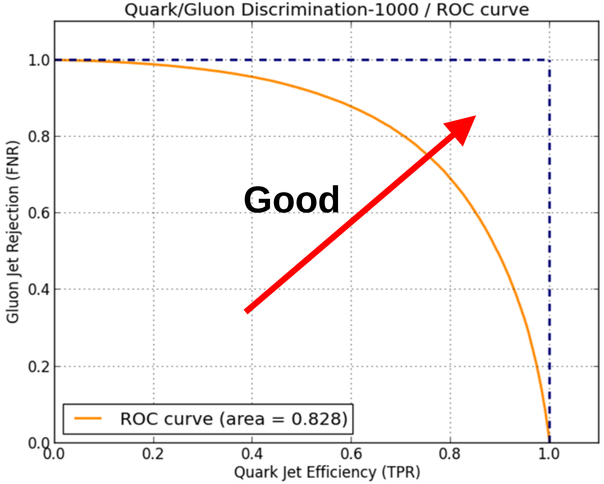
4. Results

4.1 Metrics: acc., loss and ROC curve



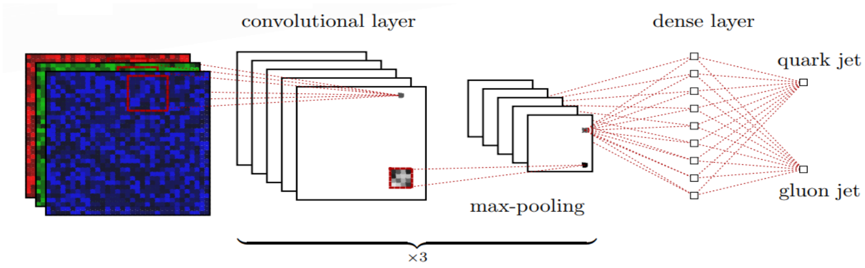


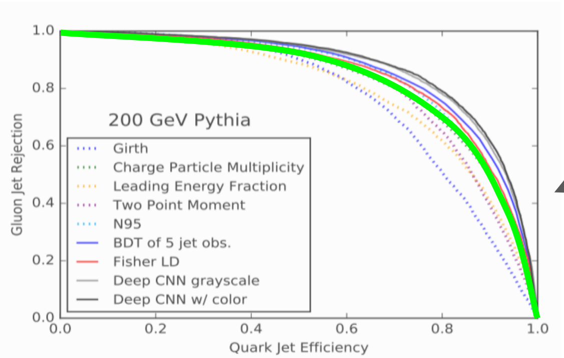
smoothing by local linear regression



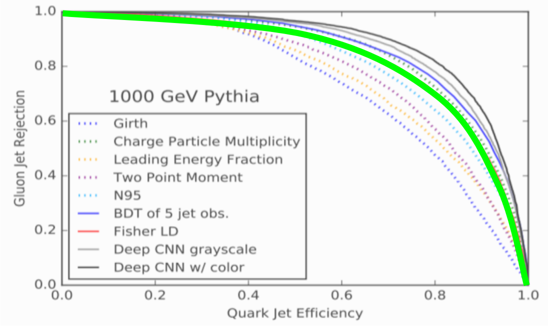
- ➢ Training took about 4.5 hours for 100 epochs on a single GPU
- ➢ Overfitting occurred after about 20000 steps.
- ➢ The best auc is 0.838 on average.

4.2 Comparison with previous research





Green one is our ROC curve.

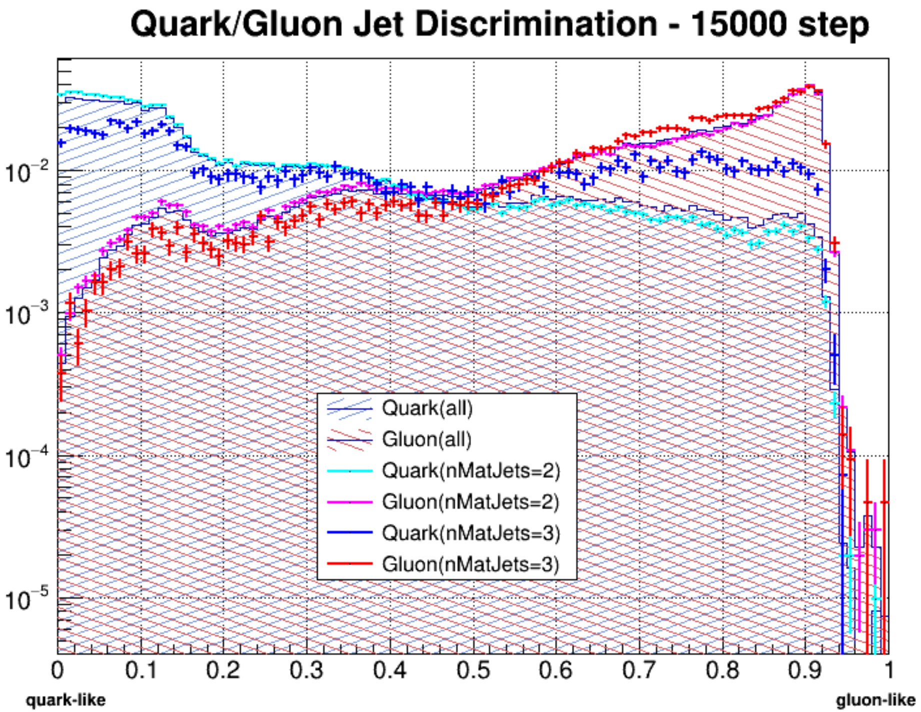


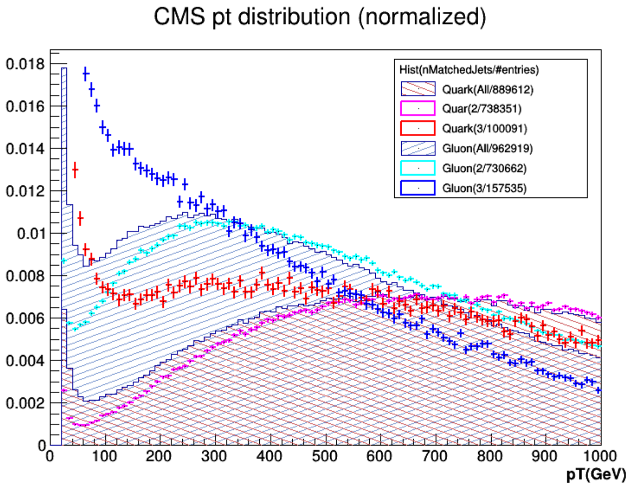
➢ Dataset in previous research
- clustering final state hadron (no detector simulation).
- jet eta cut: $|\eta| < 2.5$, jet pT cut: pT > 100 GeV
- Preprocessing: normalization + standardization
- pT binning: 200-220GeV, 1000-1100GeV

➢ Model
- 3 Conv. layer + 1 Dense layer

ref. Patrick T. Komiske, Eric M. Metodiev, and Matthew D. Schwartz, *Deep learning in color: towards automated quark/gluon jet discrimination*, arXiv:1612.01551 [hep-ph]

4.3 Histogram of NN output with # matched jets





- Due to the ambiguity of jet matching, it is expected that gluon jets are mislabeled with quark jets.

5. Conclusion

- We trained deep neural networks to discriminate quark / gluon jets.
- We need to investigate the quark / gluon jet definition and event level cleaning.
- We expect an improved model thorough
  - many normalization methods,
  - preprocessing method with minimization information loss