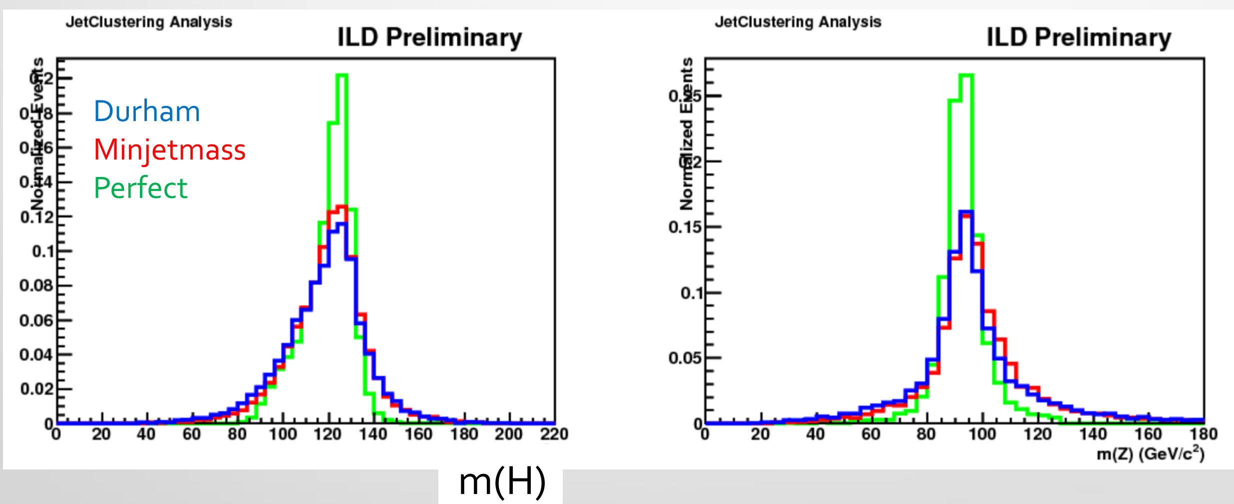
Construct
Deep Jet Clustering

Masakazu Kurata

07/01/2020

Introduction

- Jet clustering is one of the main key to obtain better physics results
  - Physics results are strongly limited by mis-clustering
  - To obtain correct jets leads to improve the mass resolution of the resonances
- Present jet clustering is far from good tool for reconstructing jets
  - e.g. Higgs self-coupling@500GeV(ZHH): ～40% improvement if perfect!

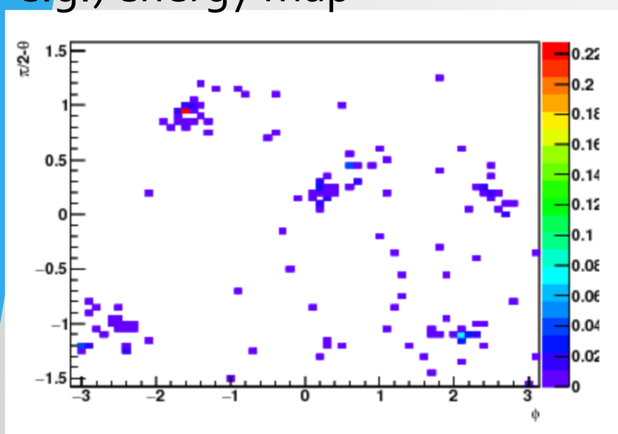


- Even at 250GeV, clustering is very important
  - Separation of ZH/ZZ/WW in hadronic events

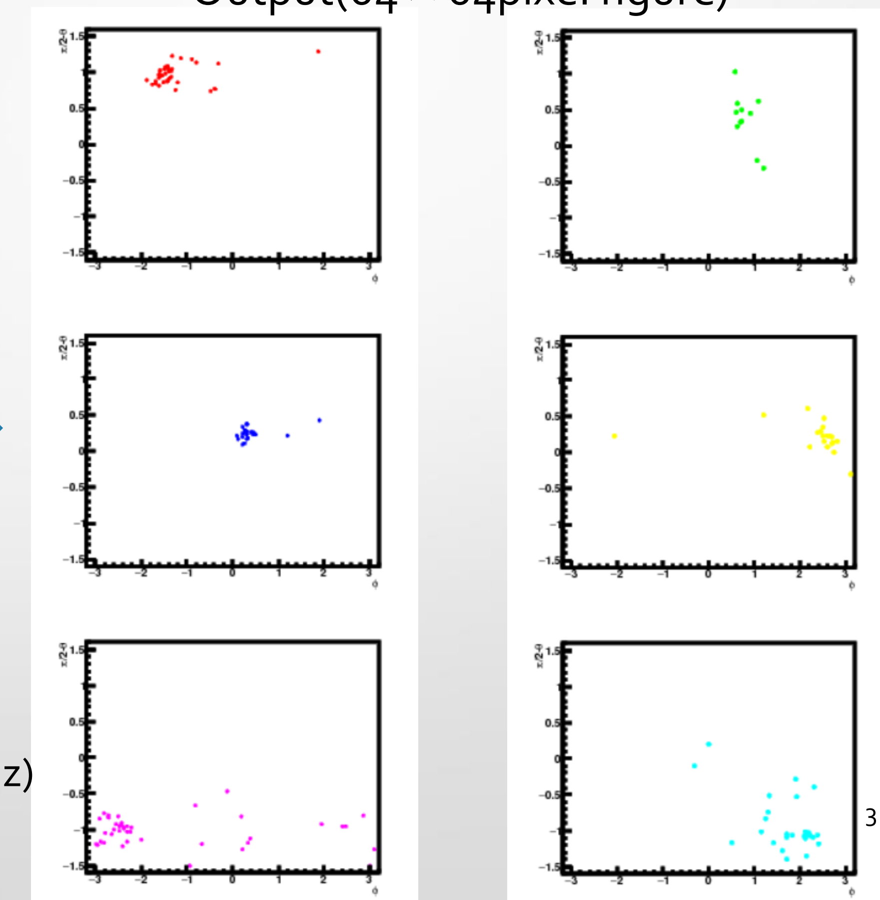
Trial

- Use Keras & tensorflow backend
- Using a certain map(s) of each event, estimate color of each track
  - Do not consider color-singlet state

Input(64×64pixel figure)
e.g.) energy map

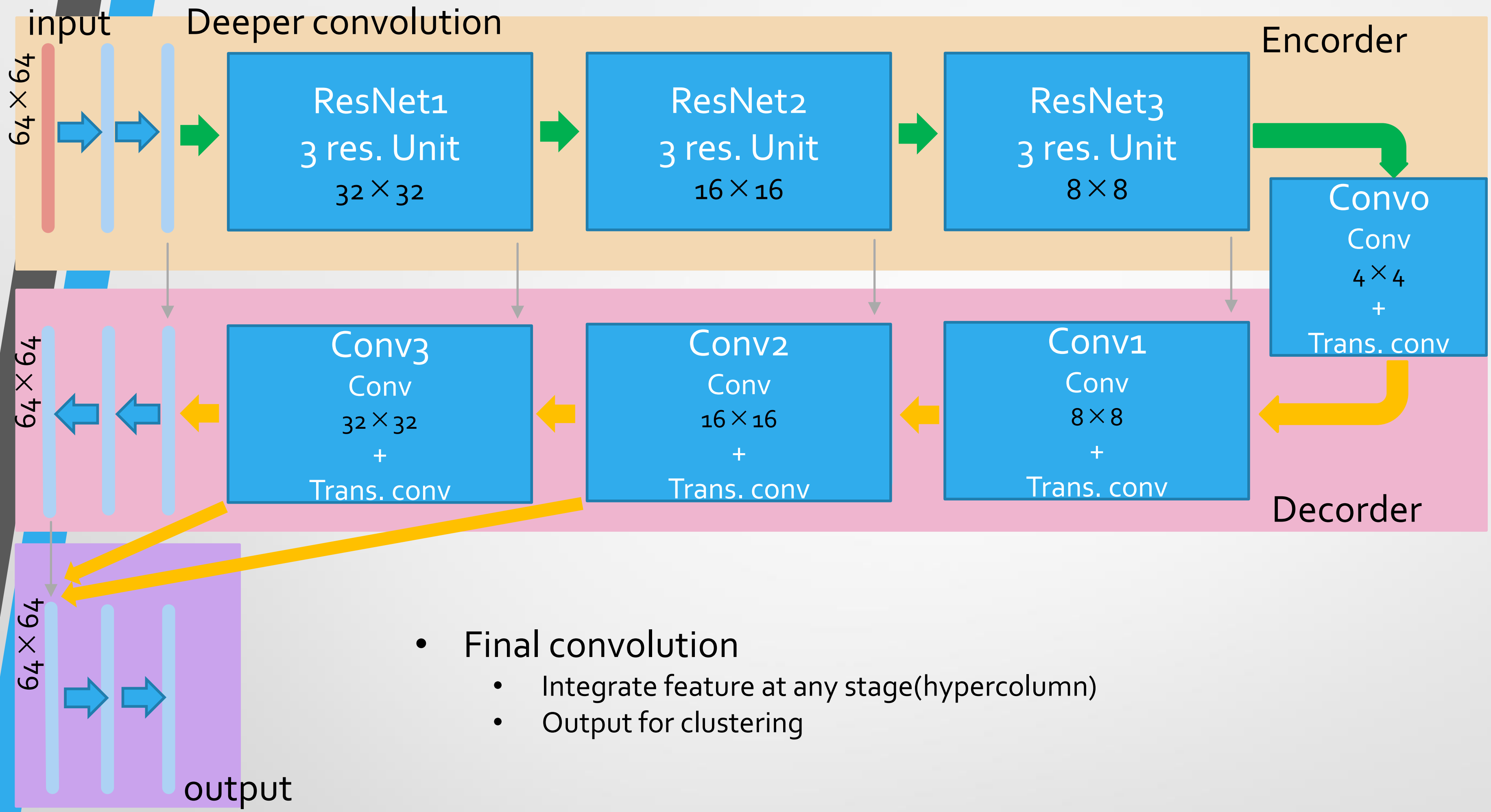


Output(64×64pixel figure)



9 input images
Energy map
Charge map
Dosig map + direction vector(x, y, z)
Zosig map
Ecal map
Hcal map

Network Architechture



- Final convolution
  - Integrate feature at any stage(hypercolumn)
  - Output for clustering

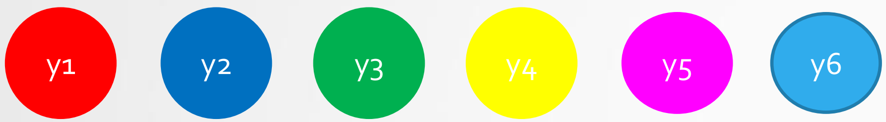
- Encorder
  - Extract global & higher order feature
  - Downsample to make network robust for distortion & shift effect
  - Lost position information

- Decorder
  - Expand obtained feature to local
  - Upsample to recover position information
  - Merge encoder nodes to get precise position information

Pseudo-labelling

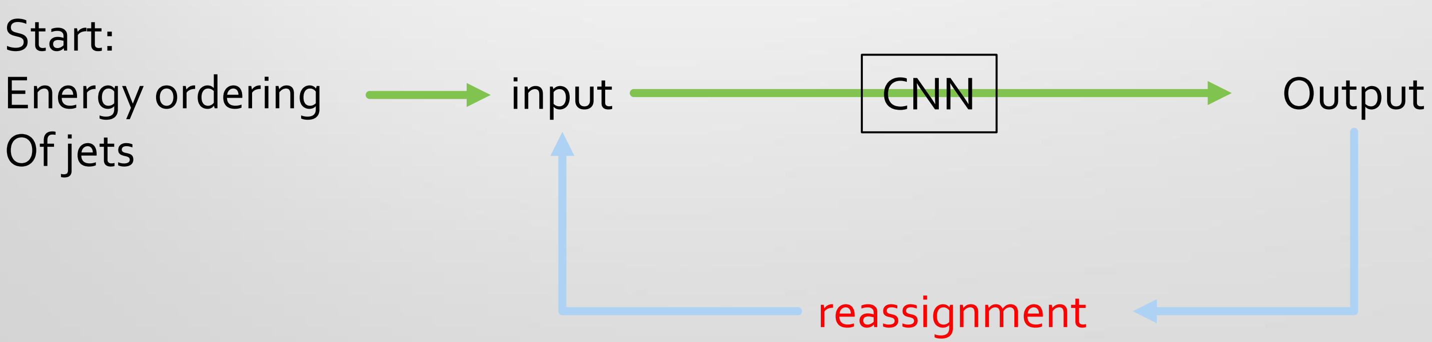
- Output: inference of the probability of the color to be assigned
  - $\sum y_i = 1.0$

output

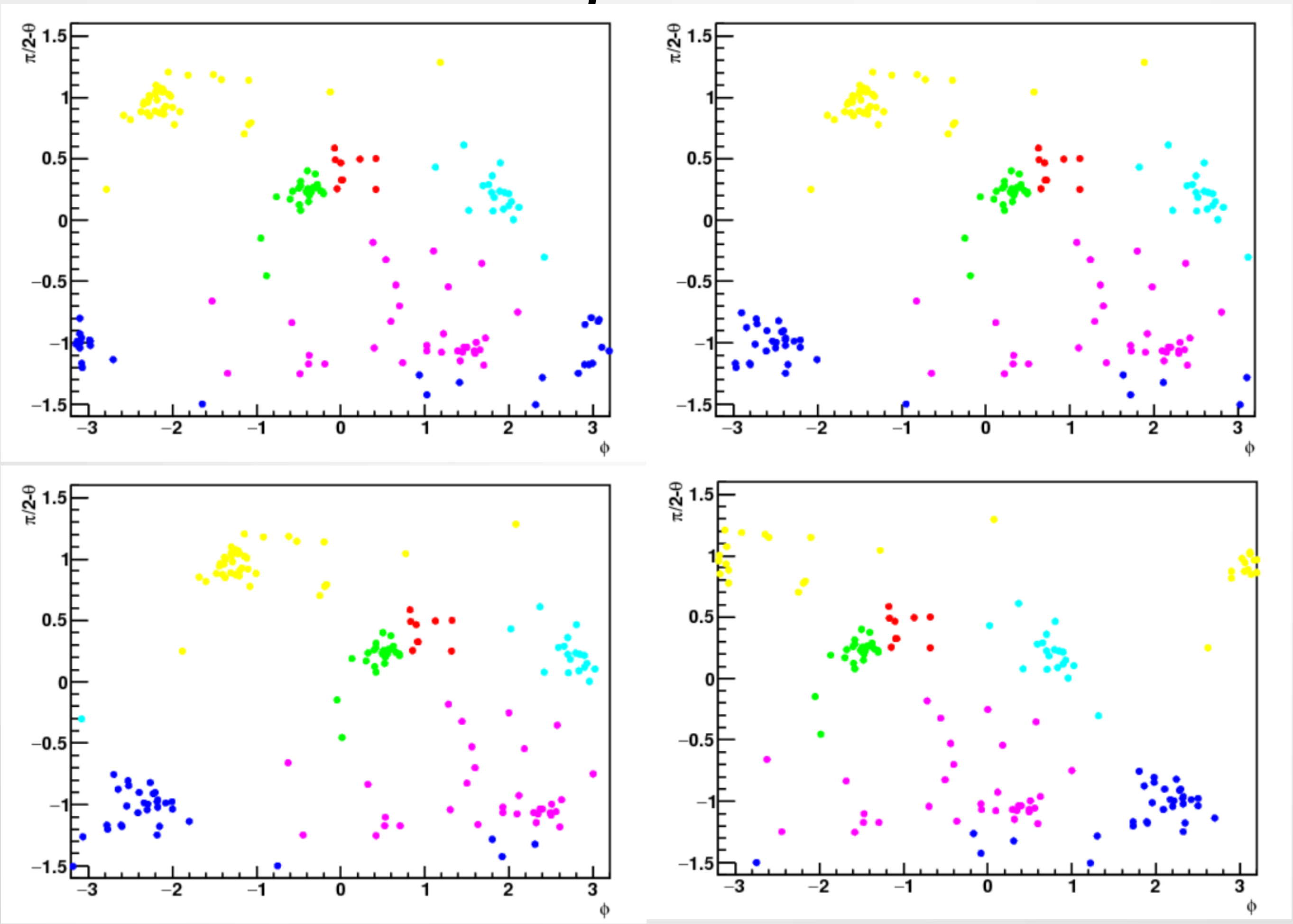


+ no particle

- The combination of color assignments is arbitrary, so assign them so that the loss function is minimized.
  - Using preliminary results after a training, re-assign the color combination
  - Minimize cross entropy $L = \frac{1}{n} \sum y_i \log p_j$



Data Augmentation



- Random shift for x axis
  - Considering periodic condition of ϕ angle (f(Φ+2π) = f(Φ))
  - To suppress over fitting
- Add random y-flip (I think not good from physics point of view, but suppress over-fitting is important)

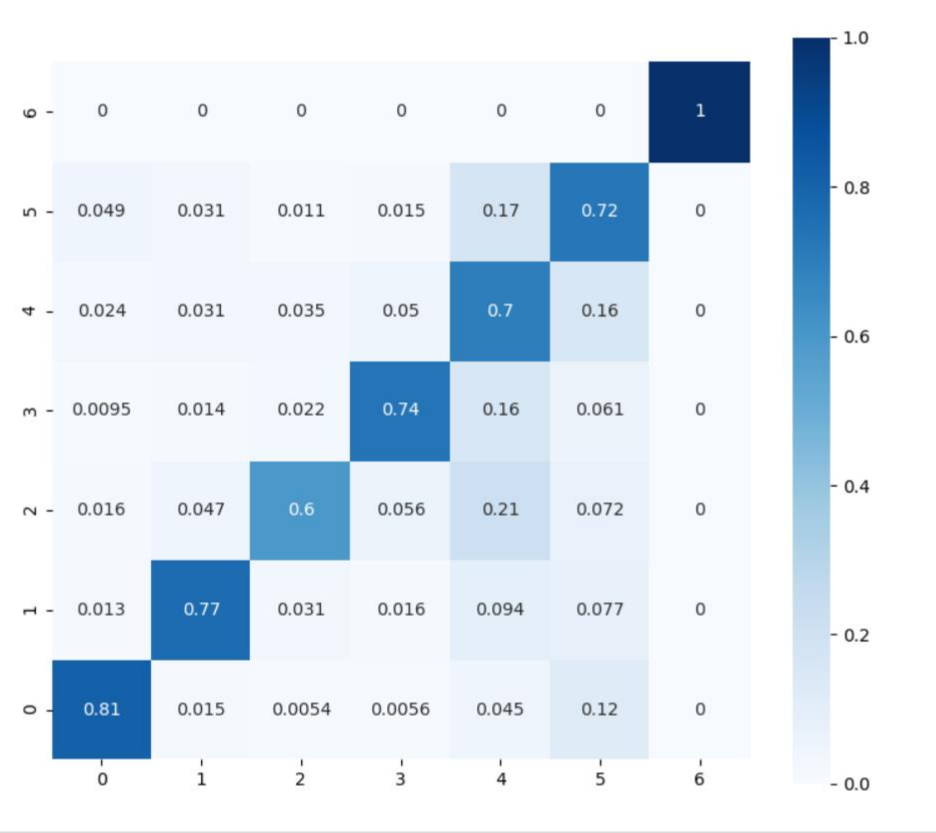
status

- Use ZHH→(qq)(bb)(bb): 6jets clustering
  - q: uds
- Use 40000 events for training(36000 train, 4000 validation)
- Don't consider color singlet state for network training
  - But, as mentioned, use the freedom of color singlet state: Data cleansing for better performance
- Input: 6 + 3 images output: 6 + 1 images

Preliminary results

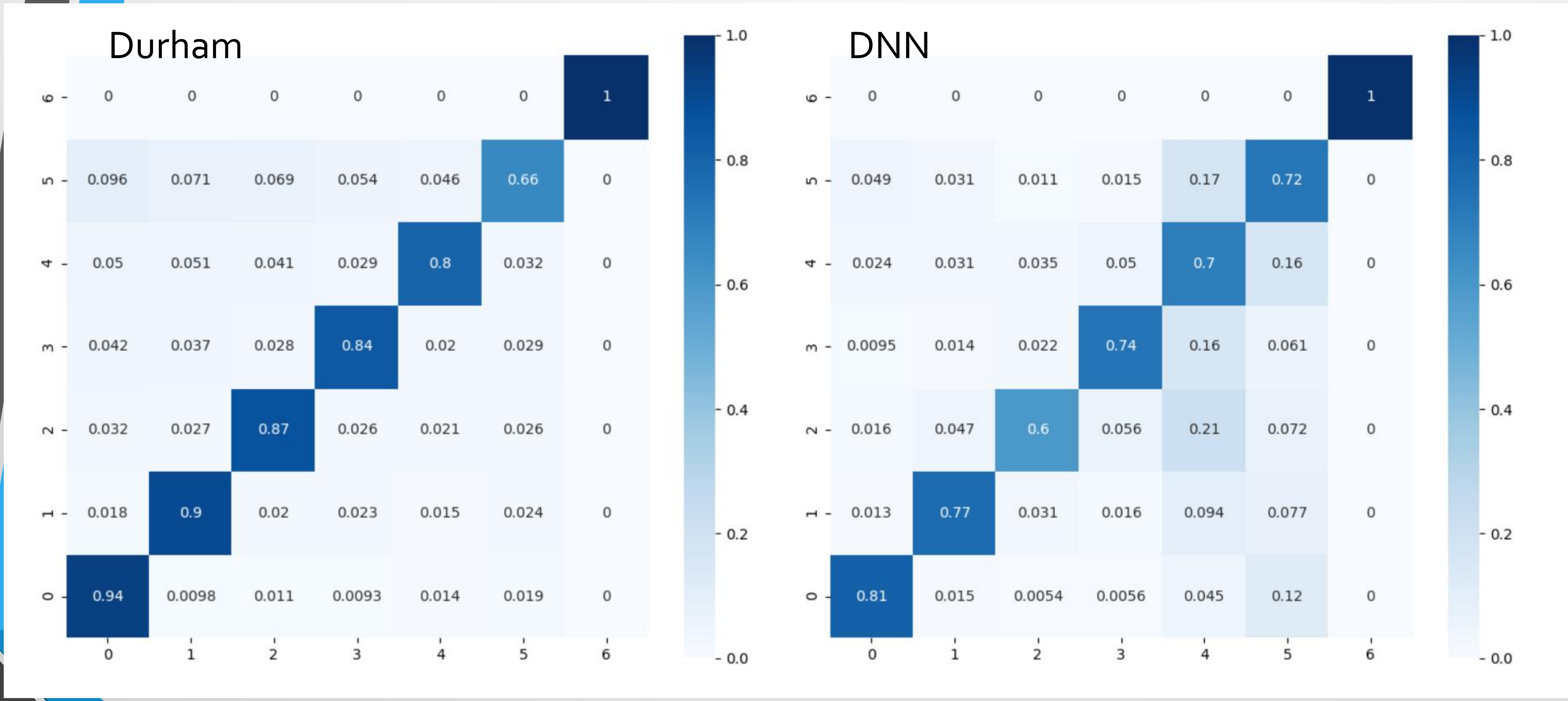
- Confusion matrix
  - Test sample: 1000 events
  - Many particles located in jet 5 & 6
    - Looks low energy particles
- Need to check resolution of Higgs mass
  - Need to improve

Keras	training	validation	test
Accuracy	0.9821	0.9846	0.9854

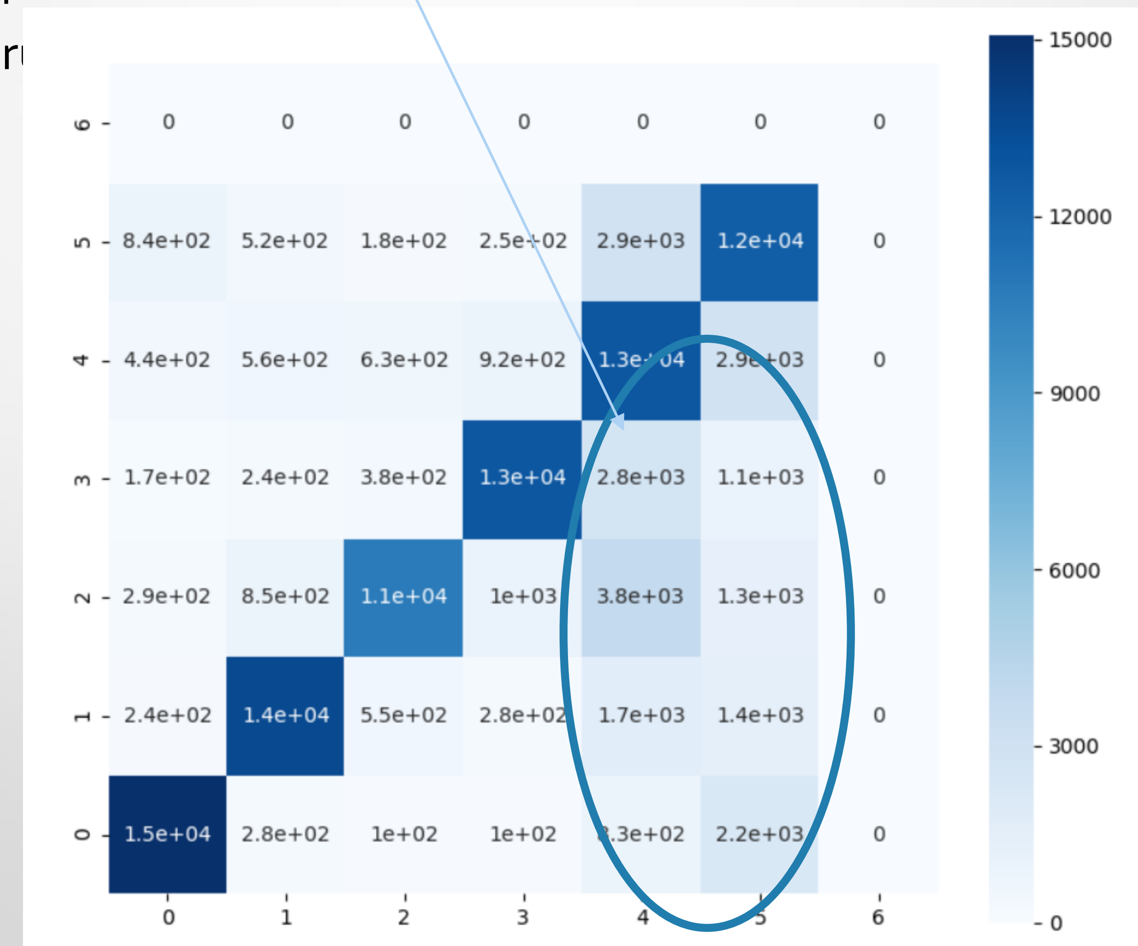


Comparison with durham

- Still Durham is better…



- These particles are not trained?
  - Low energy particles
  - “Sparse” cluster particles
  - will try to train well such kind of particles
- Other method
  - More follow Durham procedure
  - Keras is easy to constr



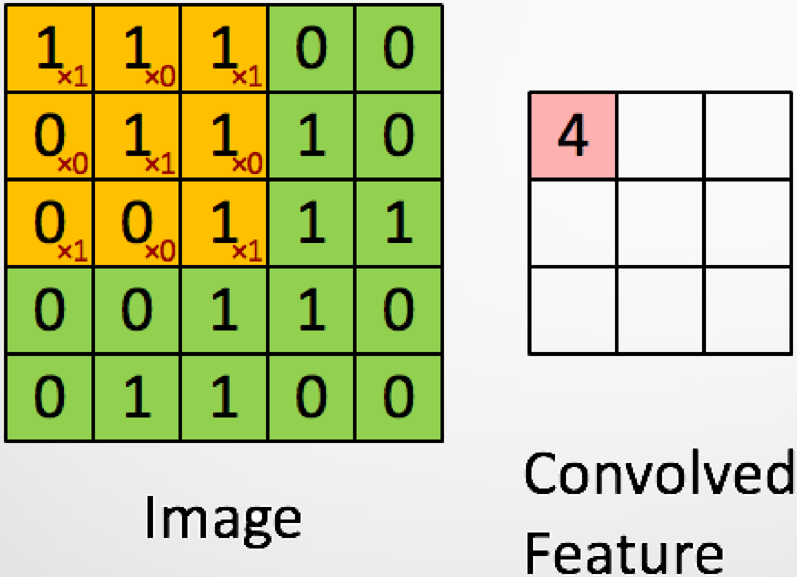
backups

Basics: convolution

- Convolution: Apply the filters to extract the feature
  - Sum of the product of each pixel and filter weights:

    $y_{kl} = \sum_{i,j} w_{ij} \cdot x_{(k+i)(l+j)}\ (+b)$

  - Slide filters over all the pixels



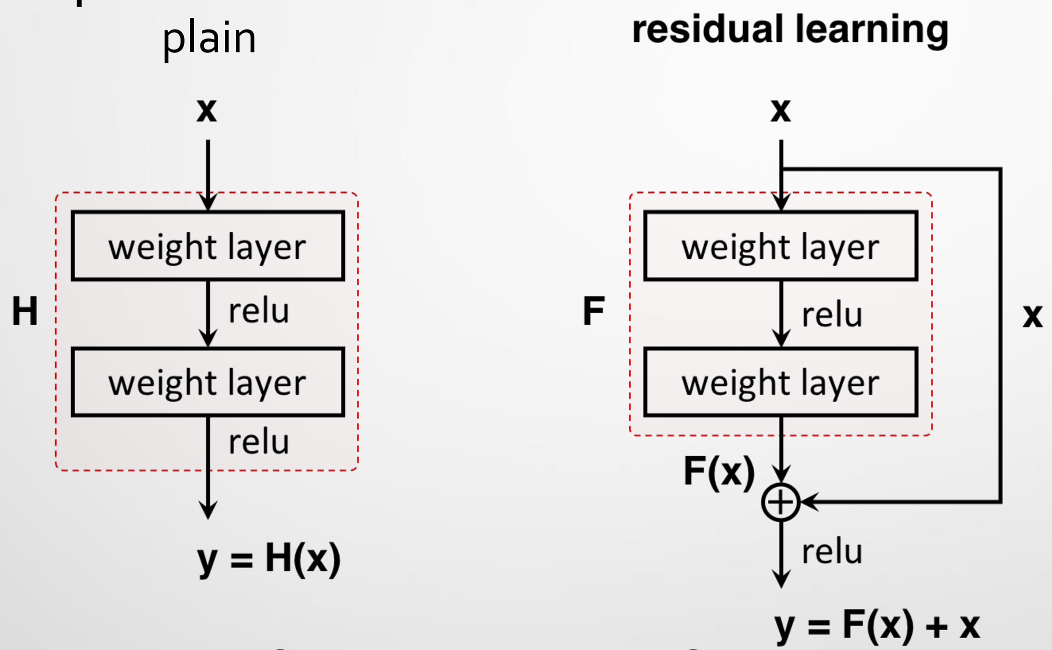
  - Filters are parameters: CNN can obtain them automatically
- After the convolutional operation, apply non-linear transform

  $z_{kl} = \sigma(y_{kl})$

  - “Non-linear” is important to get good expression
- Stack these operations

Basics: Residual convolution

- Stream is divided into 2 paths:
  - Path with convolution
  - Path without any operation
- Sum up these 2 path in downstream



- Can learn "Residuals" of previous layer features
- Can construct very deep network
  - >100 layers can be constructed
  - Deeper will be better performance

Basics: Transposed convolution

- Reverse operation of convolution
  - After adding padding, do convolution
  - Use for upsampling