**Summary of Image Representation For Pile Up Mitigation**

**Introduction**

3D image based approach for pileup mitigation encode each event as a 3D image. In such representation, each layer is thought to be a shell of a sphere in spherical coordinate system. The points on these shells either correspond to a pileup or not pileup(Fig 1). The output of this approach will be the images in which pileups would have been cleared from the respective layers (Fig 2).

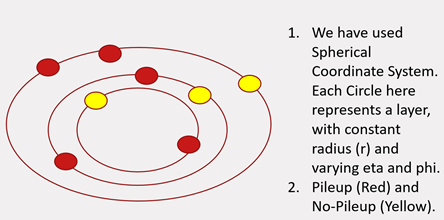


Fig 1 : Input to the model - each event is transformed into an image, and acts as a single input data-point for the model

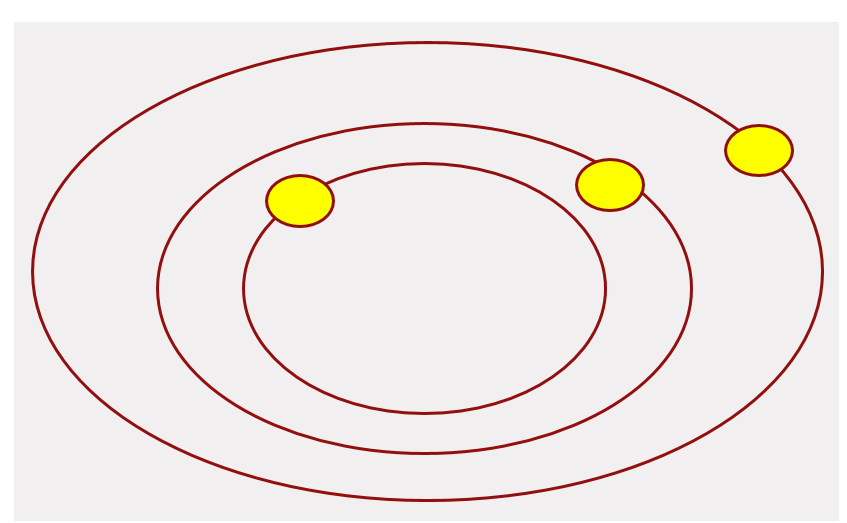


Fig 2 : Output from the model – has same dimension as the input

**Literature survey:**

The team referred following publications during the course of image based representation.

* Pileup Mitigation with Machine Learning (PUMML) (<https://arxiv.org/abs/1707.08600>)
* Deep learning in color: towards automated quark/gluon jet discrimination (<https://arxiv.org/abs/1612.01551>)
* Jet-Images – Computer Vision Inspired Techniques for Jet Tagging (<https://arxiv.org/abs/1407.5675>)
* Jet-Images – Deep Learning Edition (<https://arxiv.org/abs/1511.05190>)
* Pileup mitigation at the Large Hadron Collider with Graph Neural Networks (<https://arxiv.org/abs/1810.07988>)

**Data Analysis:**

The following section describes the data analysis conducted on 31st Oct 2019 dataset.

Represent each event as (50\*200\*200) array with pixel intensity at each location is given by weighted E values (nHits \*E): We encode the data as a 3D image of size 50\*200\*200 (layer\*x\*y).

* Total Number of events in the dataset : 100
* Global Values across all the events:
  + Min X, Max X : 48.490150451660156, 181.29637145996094
  + Min Y, Max Y : -77.65599822998047, 106.15882110595703
  + Min Layer, Max Layer : 1 , 50
* Discretize the X and Y into 200 \* 200 pixels :
  + delta\_x : 0.66 (MaxX – MinX ) /200
  + delta\_y: 0.92 (MaxY – MinY ) /200
* Sparsity of Data :
  + Total possible Nonzero pixels in an event: 2M (50\*200\*200 )
  + On an average only 244 Nonzero pixels in an event (Fig 3)
  + Each image is at least 99.96 % sparse (Fig 4)

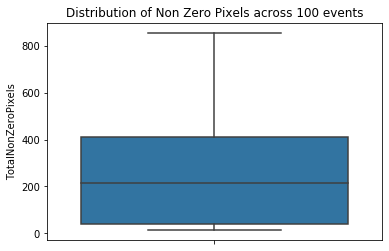


Fig 3 : Distribution of Non Zero Pixels across 100 events

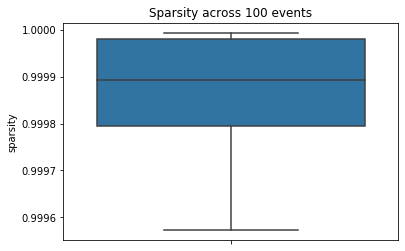


Fig 4: Distribution of sparsity across 100 events

**Challenges:**

This section details the few challenges with regards to the current dataset and its Image based representation

* **Difficulties in encoding underlying detector geometry in an image:**
  + Image based representations make specific assumptions on the detector geometry or granularity. Images are derived by binning the (eta, phi) or (x,y) plane and the binning choice should be representative of typical tracking and calorimeter resolutions.
  + The latest research suggests that binning approach discussed in different image based approaches neglect the complexity of a typical detector geometry. They are very suitable for the central barrel of a typical detector, but they overlook the existence of complicated overlap regions between the barrel and the endcap regions, specifically the irregular geometry of the endcaps. Pileup subtraction is particularly important in the endcap regions, where the solenoid magnetic field of a typical cylindrical detector pushes the abundant low-pT particles produced in pileup interactions.
  + The use of a calorimeters-inspired (eta, phi) fixed-size grid neglects the fact that charged particles are mainly detected through the inner tracker. A typical-tracker (eta, phi) resolution can hardly be represented as a fixed number, since it depends on the track pT .
* **Difficulties in incorporating information from different particles in a single image:**
  + How to incorporate additional information of the particles is unclear, as it involves combining non-additive quantities (e.g., the particle type) of multiple particles entering the same cell.
* **Sparsity of Data and Computational Cost:**
  + As depicted in Fig 3 and Fig 4, even if we encode the event into a 200\* 200 pixel Image, it leads to a very sparse representation: in the given dataset, more than 99% of the pixels are blank. This makes the CNNs highly computationally inefficient on these event images.