ML status summary

13th sept

Data (on old Xe Mak.)

- Moved to different data handling / pipeline
  straight from raw images to numpy
  - fixes some issues such as normalisation being lost
  - easier to track pit metadata

  - faster
- Labelled data at higher preselection density cut (30 → 60), approx 1400 → 7000 samples
  ---
  Will fix datasets at 200, 500,and 1000 examples per class, will allow comparison / benchmark of the impact data quantity has. K-folding within these sets provides an estimate of performance variation due to choice of training data (Assume variable ‘true label truth’)

Performance

- Network performance on new normalisation not ‘better’ although more stable.
  - Find new set of optimal HP’s

Interleaved foils

- Stack with LHC exposed foils interleaved with clean test beam foils arrived june.
- Provides a confirmed TRUTH label
- Scanning microscope down / fixed 17th sept
- Different test beam ion and plastic thickness / nature, may require different approach
  - can still do 3d approach, eg, feeding front and back imaging as 2*X*Y array