



ML status - 30th october

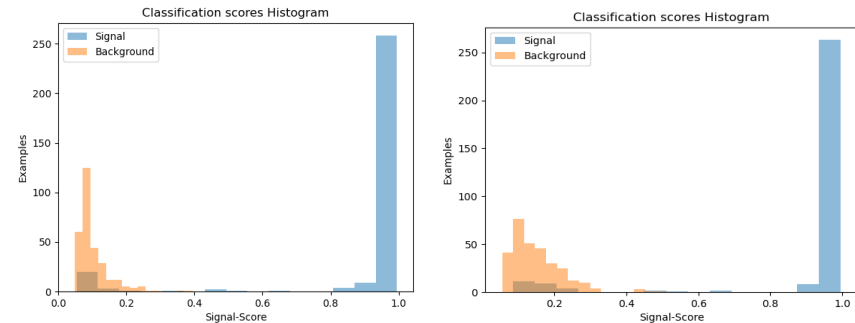
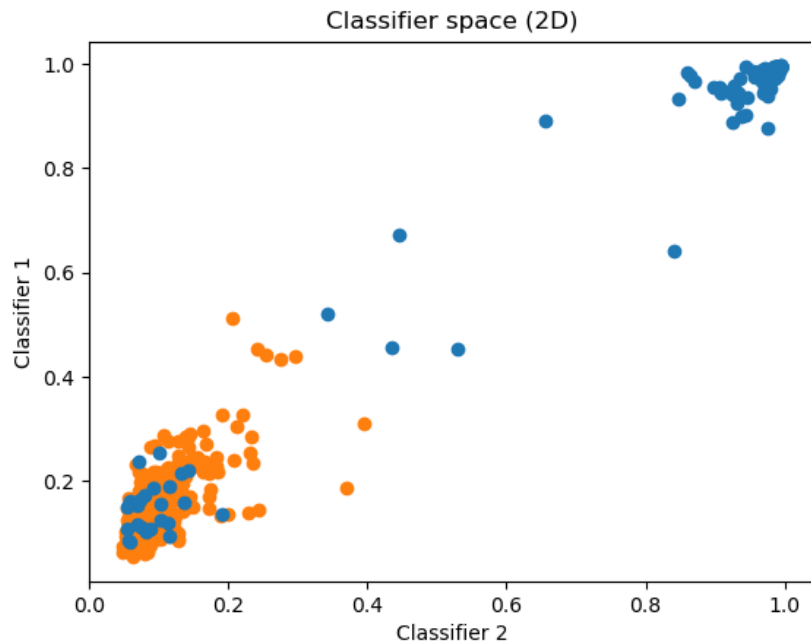
l.millward@qmul.ac.uk

Error sources

- Two error sources;
 - Signal confused for background
 - Irreducible signal loss?
 - Background confused for signal
- Change to dataset: Double confirmation* on all 'signal' etch-pits *(single foil pits still interesting as indicate lighter ionisations)
- 500 sig + 500 bkg training set (300 for val)
- Ensemble of CNNs trained on 5 different data-folds

Ensemble

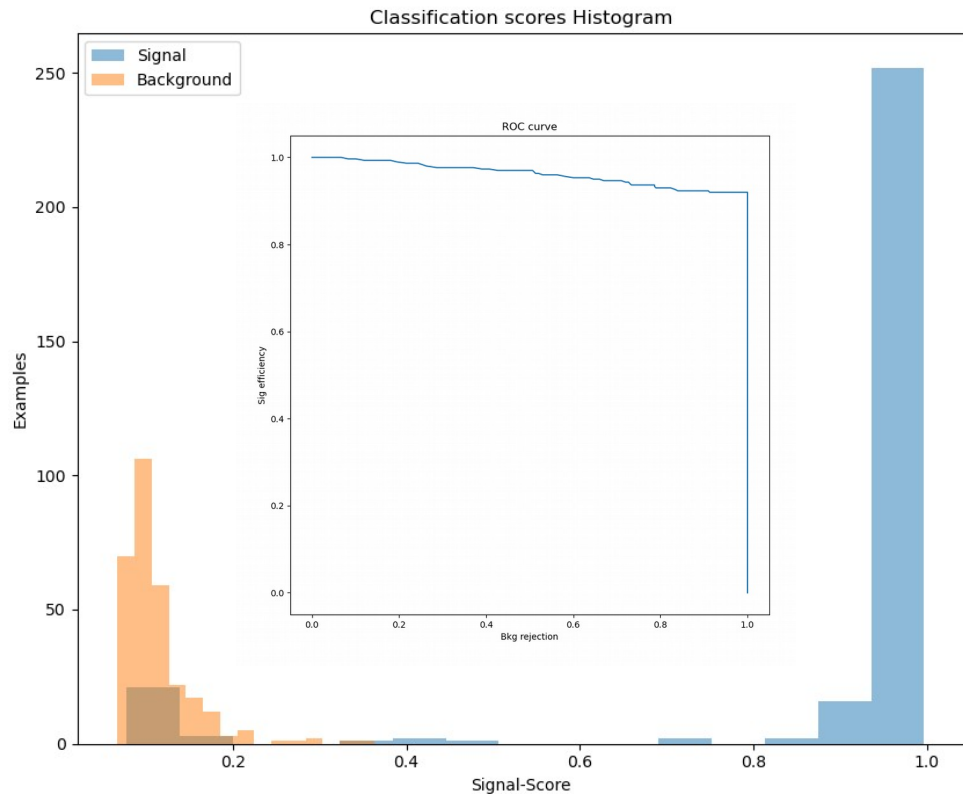
- Validation set of 300 examples



Output: Ensemble of 5 classification scores
(2 shown for simplicity)

- Using double confirmed pits as signal prototype → bkg doesn't get confused for signal
- Error comes from signal loss
- Most signal loss occurs with the same pits for all classifiers

Ensemble average



- Taking average, pretty irrelevant WHICH working point we choose between 0.2 and 0.4

- Eg,
wp .2 : 11fp / 24tn
wp .3 : 3fp / 24tn
wp .4 : Zero fp / 26tn

ie,
consistent 10% signal loss
3%/1%/0%

True negatives

- Examination of true negatives
- Pit loss in exposed foil - Debateable how recoverable
- Inspection in clean foils indicates small etch-pits

