Hashing Classification for charged particle tracking

Luiza Adelina Ciucu (ATLAS)

03 July 2020

UNIVERSITÉ
DE GENÈVE

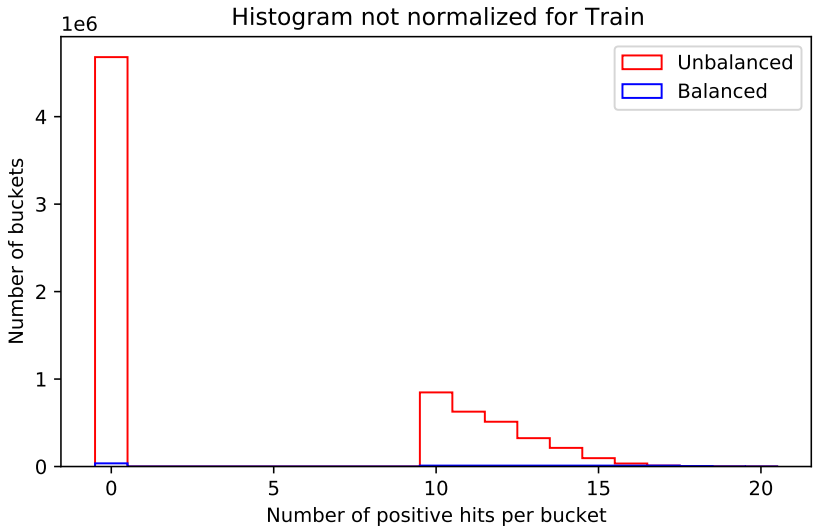
FACULTY OF SCIENCE
Physics Section

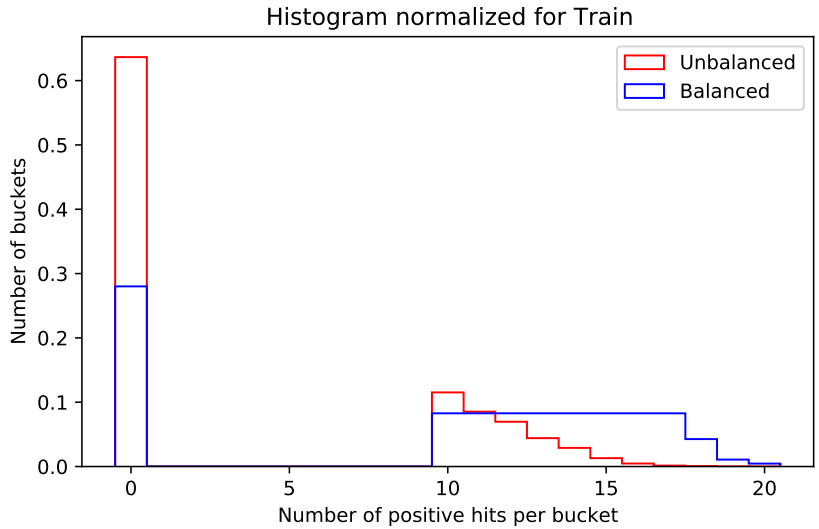
Introduction

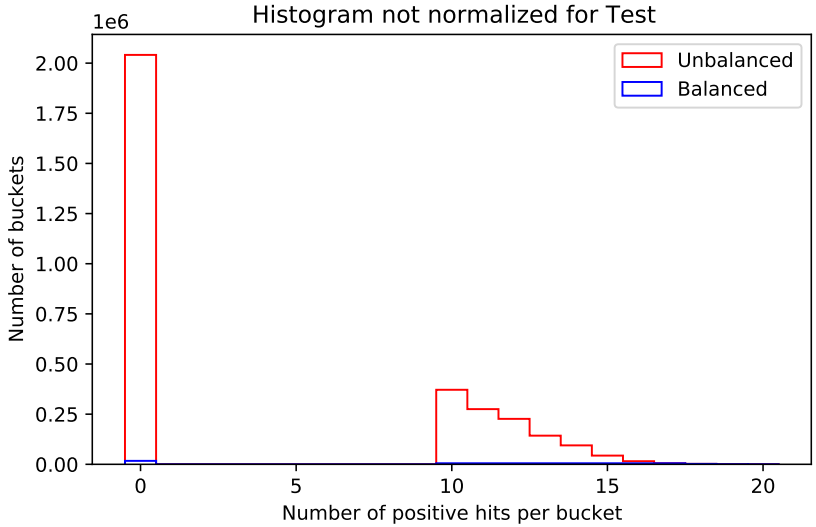
- 100 events. For each group of 10: 7 train, 3 test.
- If nbPositiveHit<10, set nbPositiveHit=0 and output made only of -1.
- For Train used Balanced (130k), for Test use Unbalanced (3.2M).
- Balancing Train in two steps, as shown last time.
  - Make peak flat between 10-17 (with value of 17).
  - Reduce nbPositiveHit=0 until 50% Pos, 50% Neg.
- New studies today
  - Last time Test was also Balanced. Now Test is Unbalanced.
  - With constraint of our output of -1 and 1:
  - Try other output layer activation functions: tanh, squared non linear, soft sign.
  - Try other loss functions: squared hinge and hinge.
  - Try 120 epochs vs 50 epochs.

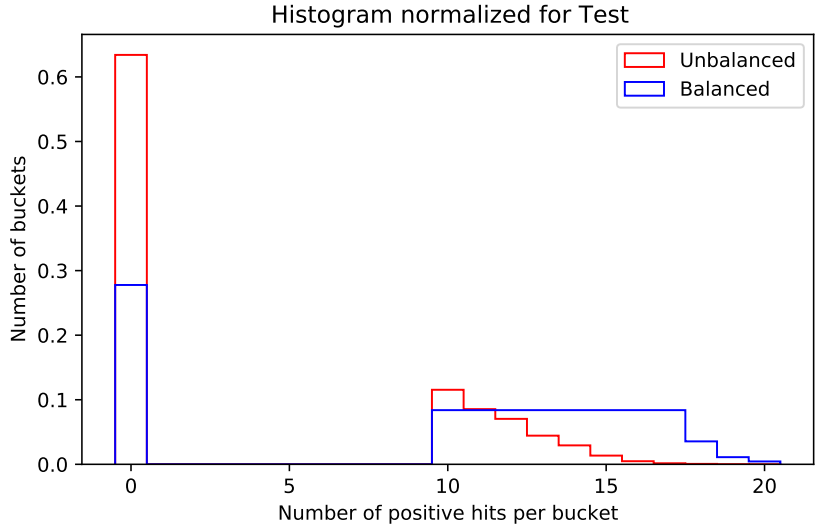
Balancing results N=17

- For Train used Balanced (130k), for Test use Unbalanced (3.2M).



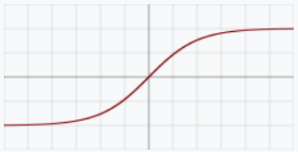
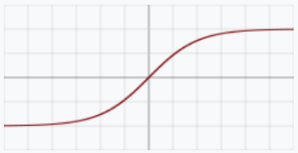
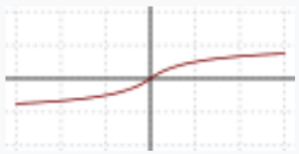




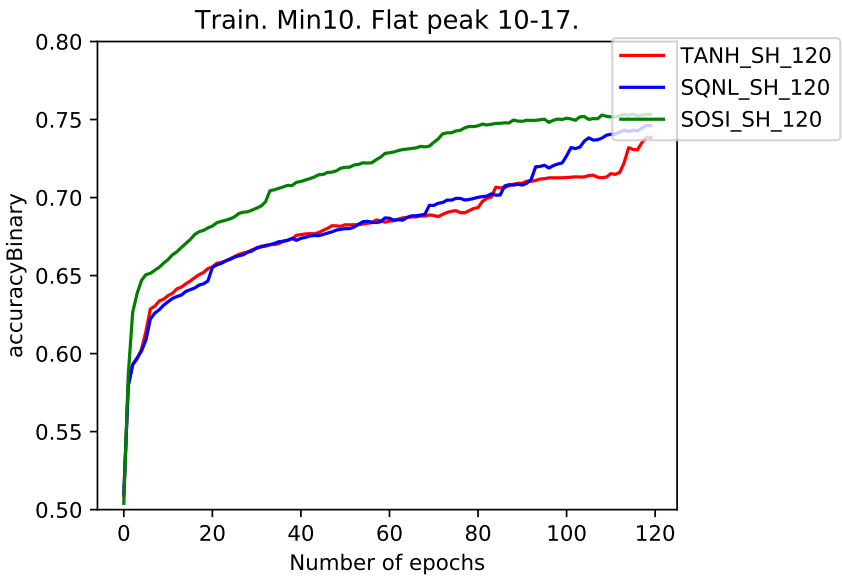


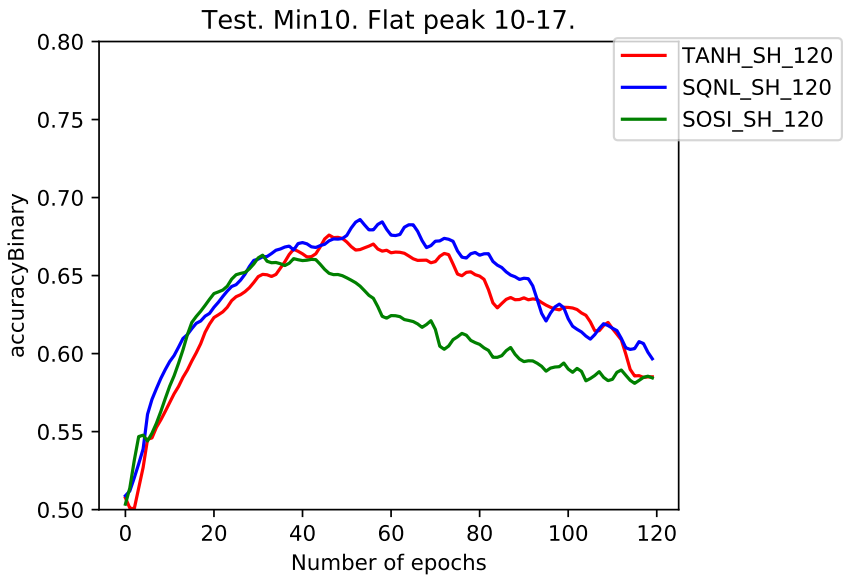
Study 1: varying output layer activation functions

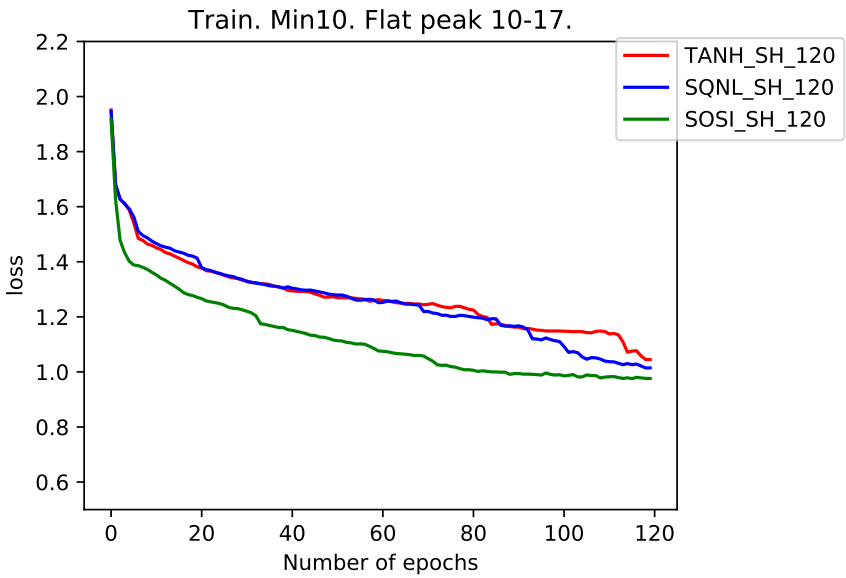
- Tanh, Squared non linear, Soft Sign.

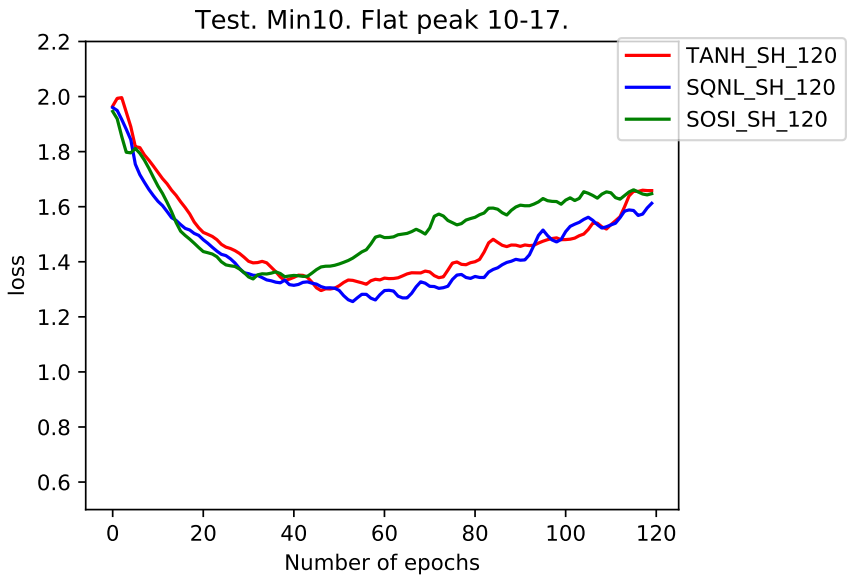
TanH		$f(x) = \tanh(x) = \frac{(e^x - e^{-x})}{(e^x + e^{-x})}$
SQNL[10]		$f(x) = \begin{cases} 1 & : x > 2.0 \\ x - \frac{x^2}{4} & : 0 \leq x \leq 2.0 \\ x + \frac{x^2}{4} & : -2.0 \leq x < 0 \\ -1 & : x < -2.0 \end{cases}$
ElliotSig[11][12] Softsign[13][14]		$f(x) = \frac{x}{1 + \lvert x \rvert}$

Accuracy and Loss from training





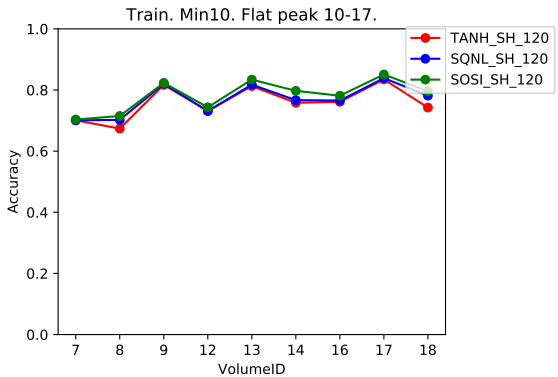


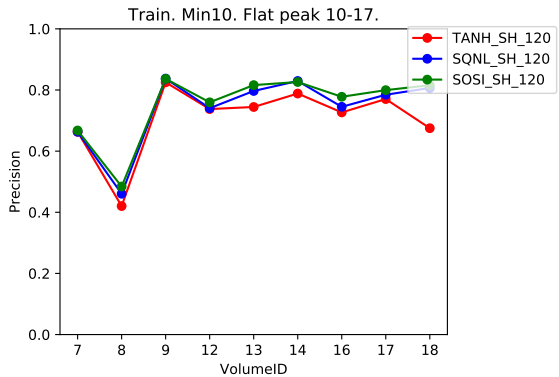


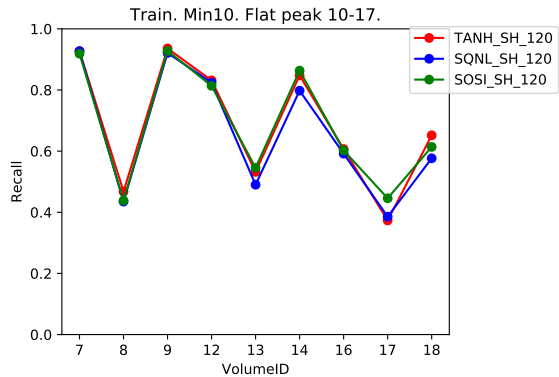
Metrics for each VolumeID.

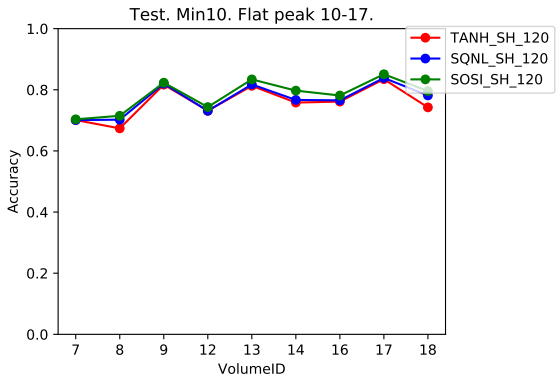
- Three activation functions quite similar.

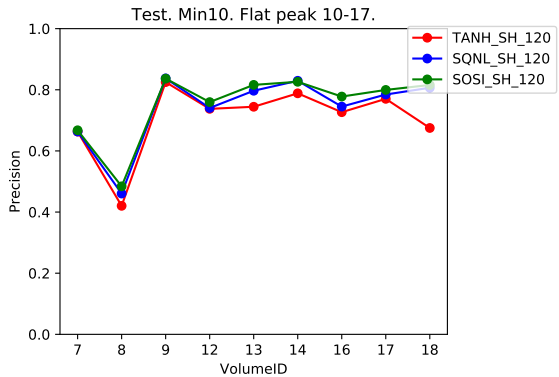
Accuracy	Precision	Recall
$\frac{TP+TN}{TP+FP+FN+TN}$	$\frac{TP}{TP+FP}$	$\frac{TP}{TP+FN}$

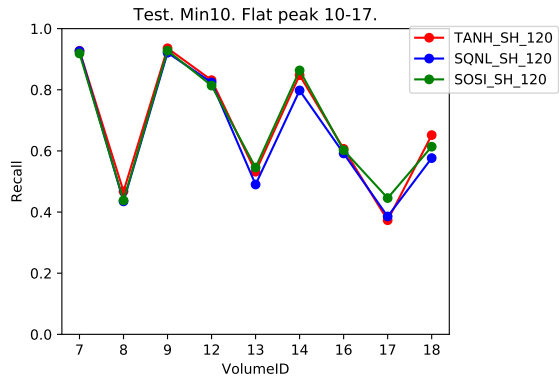




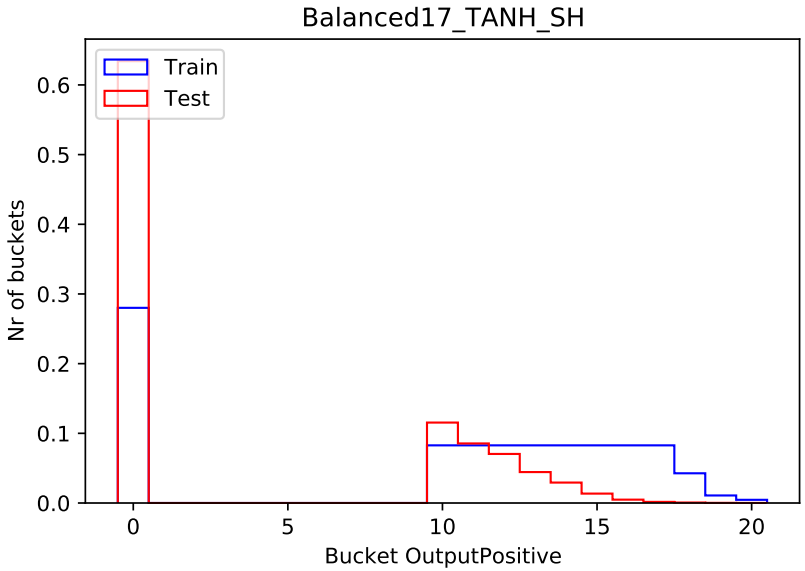


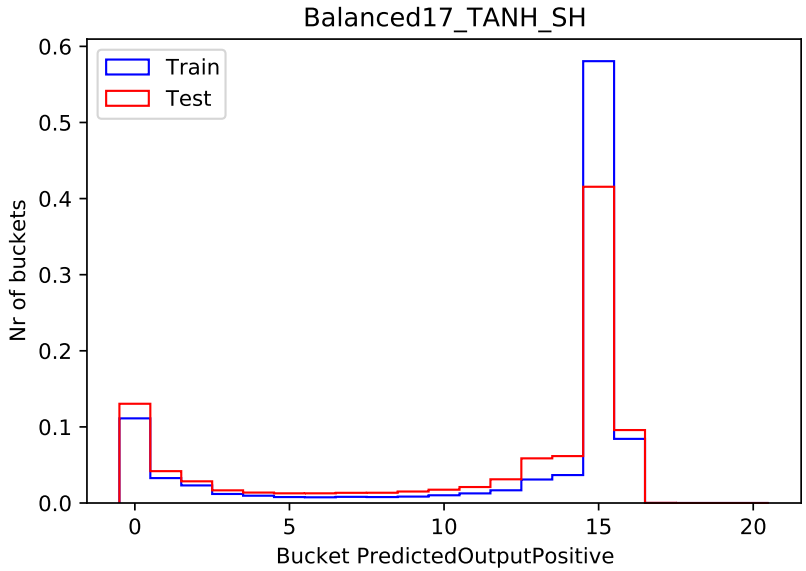


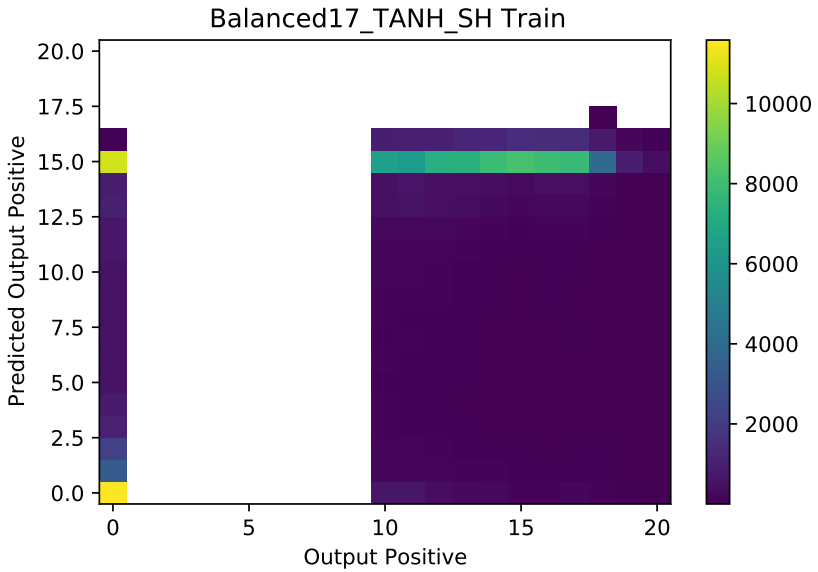


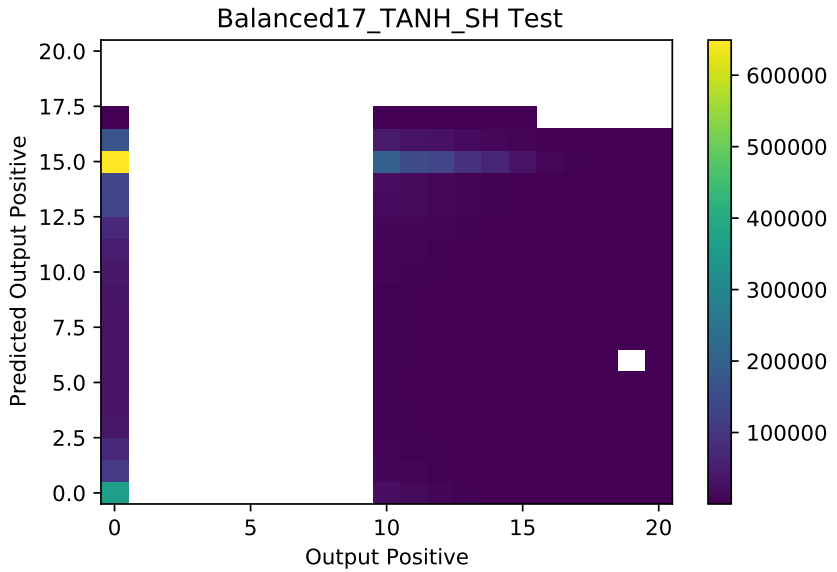


2D plots TANH SH 120

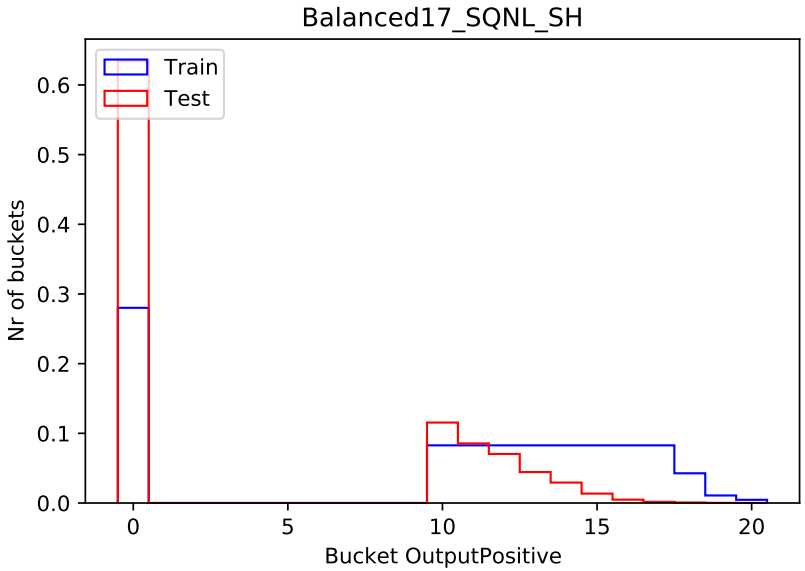


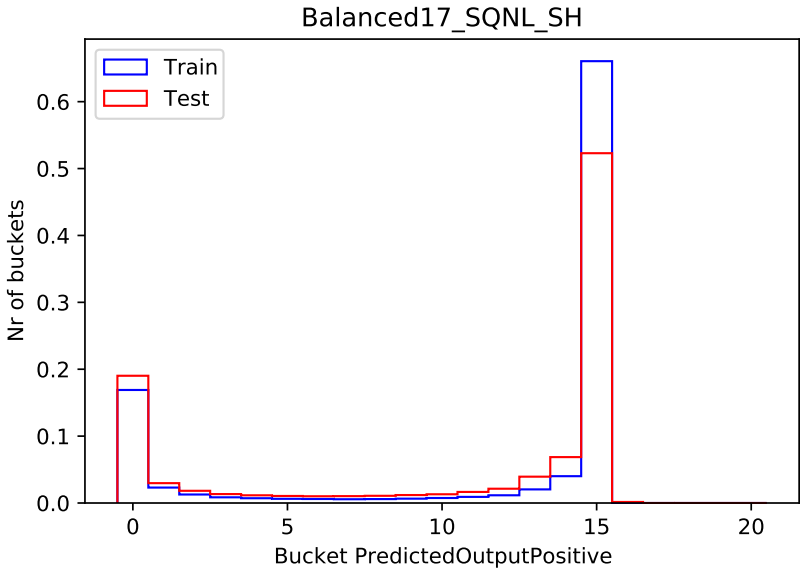


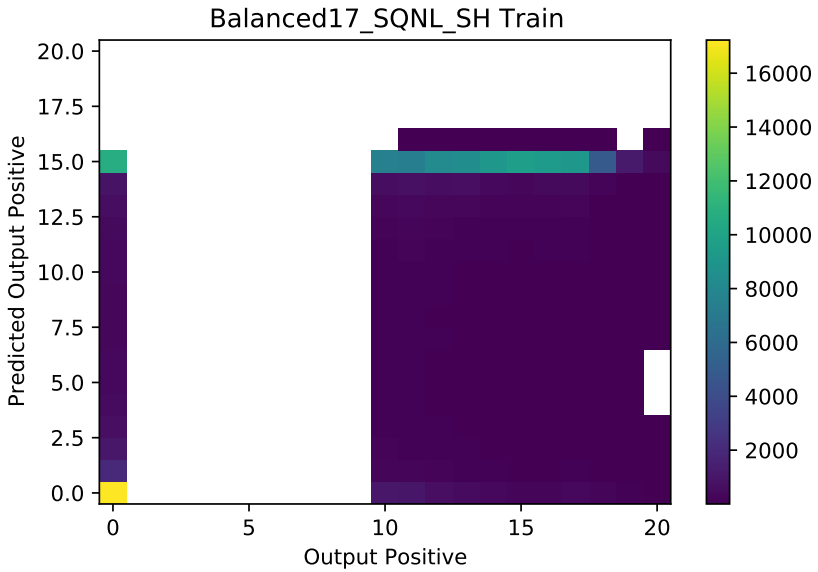


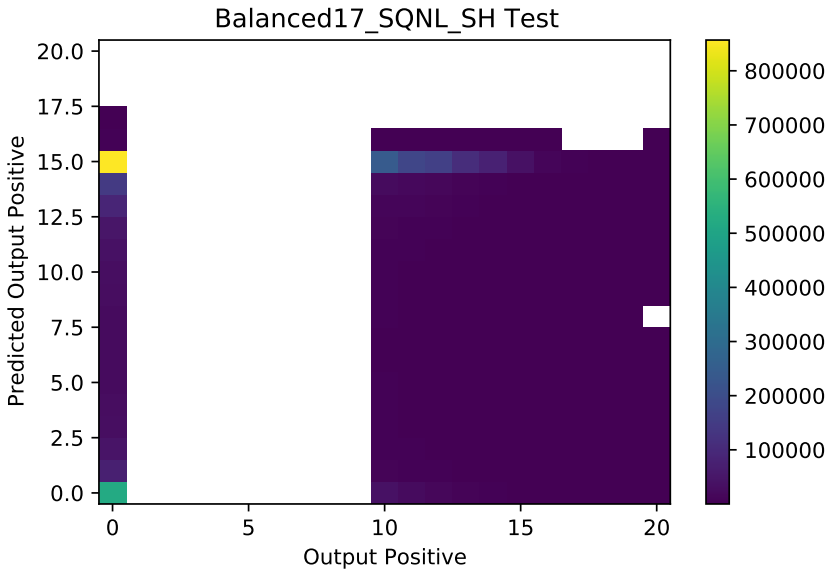


2D plots SQNL SH 120

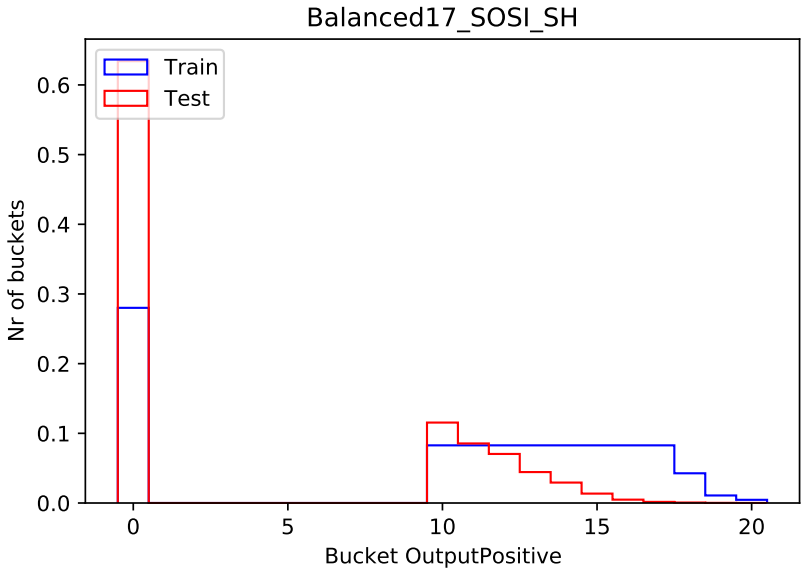


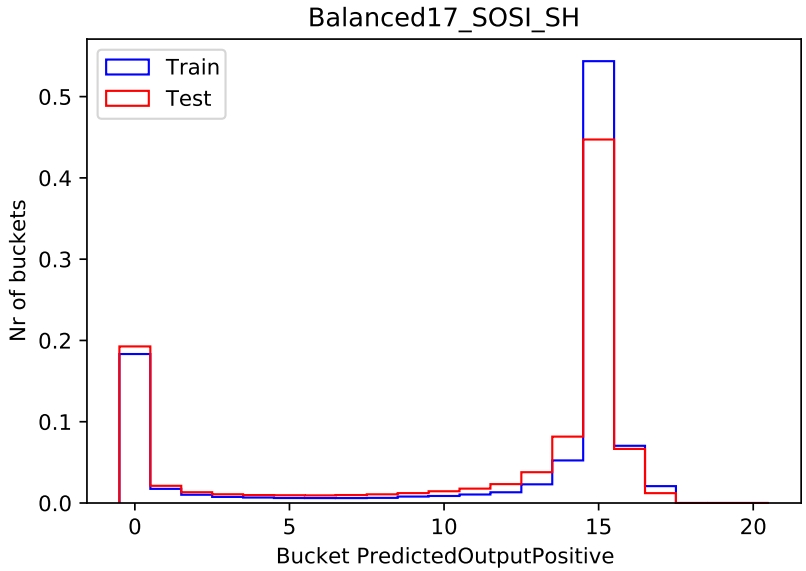


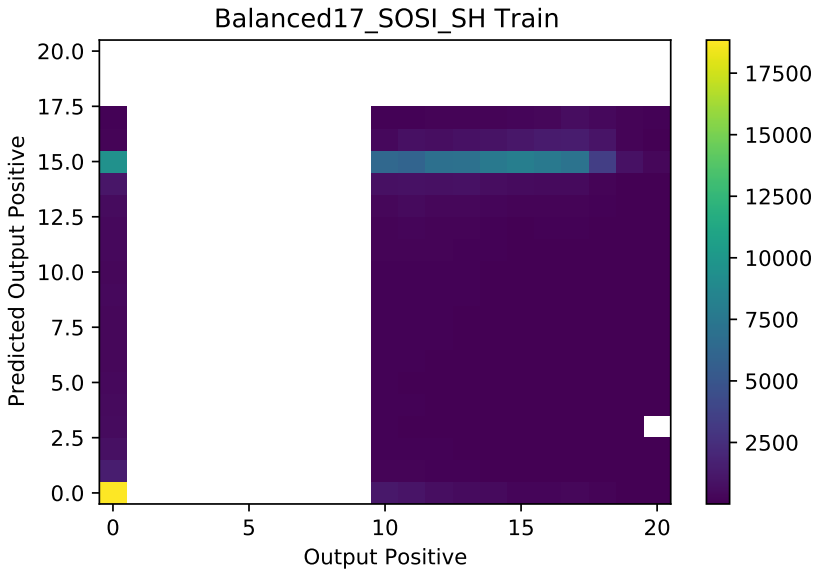


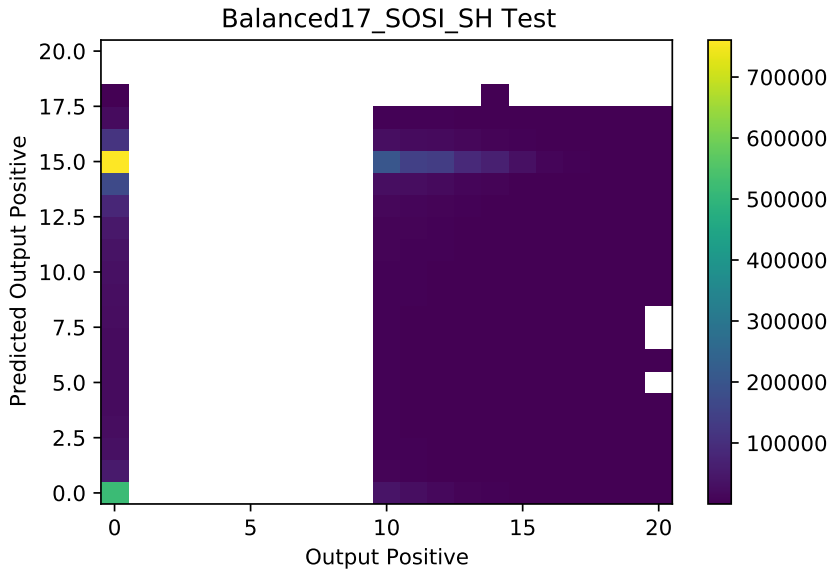


2D plots SOSI SH 120









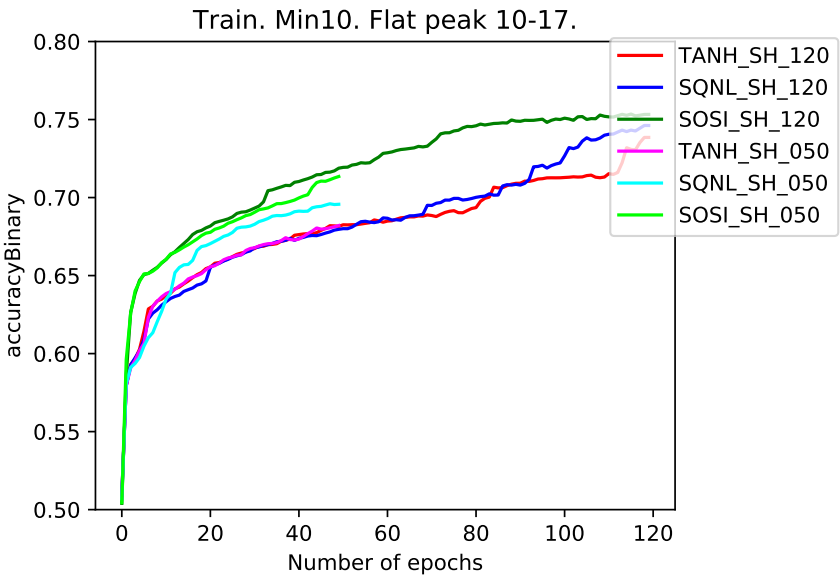
Study 1: varying output layer activation functions

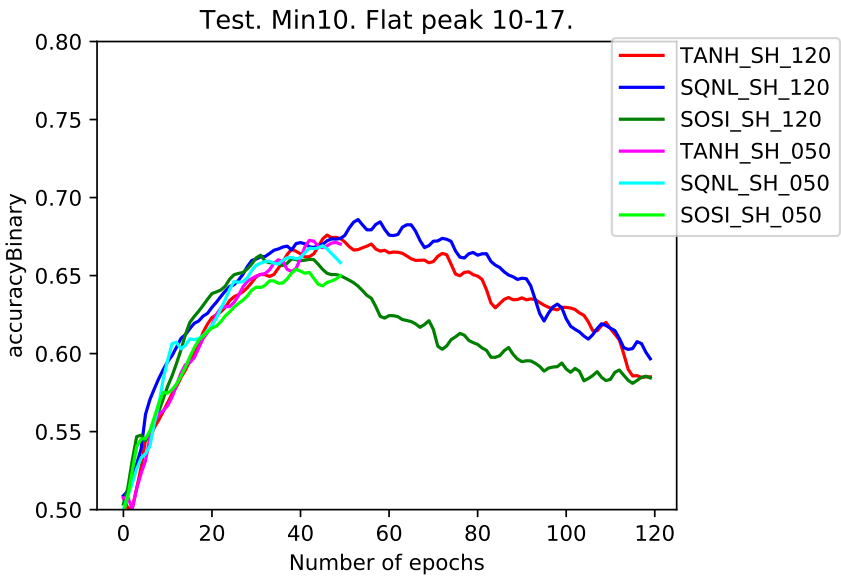
- Conclusions:
- They are very similar.
- But it looks like that after 50 epochs they start to over-train.
- So ran again with 50 epochs.
- But as you will see, the results look a bit worse, including for Test.

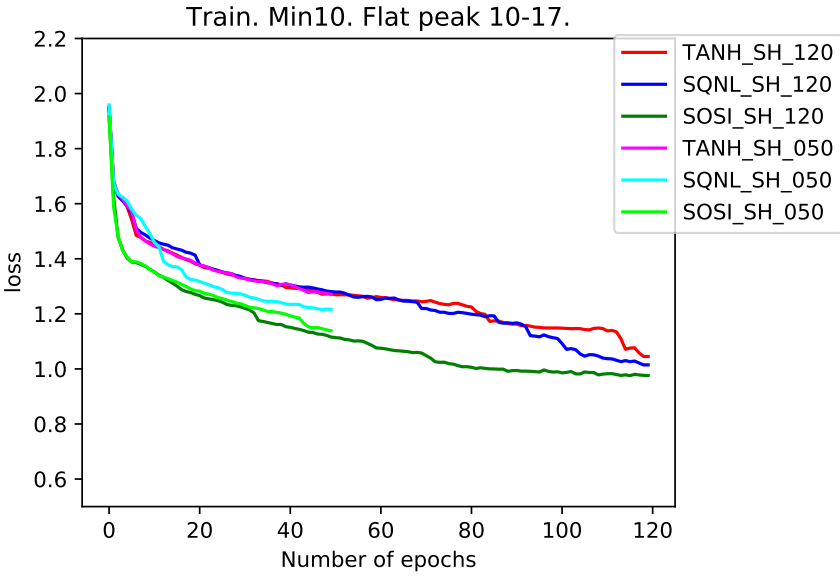
Study 2: 120 vs 50 epochs

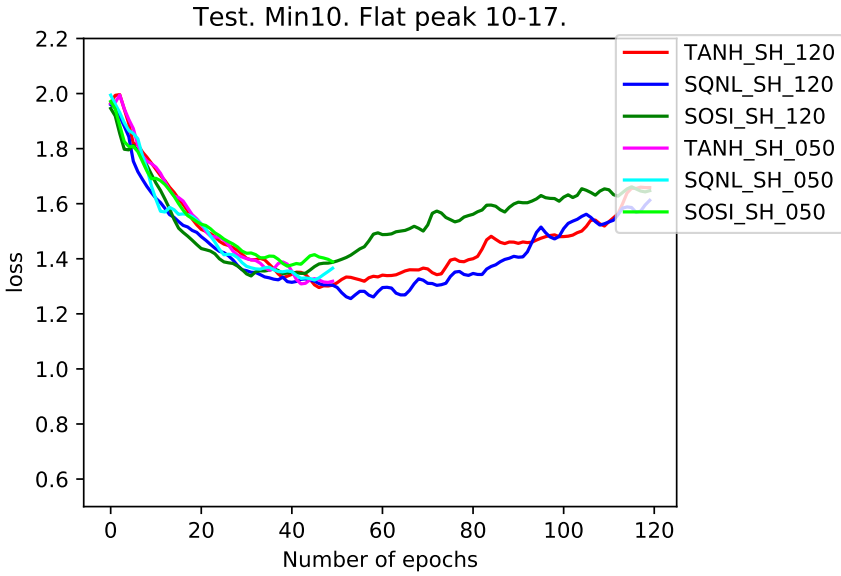
- For same activation function, same colour, different nuances for 120 and 50 epochs.

Accuracy and Loss from training





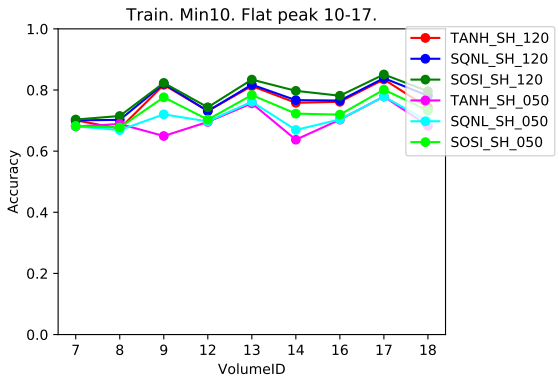


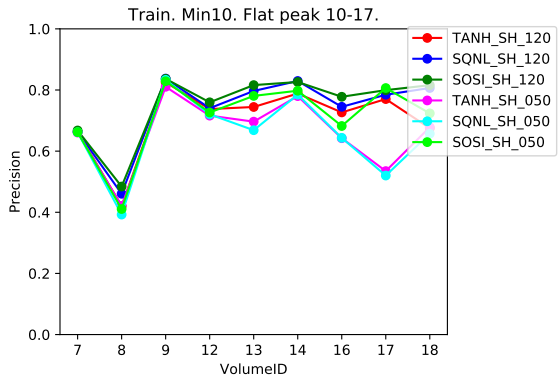


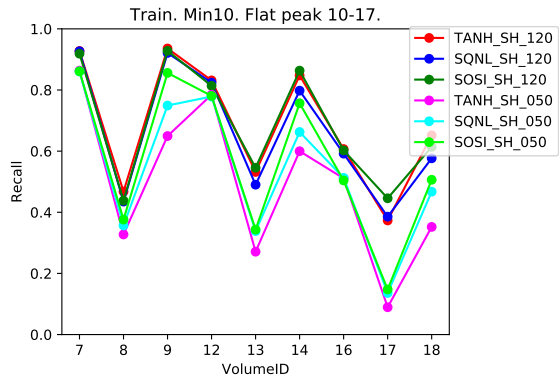
Metrics for each VolumeID.

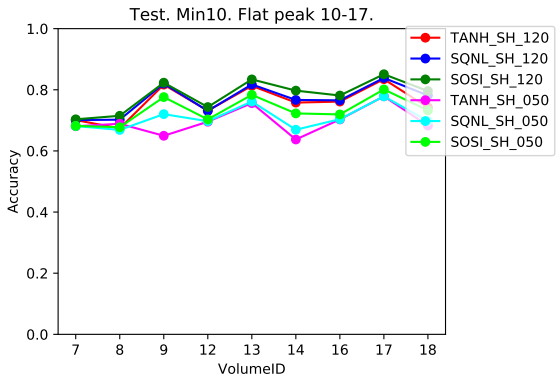
- 50 epochs slightly worse than 120 epochs.

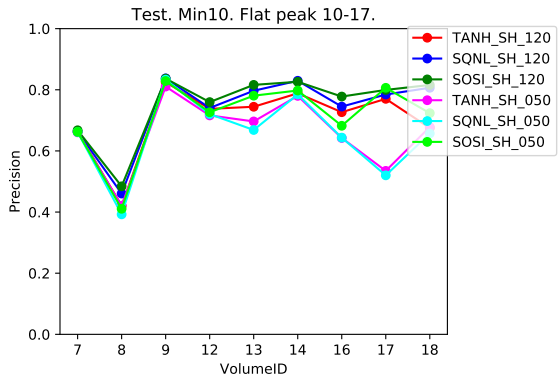
Accuracy	Precision	Recall
$\frac{TP+TN}{TP+FP+FN+TN}$	$\frac{TP}{TP+FP}$	$\frac{TP}{TP+FN}$

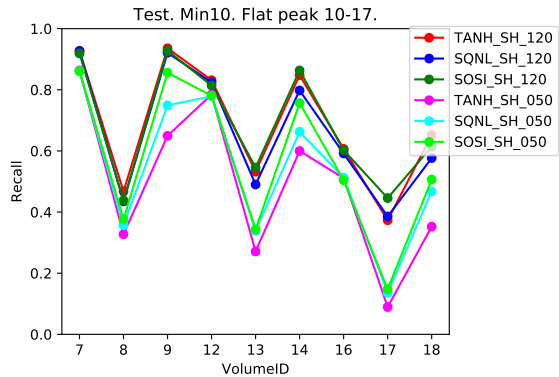




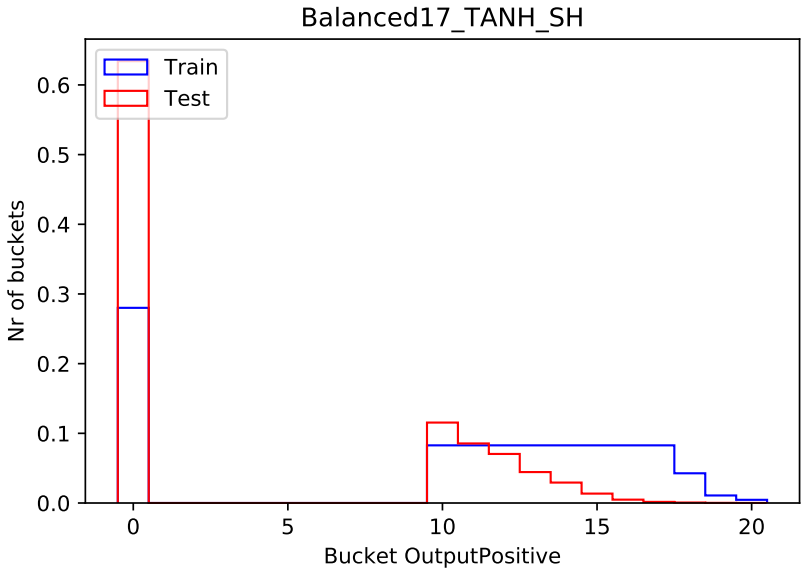


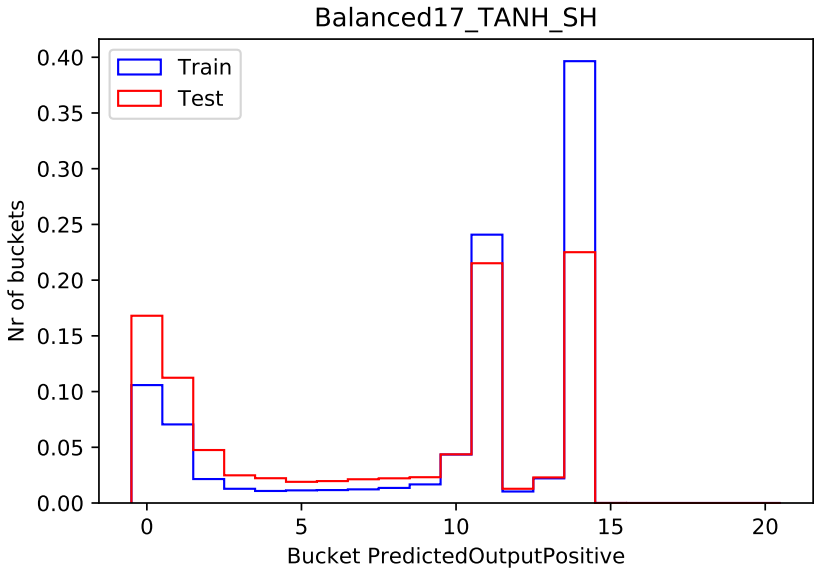


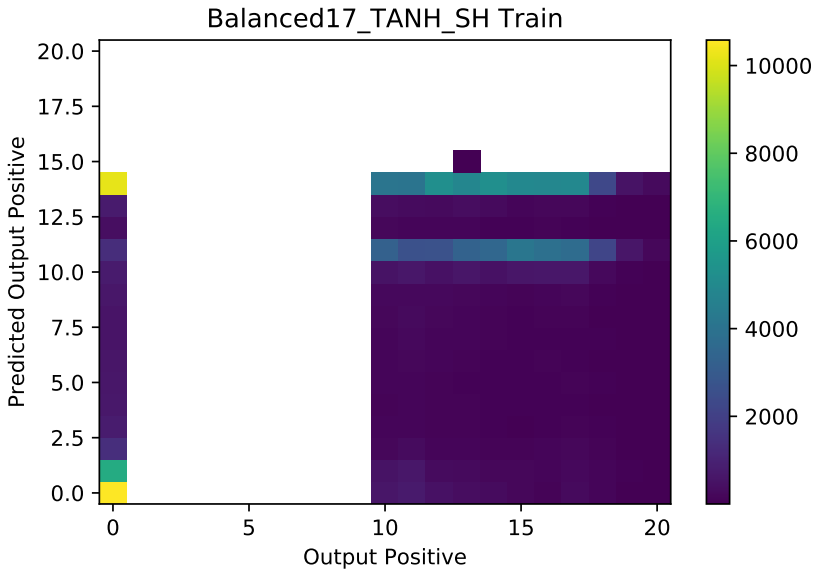


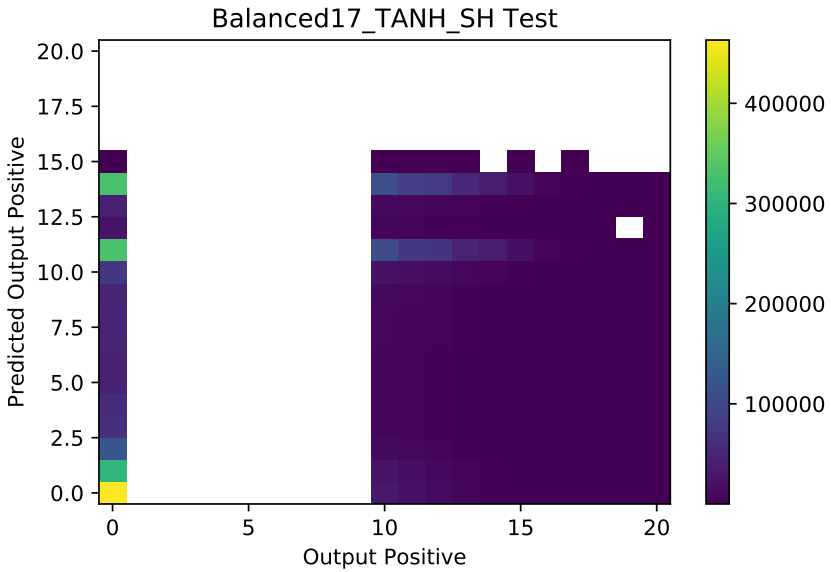


2D plots TANH SH 50

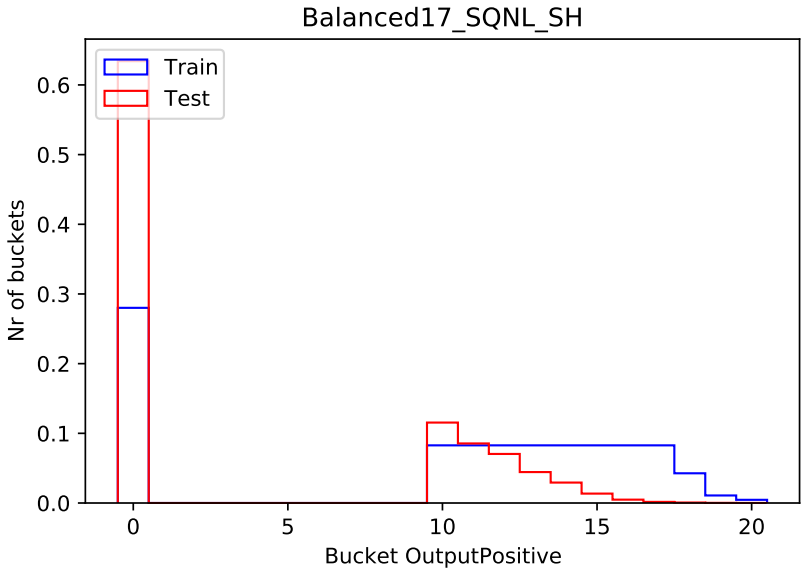


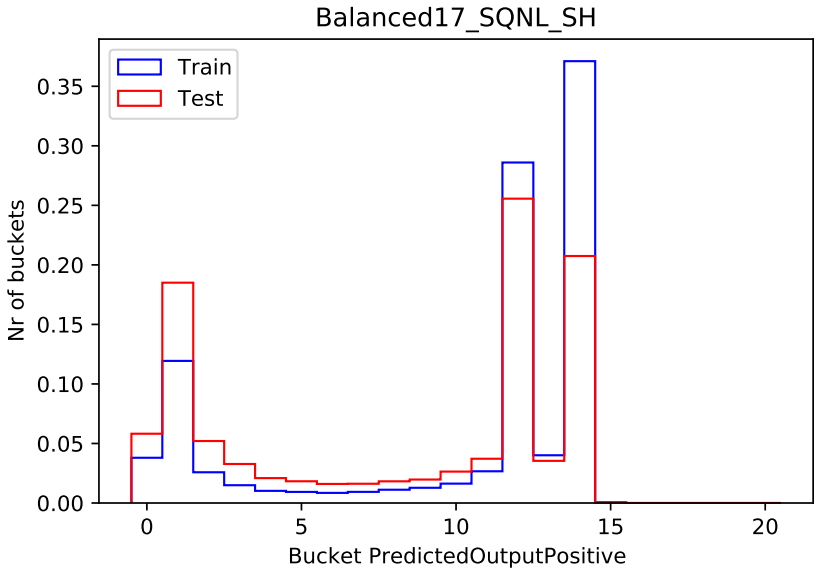


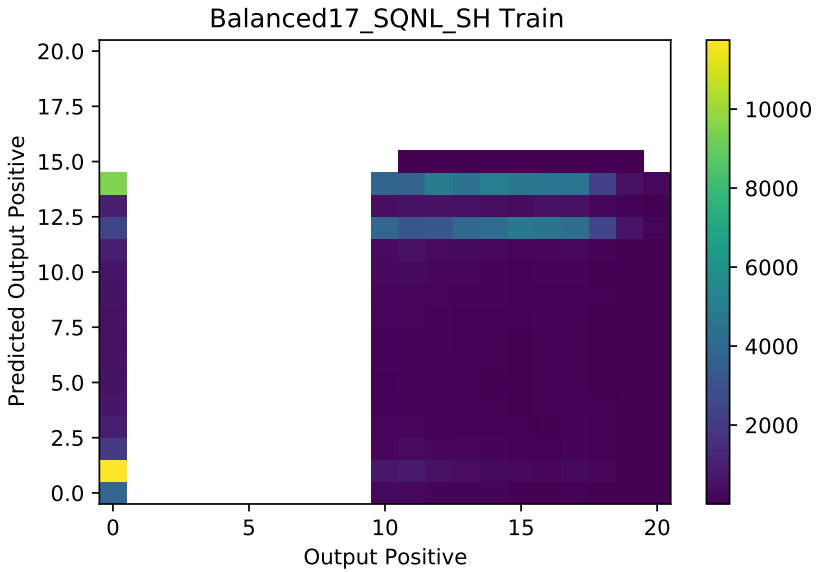


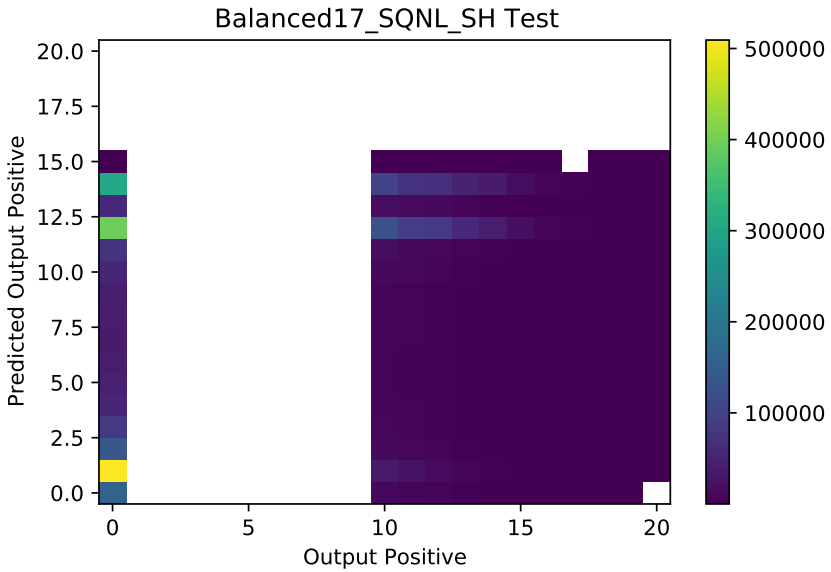


2D plots SQNL SH 50

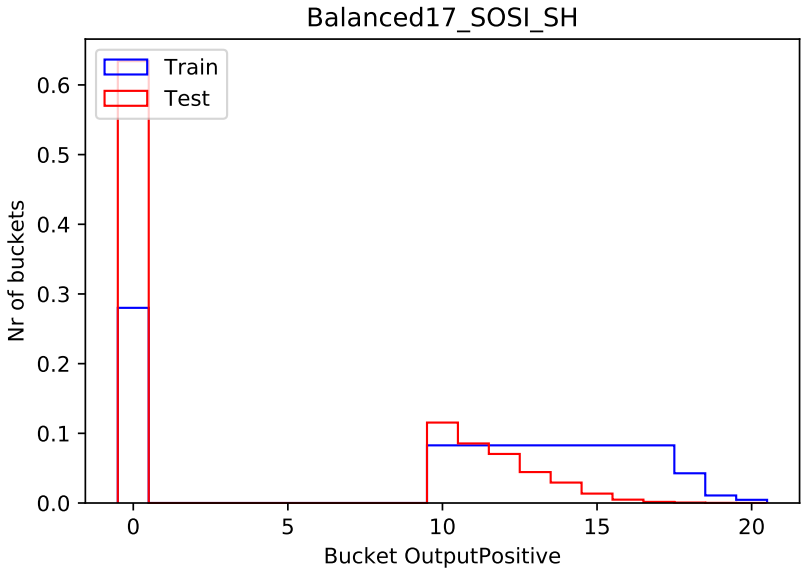


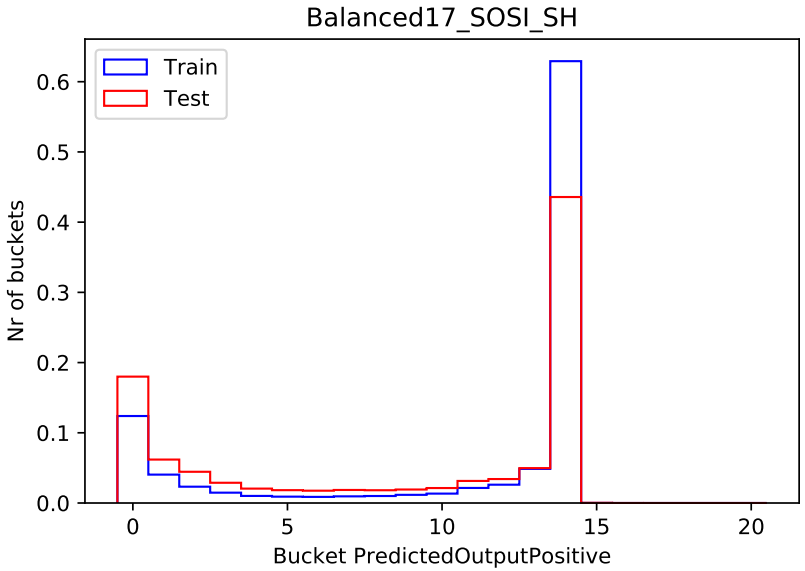


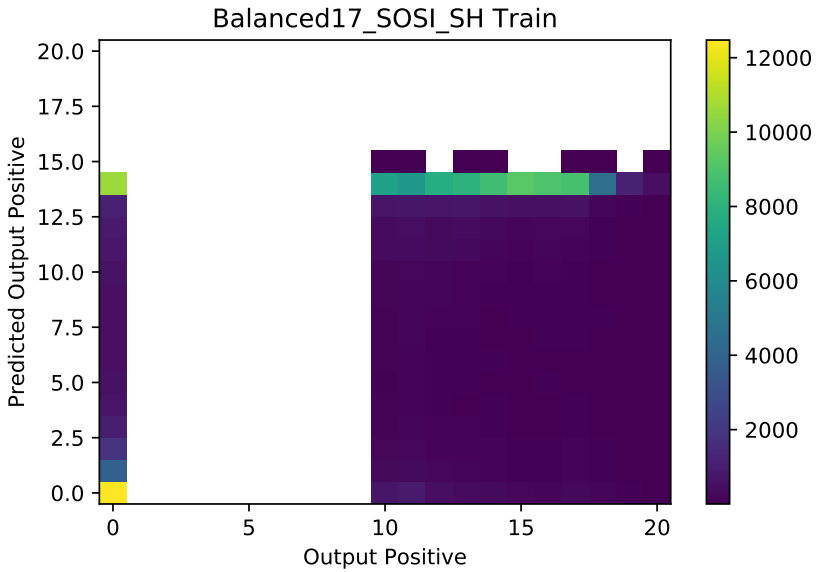


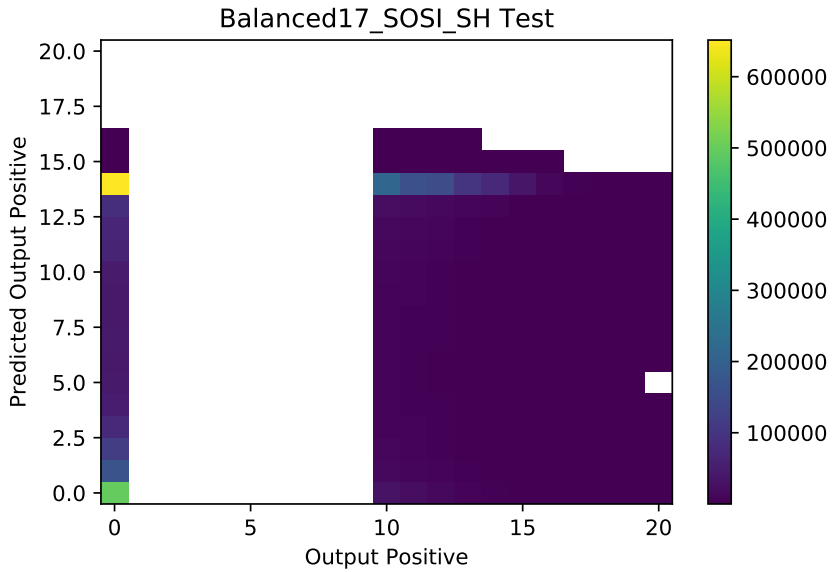


2D plots SOSI SH 50









Study 2: varying number of epochs

- Conclusions:
- 120 vs 50 epochs.
- Looking at 120 epochs it appeared that from 50 epochs it started to overtrain.
- So ran again with 50 epochs.
- But results look a bit worse, including for Test.

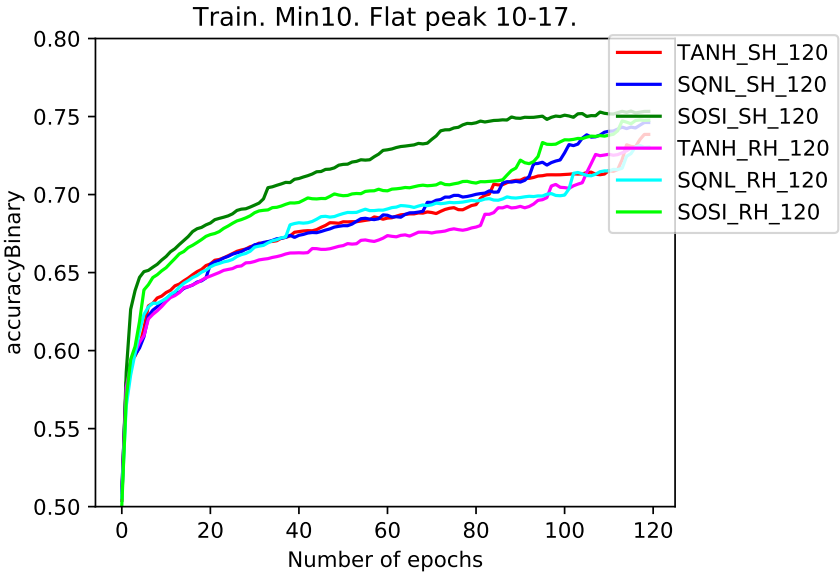
Conclusion

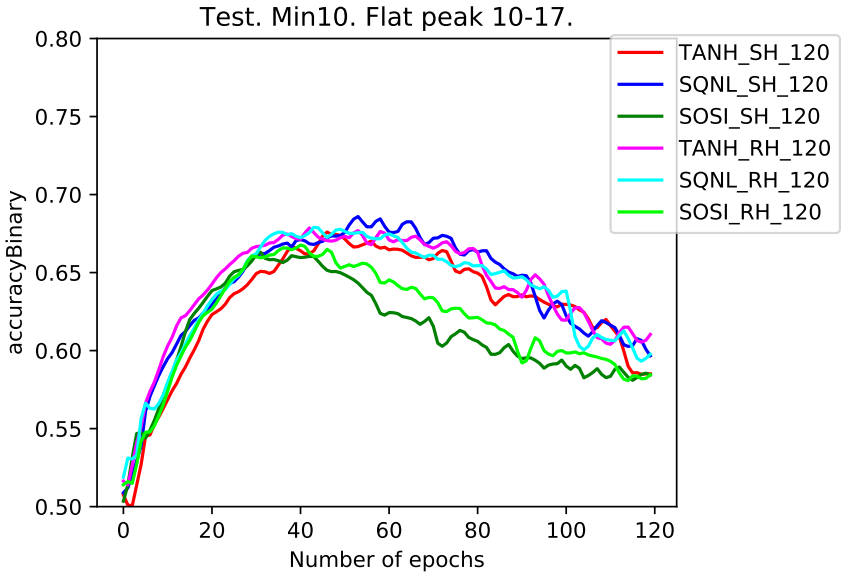
- 100 events. For each group of 10: 7 train, 3 test.
- If nbPositiveHit$<$10, set nbPositiveHit=0 and output made only of -1.
- For Train used Balanced (130k), for Test use Unbalanced (3.2M).
- Balancing Train in two steps, as shown last time.
  - Make peak flat between 10-17 (with value of 17).
  - Reduce nbPositiveHit=0 until 50% Pos, 50% Neg.
- New studies today
  - Last time Test was also Balanced. Now Test is Unbalanced.
  - With constraint of our output of -1 and 1:
  - Very similar with various layer activation functions: tanh, squared non linear, soft sign.
  - Very similar with various loss functions: squared hinge and hinge.
  - Overtrain at 50 epochs maybe a false alarm? As 50 epochs slight worse than 120 epochs. Try 120 epochs vs 50 epochs.

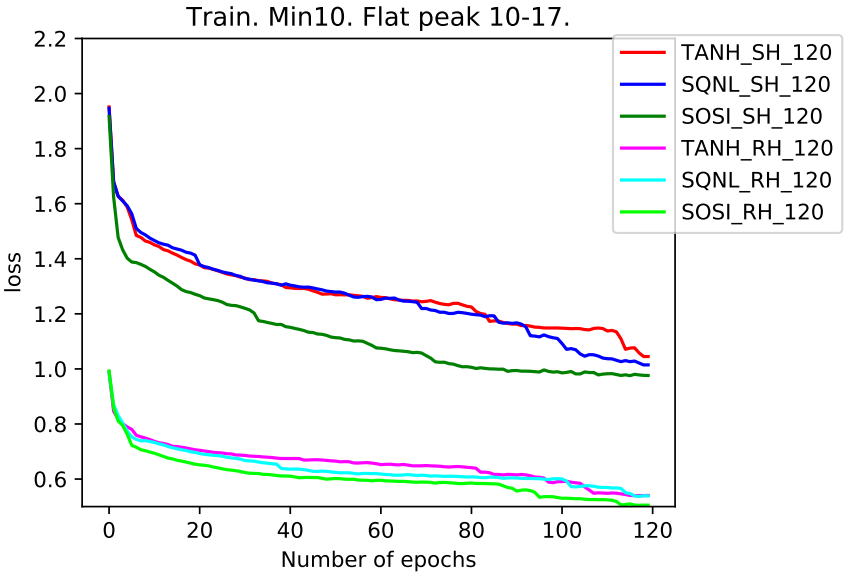
Study 3: for 120 epochs, vary loss function

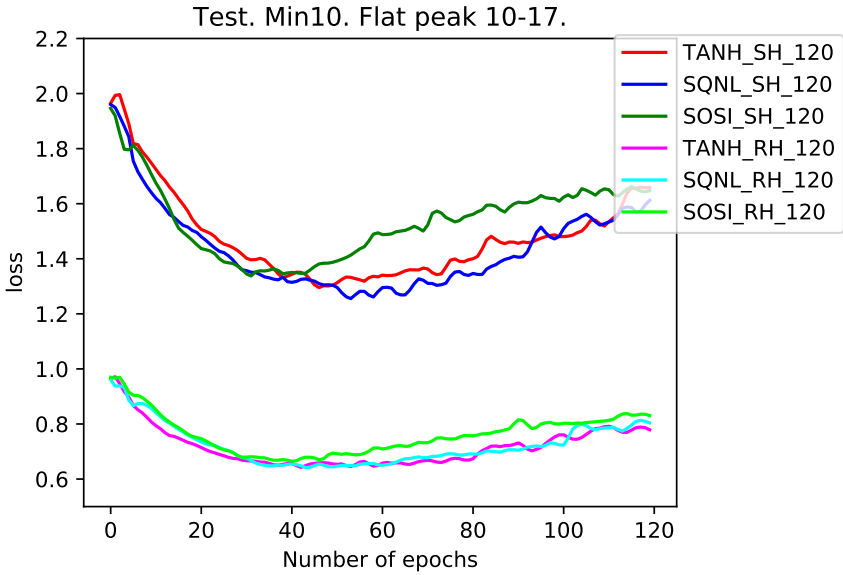
- Squared hinge, Hinge.

Accuracy and Loss from training





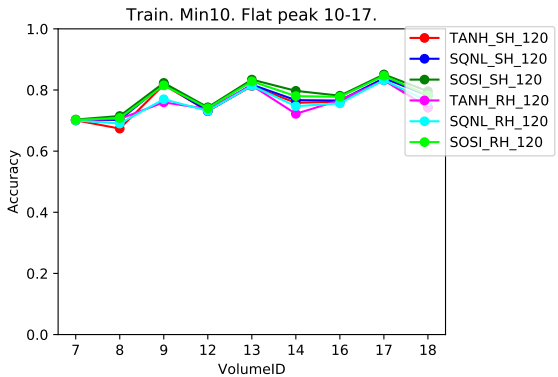


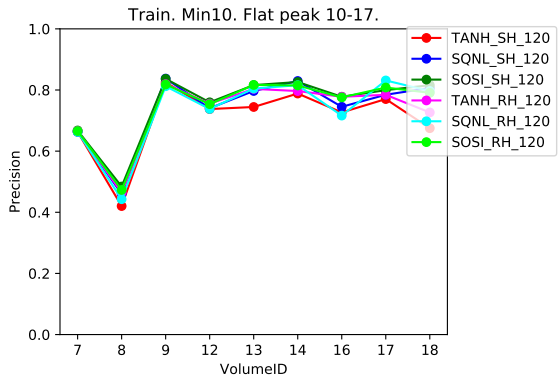


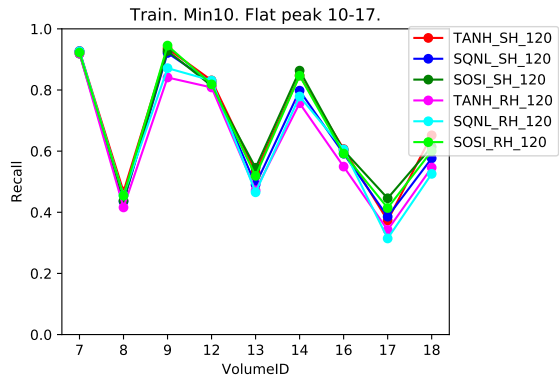
Metrics for each VolumeID.

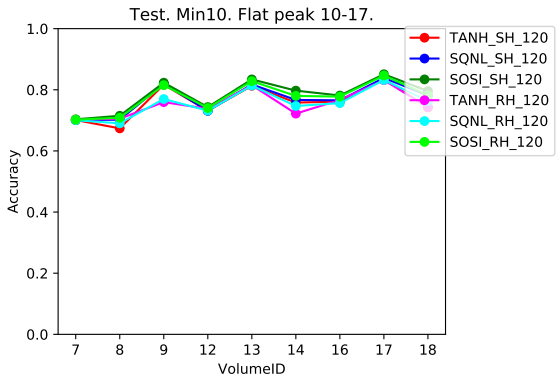
- Squared hinge vs hinge quite similar.

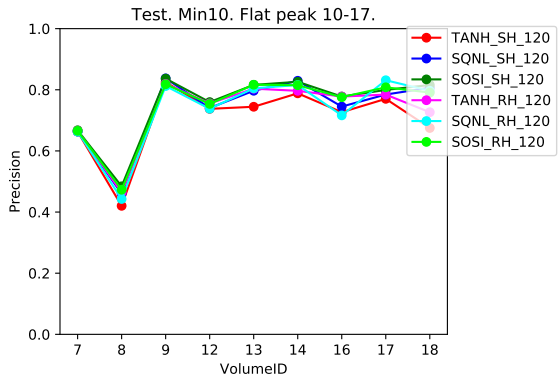
Accuracy	Precision	Recall
$\frac{TP+TN}{TP+FP+FN+TN}$	$\frac{TP}{TP+FP}$	$\frac{TP}{TP+FN}$

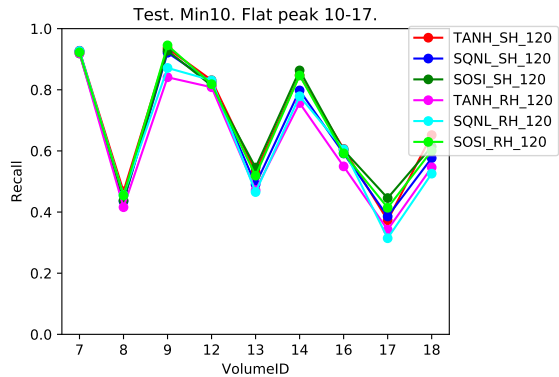




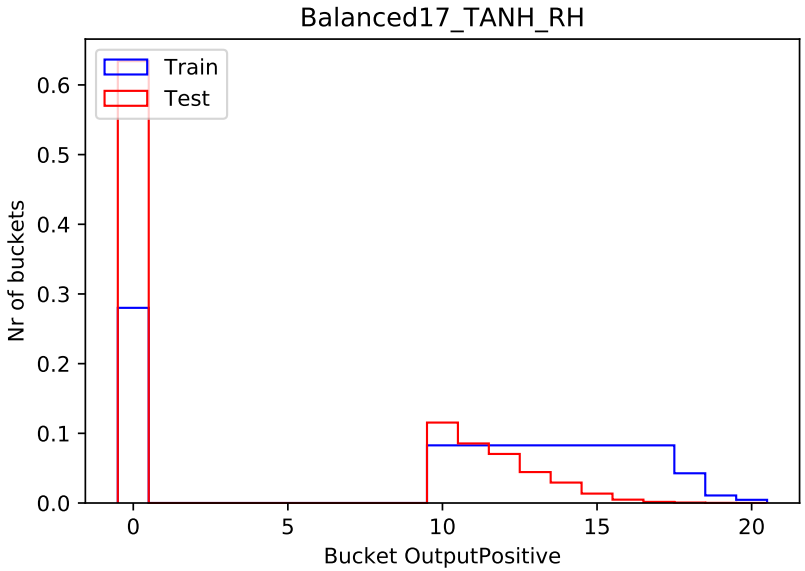


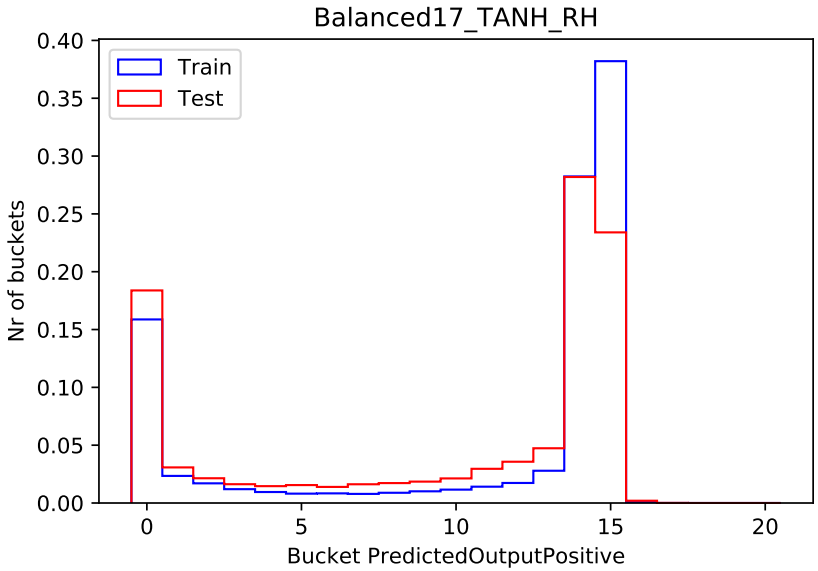


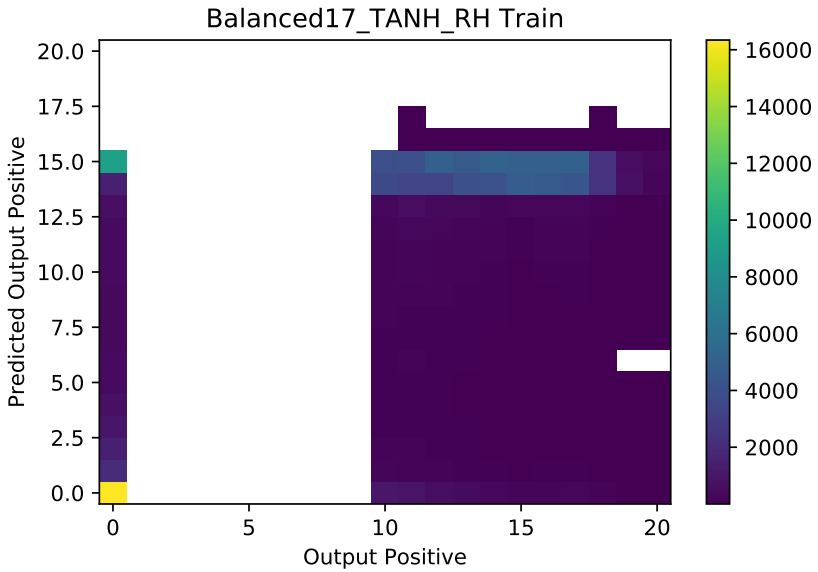


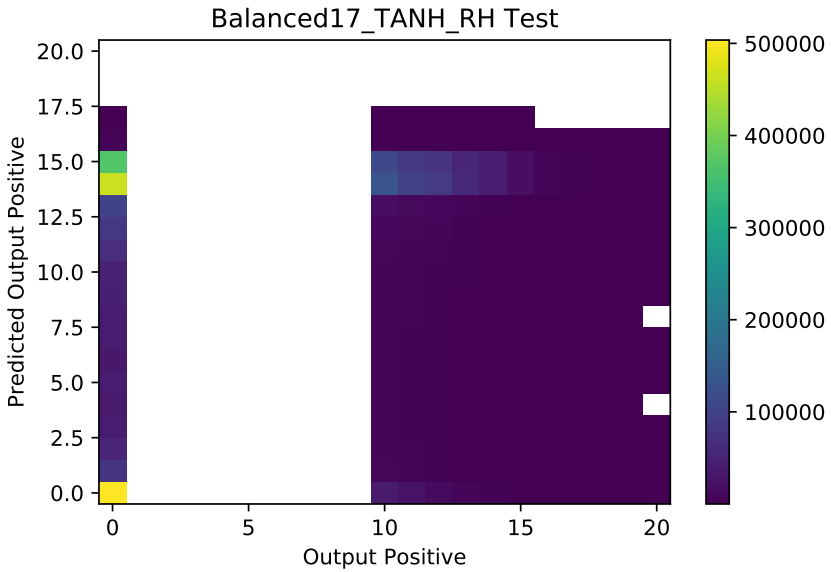


2D plots TANH RH 120

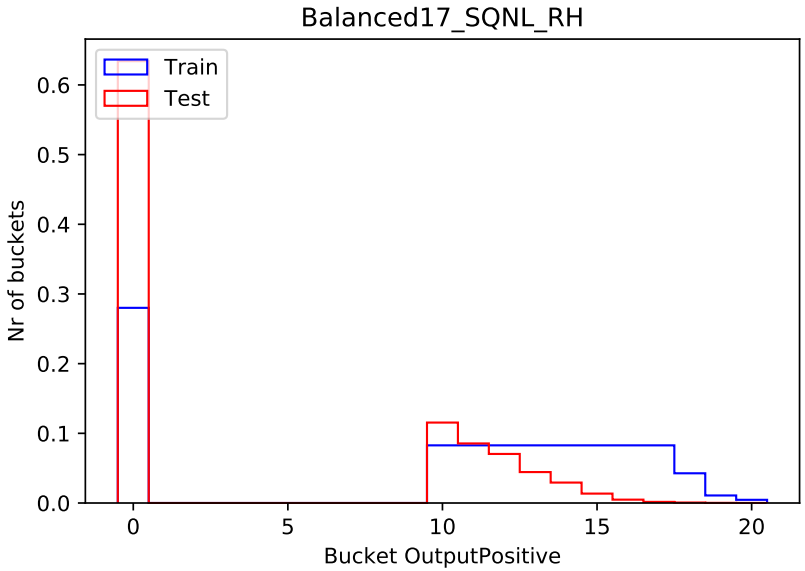


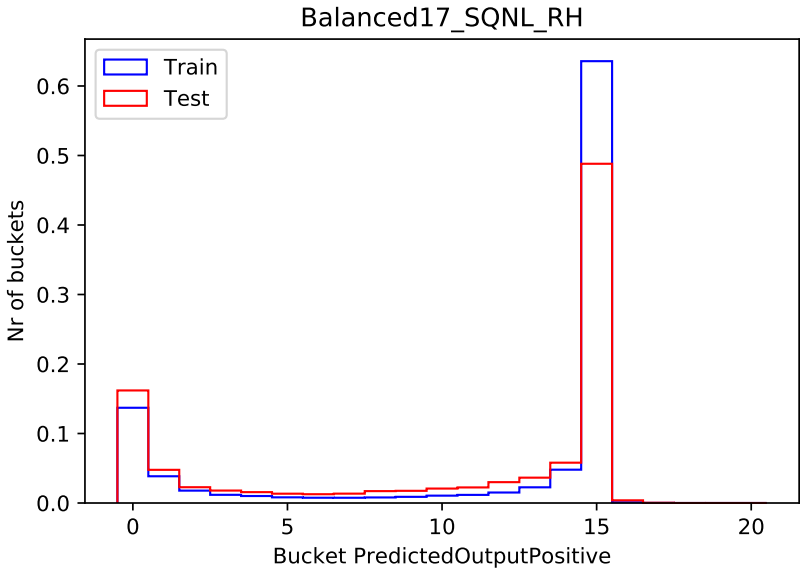


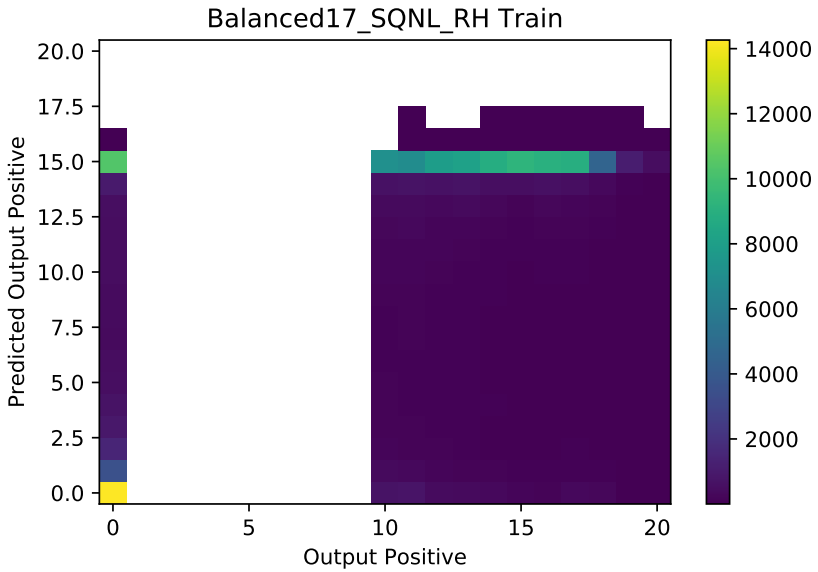


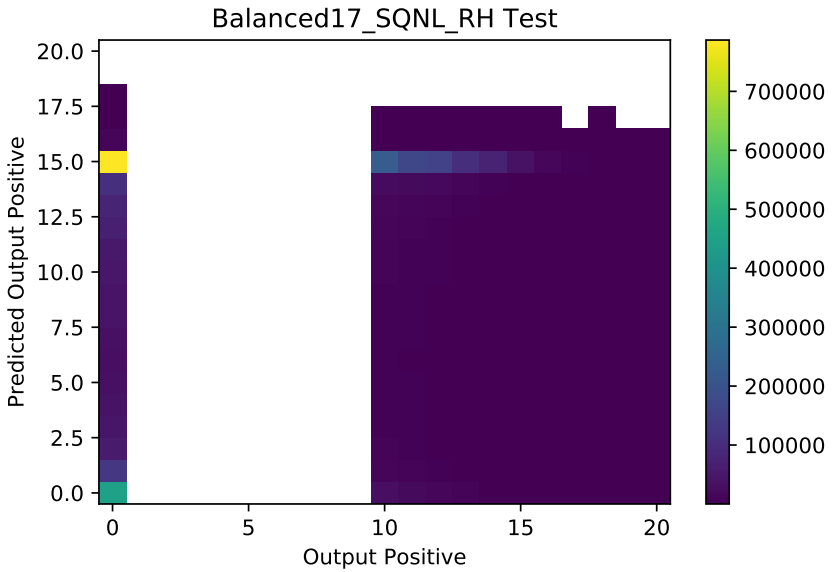


2D plots SQNL RH 120

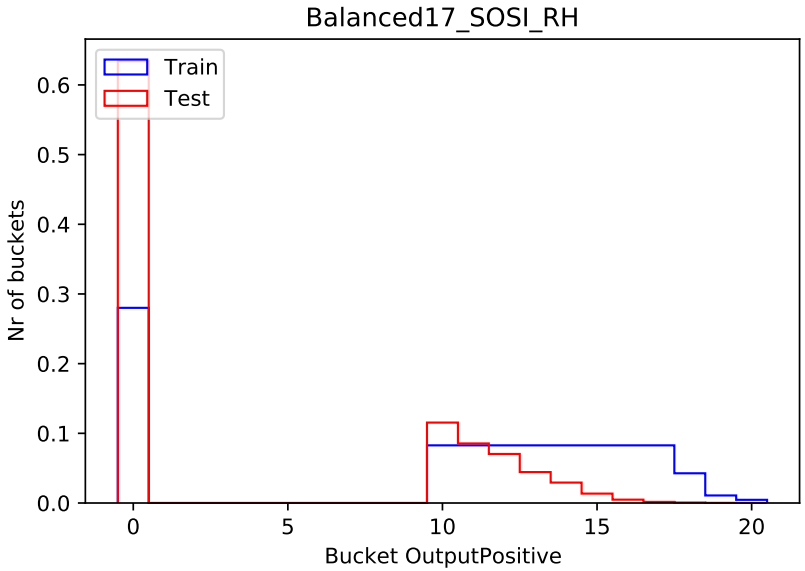


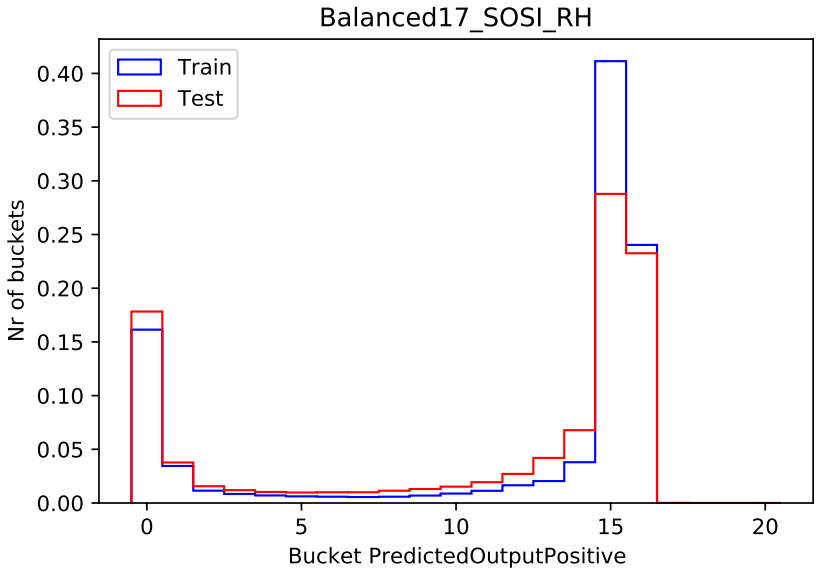


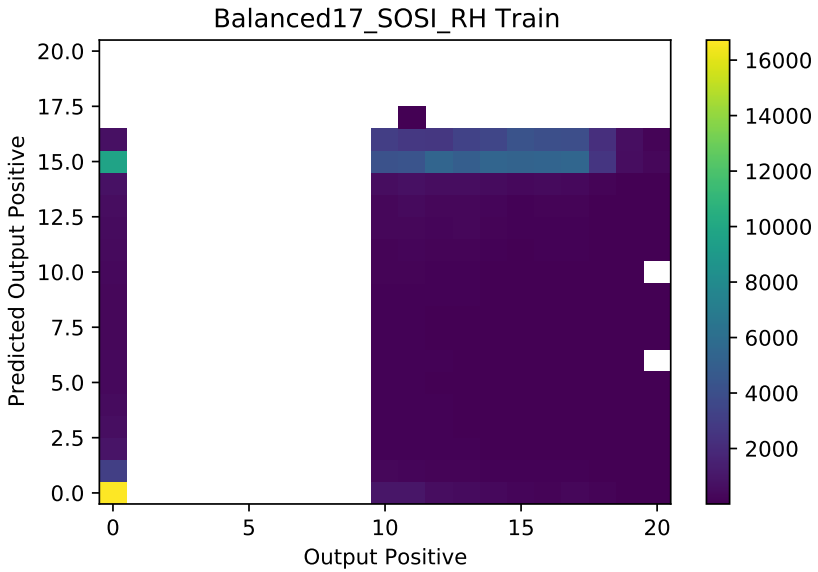


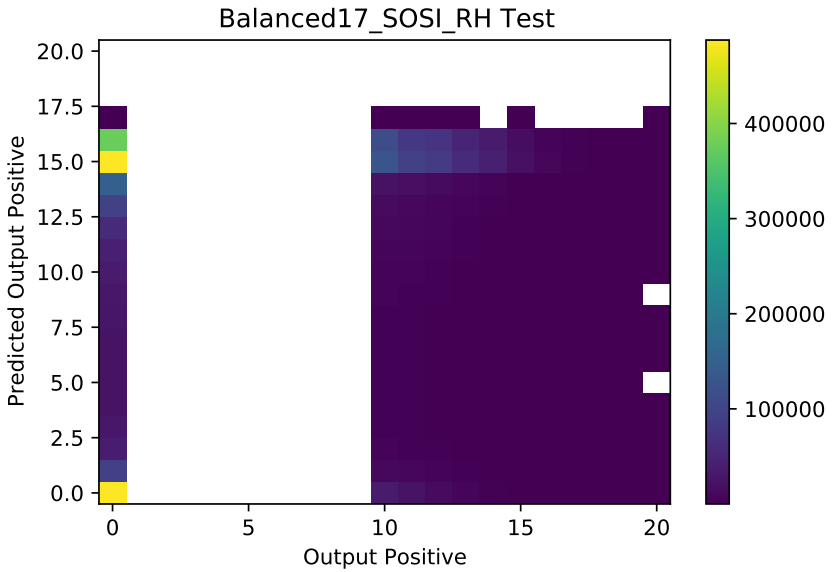


2D plots SOSI RH 120









Study 3: for 120 epochs, vary loss function

- Conclusions
- Results are similar for squared hinge and hinge.