04/19/2024 GELATO Weekly

Max Cohen

Penn

Updates from this week

- Taking a closer look at anomaly events and the reconstruction
    - The network seems to do very poorly
    - Anomalies seem to be correlated with multiplicity

- The network was doing so poorly that I conducted a few tests to understand this further
    - Training over Ztautau
    - Training with a bigger network
    - Training without padding in the loss or the AD score

A closer look at anomalies

Example 1:

```
event:
[[[ 0.18945431  0.4822886  -0.30662119]
  [ 0.14006709  0.0747449   2.48758316]
  [ 0.10698286 -4.40890455  2.29233575]
  [ 0.10339661 -3.51014018 -0.84333491]
  [ 0.08084805  4.40199375 -2.6154201 ]
  [ 0.05505495  3.06775713 -2.3277216 ]
  [ 0.05356225  4.37638283  0.78126425]
  [ 0.04481367 -2.55221438 -0.62268186]
  [ 0.04245868  1.20024025  2.29764962]
  [ 0.03930756  2.72266507  3.05115509]
  [ 0.          0.          0.        ]
  [ 0.          0.          0.        ]
  [ 0.          0.          0.        ]
  [ 0.          0.          0.        ]
  [ 0.          0.          0.        ]
  [ 0.          0.          0.        ]
  [ 0.06237572  0.03229186  2.51828933]
  [ 0.06237572  0.03229186  2.51828933]
  [ 0.01930252  0.49109831 -0.34072194]
  [ 0.          0.          0.        ]]]
```

```
reconstruction:
[[[ 0.07721651 -2.3509588   0.15923315]
  [ 0.09922105 -1.2720616   3.5524993 ]
  [ 0.08455377 -3.1708102  -0.55961454]
  [ 0.08034864 -2.2868004   1.5503771 ]
  [ 0.07941206 -4.665706   -1.0439769 ]
  [ 0.07485644 -3.398923   -0.3168471 ]
  [ 0.07140426 -4.121166    0.00983873]
  [ 0.06736635 -3.144613   -0.42728975]
  [ 0.06484393 -2.0061798  -0.45480698]
  [ 0.06224328 -1.5839988  -0.27955362]
  [ 0.07932642 -0.80624723  2.311302  ]
  [-0.00563699 -1.5004848   1.2698274 ]
  [ 0.510247    0.26875934  3.2155108 ]
  [ 0.01160312 -0.35231897  0.63435763]
  [ 0.02050136 -0.08960898  0.3395575 ]
  [ 0.03033541 -0.33366236  0.25899255]
  [ 0.03270582 -0.14314073  2.5033002 ]
  [ 0.03578089 -0.18150991  2.4510999 ]
  [ 0.02369734 -0.41831353  1.6878198 ]
  [ 0.08133683 -1.1599902  -0.2116593 ]]]
```

A closer look at anomalies

Example 2:

```
event:
[[[ 0.21395209 -3.50761843  1.22508359]
  [ 0.08297126 -4.30362606  1.22082376]
  [ 0.08168574  4.16133118 -1.68941939]
  [ 0.07821483 -3.7948966  -1.90423143]
  [ 0.07677587  0.07985659 -1.69648123]
  [ 0.07105421 -2.05840516  2.74635196]
  [ 0.06233085 -4.58569002  3.04448748]
  [ 0.06131445 -4.19437599  2.63759494]
  [ 0.05960928 -2.22904682 -2.74107647]
  [ 0.05935701 -4.64313602 -1.22328722]
  [ 0.          0.          0.        ]
  [ 0.          0.          0.        ]
  [ 0.          0.          0.        ]
  [ 0.          0.          0.        ]
  [ 0.          0.          0.        ]
  [ 0.          0.          0.        ]
  [ 0.02013095  1.236992    2.64972663]
  [ 0.01831481  0.80900902  2.89349794]
  [ 0.01309663  1.79635572  0.7123993 ]
  [ 0.10119201  0.          0.79279119]]]
```

```
reconstruction:
[[[ 0.26940504 -2.2142043   0.82784677]
  [ 0.167596   -1.1319718   0.67020804]
  [ 0.11622629  2.5782375   0.5933505 ]
  [ 0.10308573 -4.7130303  -0.41124877]
  [ 0.08614404  0.11469989  0.42081082]
  [ 0.07444128  2.4524264  -0.2210842 ]
  [ 0.06280955  2.18891    -0.06317818]
  [ 0.05588151  0.7704496   0.00557563]
  [ 0.05215904  0.54938245 -0.08000983]
  [ 0.04694394  0.20393932 -0.11935087]
  [ 0.11315463 -1.4846318   1.557439  ]
  [ 0.01474254 -1.5232285   1.1909747 ]
  [ 0.08901621 -1.814821    1.7658699 ]
  [ 0.06608011 -0.5555796   0.06962078]
  [ 0.08023279 -0.43822476 -0.10499026]
  [ 0.10256653 -0.8053799  -0.4873036 ]
  [ 0.0832797  -1.033901    1.7371817 ]
  [ 0.06058853 -1.0926914   1.783702  ]
  [ 0.03569566 -0.7976829   0.7120389 ]
  [ 0.08548179  1.6195887  -0.10486908]]]
```

A closer look at anomalies

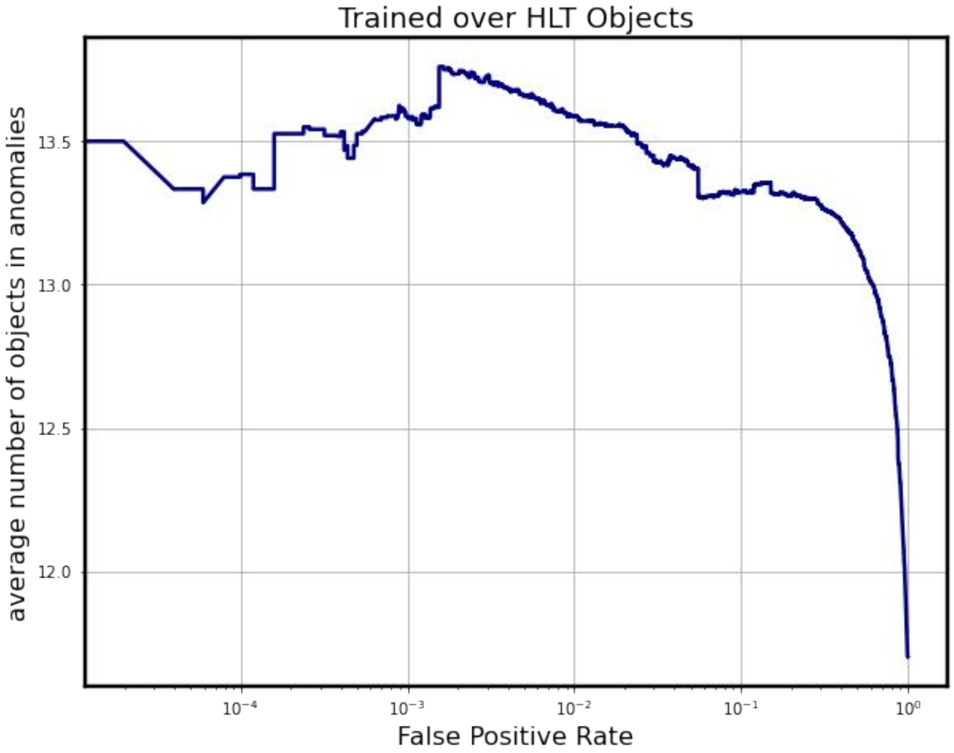
Example 3:

```
event:
[[[ 0.35784967  2.42339802  0.22014889]
  [ 0.13097649  0.82441962 -2.80768776]
  [ 0.12141488  0.24363174  3.08500719]
  [ 0.06966665 -4.06942606 -2.08009696]
  [ 0.05586545  4.62202978  0.52986825]
  [ 0.0534029   4.3804183  -1.01299679]
  [ 0.05219175 -3.57342696 -1.95683408]
  [ 0.03930232 -1.41790509 -2.29966307]
  [ 0.03509849 -2.41148853 -2.62070775]
  [ 0.03324422 -2.3036406   2.16999936]
  [ 0.          0.          0.        ]
  [ 0.          0.          0.        ]
  [ 0.          0.          0.        ]
  [ 0.01693964  0.22537266 -3.12981367]
  [ 0.          0.          0.        ]
  [ 0.          0.          0.        ]
  [ 0.0118221   0.23994373  3.14154172]
  [ 0.0118221   0.23994373  3.14154172]
  [ 0.          0.          0.        ]
  [ 0.01040335  0.          0.78312886]]]
```

```
reconstruction:
[[[ 3.9305717e-02  3.1200204e+00  5.8477545e-01]
  [ 9.5240787e-02 -3.5604084e+00 -1.5456468e-02]
  [ 9.0089358e-02  4.4188509e+00 -1.2309506e+00]
  [ 9.5047869e-02  4.4721823e+00  7.6676607e-01]
  [ 9.2653990e-02  4.9043503e+00 -1.2031083e+00]
  [ 9.0351768e-02  4.9699178e-01 -1.6164252e-01]
  [ 8.5072577e-02 -4.2029576e+00  3.3758929e-01]
  [ 8.0406517e-02 -5.1678813e-01 -5.1651448e-02]
  [ 7.5475812e-02  2.5767875e-01 -4.5509994e-01]
  [ 6.9280580e-02 -7.4544466e-01  2.6863626e-01]
  [ 6.7399137e-02  1.4028430e+00  6.7950714e-01]
  [-8.3197214e-02  1.1752083e+00  3.3030117e-01]
  [ 1.0614657e+00 -1.0375603e+00  8.2497489e-01]
  [ 9.7243190e-03  6.4878923e-01 -2.0337856e-01]
  [ 3.7610263e-02 -2.1761549e-01 -4.2223656e-01]
  [ 8.4569298e-02  9.9972039e-01 -2.7338648e-01]
  [-3.8064048e-03  8.1545663e-01  9.0011215e-01]
  [ 1.5678771e-02  8.6523509e-01  9.5954943e-01]
  [ 2.0135742e-02  6.1568207e-01  6.8408871e-01]
  [ 6.9518521e-02  1.8717062e+00 -4.6388734e-02]]]
```

Anomalies and Multiplicity

Anomalies seem to be correlated multiplicity



I’m not convinced I understand what the model is doing/learning

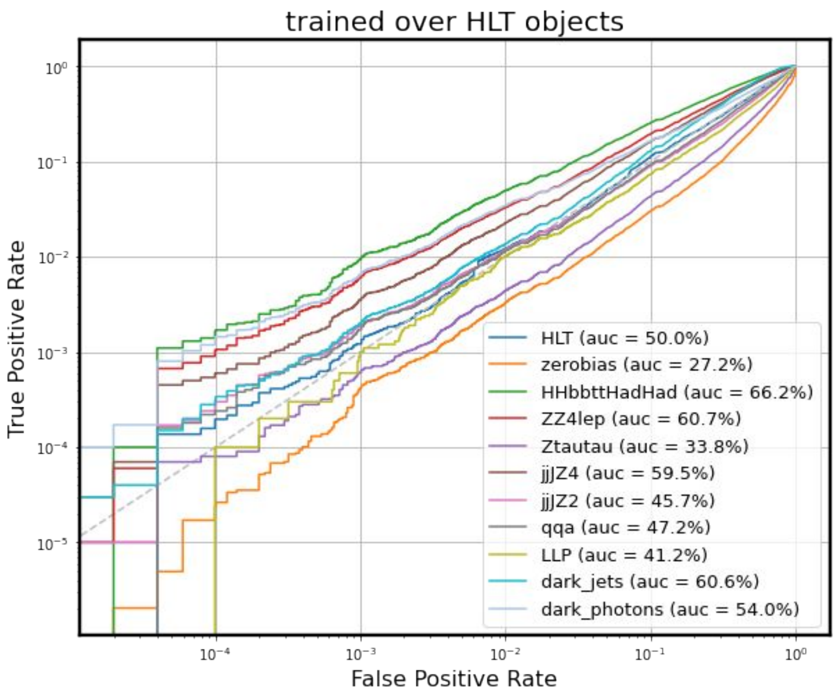
So I started running a few tests…

Training with a larger network

Maybe the EB events are more complicated than the 40MHz dataset, and therefore we need a larger encoder and decoder

So I tried a new architecture: 60 → 128 → 64 → 32 → 3

- Val loss was slightly lower during training
- Network still performs similarly...

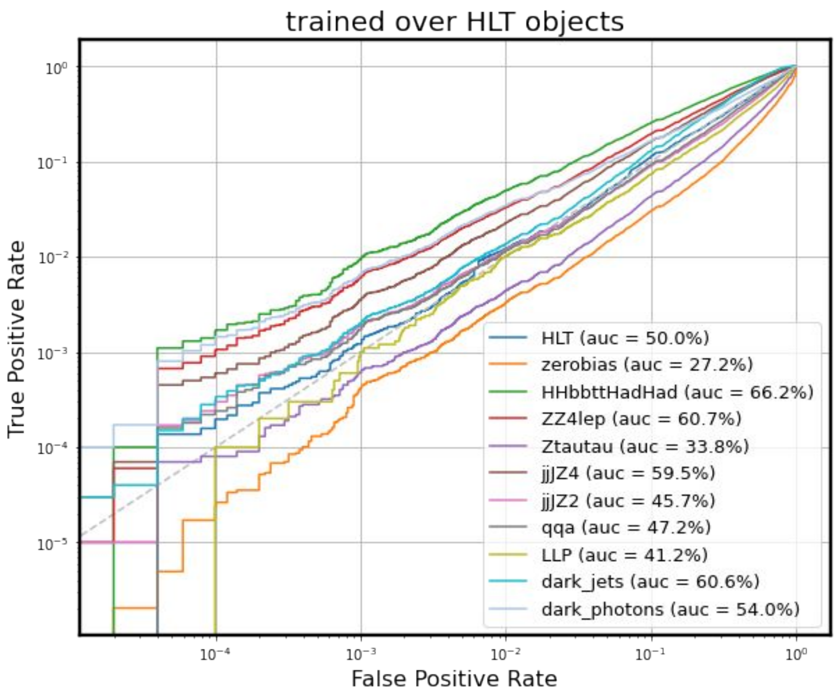


Training with a larger network

Maybe the EB events are more complicated than the 40MHz dataset, and therefore we need a larger encoder and decoder

So I tried a new architecture: 60 → 128 → 64 → 32 → 3

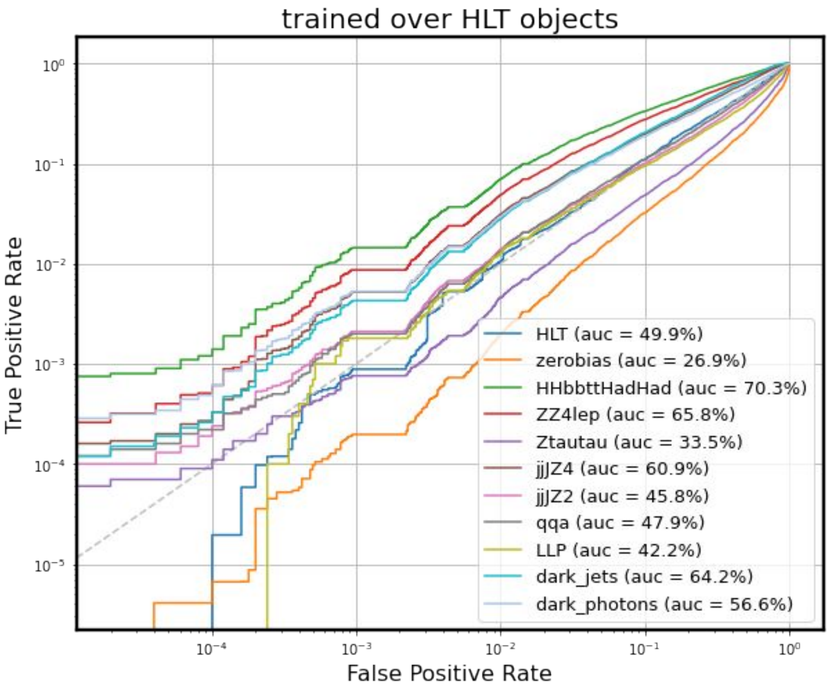
- Val loss was slightly lower during training
- Network still performs similarly…



Training with a larger network

I also tried 60 → 128 → 64 → 32 → 8

Still not great...

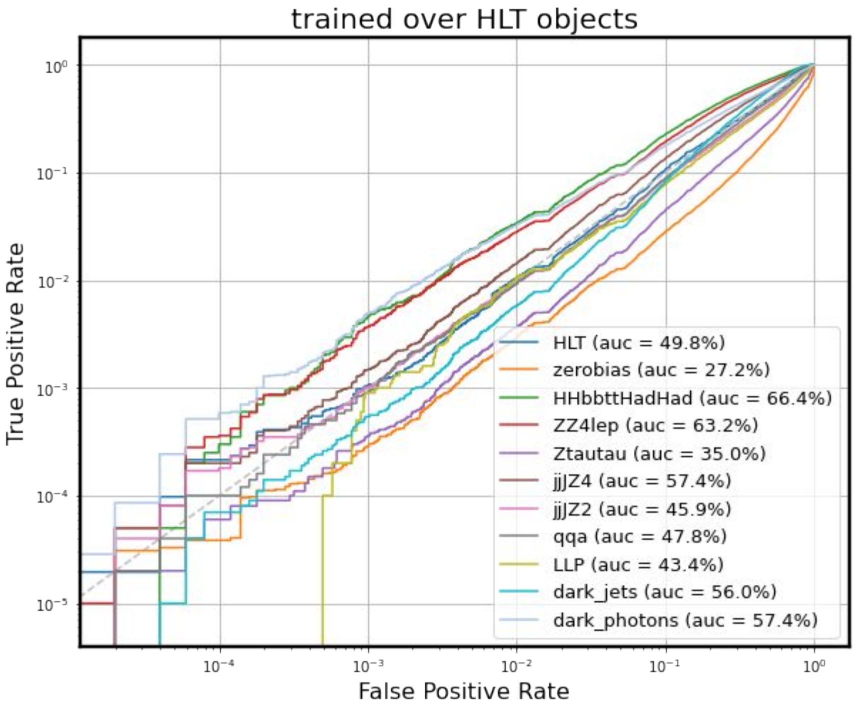


Training without zero padding

- Recall that I have been using masked MSE loss, where I am zeroing out parts of the MSE corresponding to missing objects in the input

- I tried training without this mask, hoping to force the network to learn more

- Network seems to learn nothing

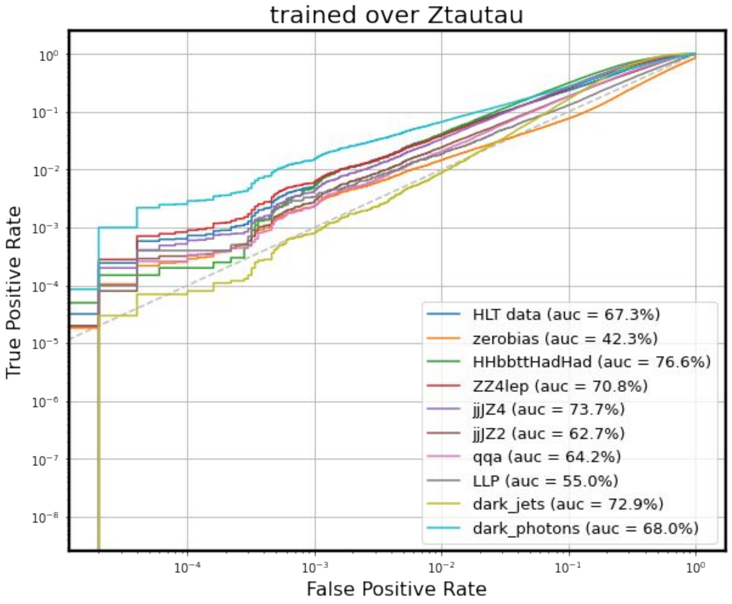


Training over Ztautau

- None of these results are very encouraging. Maybe I can do a test to ensure there are no bugs in the code

- Accordingly, I tried training over Ztautau, which should be very different than many of the other signals.
    - It shouldn't be hard for the network to perform well over some of the signals

- Still no good performance over any signal



So what's going on?

- Is there some bug in my code?
  - I'm not sure where this bug would be, since the input datasets look okay, and the network is very straightforward
  - Maybe some error keeping track of the events when calculating loss / AD scores?
  - I've spend a bit of time looking and haven't found anything
  - I also tried the 1000 year old technique of pasting the code into Chat GPT and asking "what's the bug in this code", with no success
  - Also worth noting that the network trained over L1 objects has substantially lower val loss during training than the network trained over HLT objects

- Any advice?