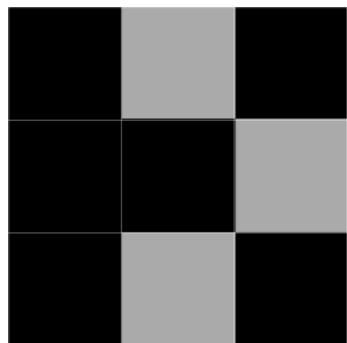
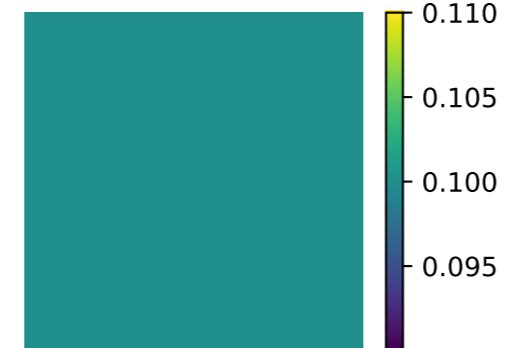


CNN Feature Analysis: Grid 3x3

Original Input



One-Hot Embedding
10 channels



Grid Size: 3x3 | CNN: 3 conv layers | Hidden Dim: 128

Layer Statistics:

EMBED: Ch=10, Size=(3, 3), Mean=0.100±0.213, Range=[0.00,1.00]
Diversity=0.4444, Sparsity=0.900, Entropy=0.29

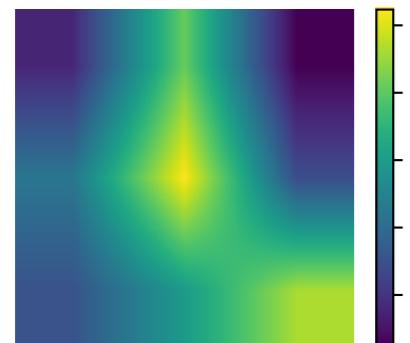
CONV1: Ch=32, Size=(3, 3), Mean=0.578±0.344, Range=[0.00,4.86]
Diversity=0.6151, Sparsity=0.559, Entropy=1.07

CONV2: Ch=32, Size=(3, 3), Mean=0.273±0.264, Range=[0.00,3.17]
Diversity=0.5980, Sparsity=0.670, Entropy=0.78

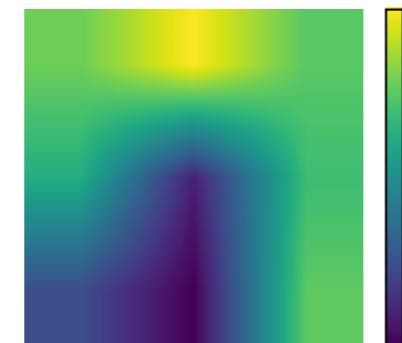
CONV3: Ch=32, Size=(3, 3), Mean=0.510±0.340, Range=[0.00,2.25]
Diversity=0.1267, Sparsity=0.174, Entropy=1.73

SPATIAL_FEATURES: Ch=32, Size=(3, 3), Mean=0.510±0.340, Range=[0.00,2.25]
Diversity=0.1267, Sparsity=0.174, Entropy=1.73

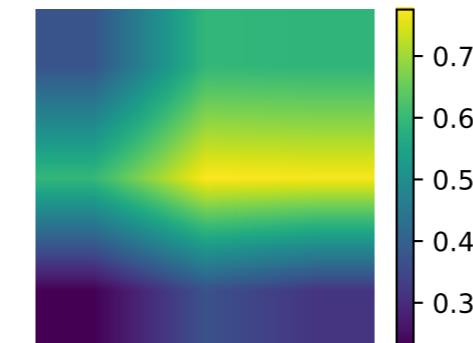
CONV1 Average
1 channels



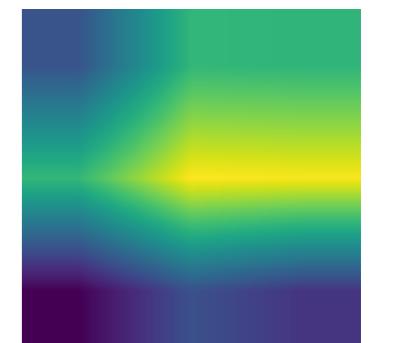
CONV2 Average
1 channels



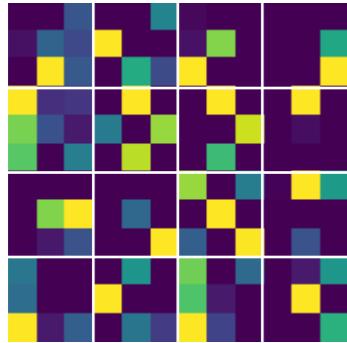
CONV3 Average
1 channels



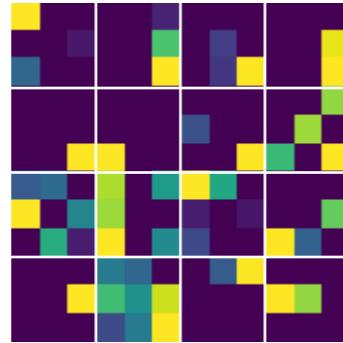
Spatial Features (avg)
→ Input to Slot Attention



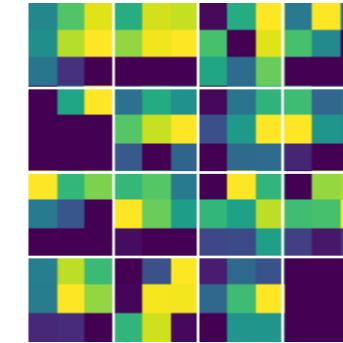
Feature Channels (showing 16/32)



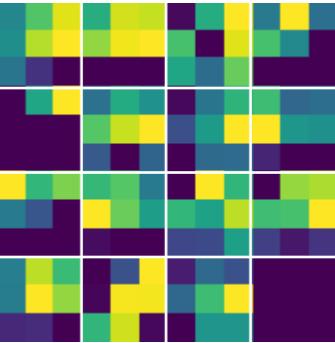
Feature Channels (showing 16/32)



Feature Channels (showing 16/32)



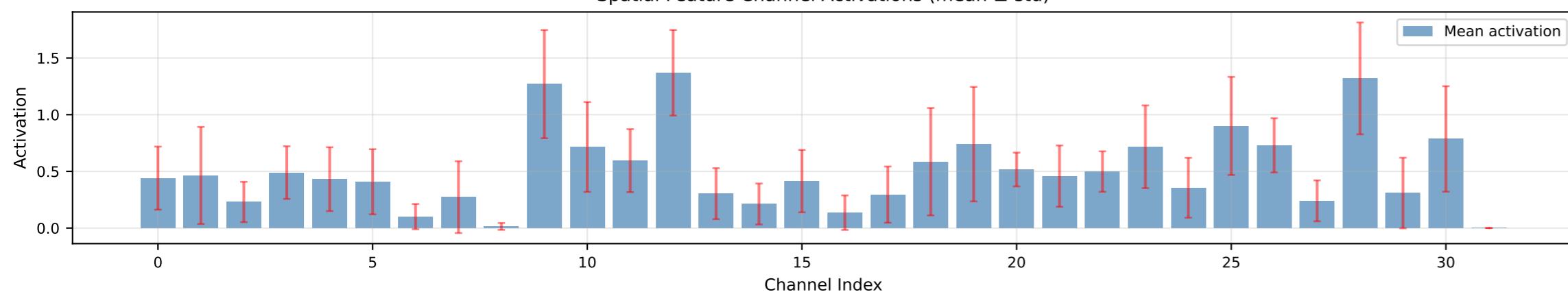
Feature Channels (showing 16/32)



DIAGNOSIS SUMMARY:

- GOOD SPARSITY
- △ LOW ENTROPY: Low information

Spatial Feature Channel Activations (mean ± std)



Spatial Location Similarity
(sample of 9 locations)

