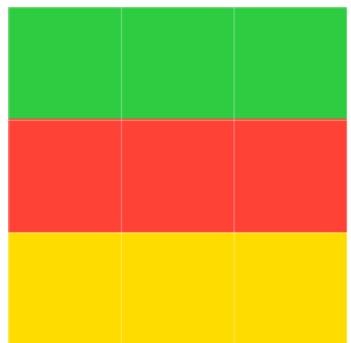
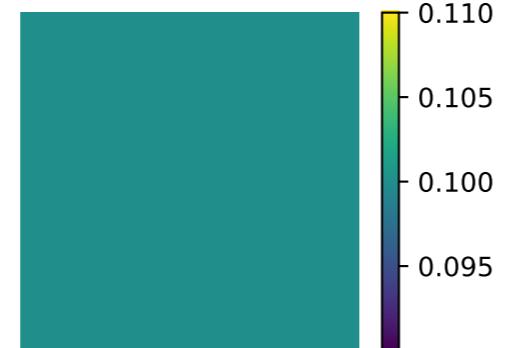


CNN Feature Analysis: Grid 3x3

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±0.153, Range=[0.00,1.00]
Diversity=0.6667, Sparsity=0.900, Entropy=0.33

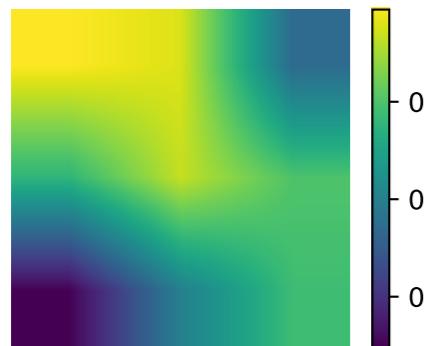
CONV1: Ch=32, Size=(3, 3), Mean=0.568±0.484, Range=[0.00,5.47]
Diversity=0.6053, Sparsity=0.628, Entropy=0.89

CONV2: Ch=32, Size=(3, 3), Mean=0.330±0.322, Range=[0.00,3.31]
Diversity=0.5726, Sparsity=0.649, Entropy=0.83

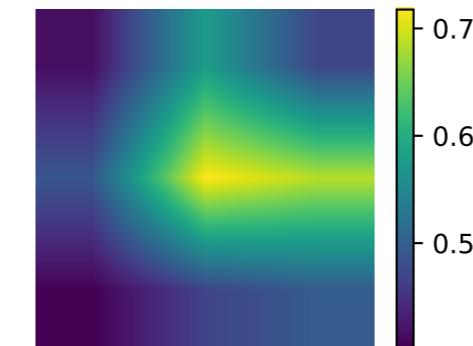
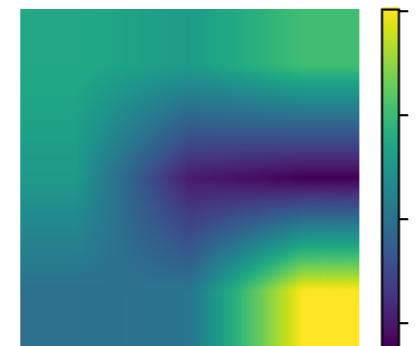
CONV3: Ch=32, Size=(3, 3), Mean=0.524±0.350, Range=[0.00,2.54]
Diversity=0.1875, Sparsity=0.188, Entropy=1.74

SPATIAL_FEATURES: Ch=32, Size=(3, 3), Mean=0.524±0.350, Range=[0.00,2.54]
Diversity=0.1875, Sparsity=0.188, Entropy=1.74

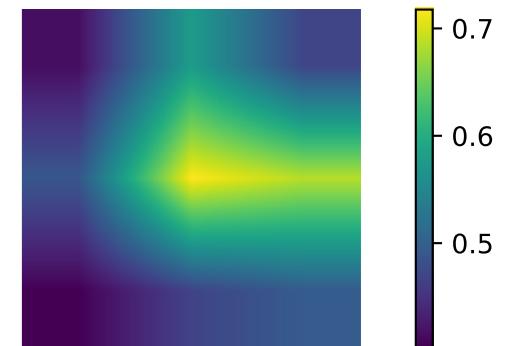
CONV1 Average
1 channels



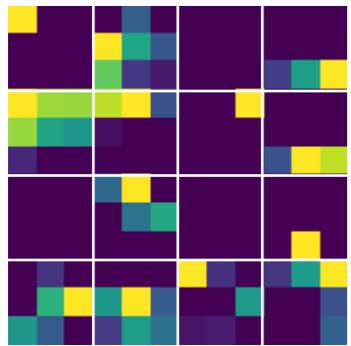
CONV2 Average
1 channels



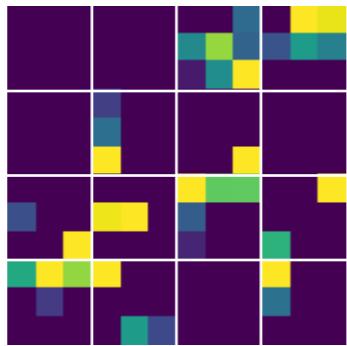
Spatial Features (avg)
→ Input to Slot Attention



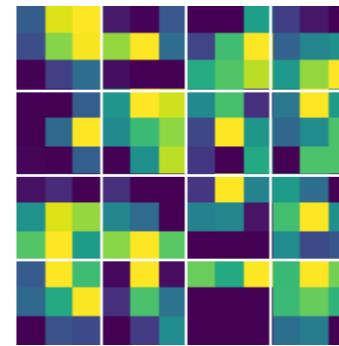
Feature Channels (showing 16/32)



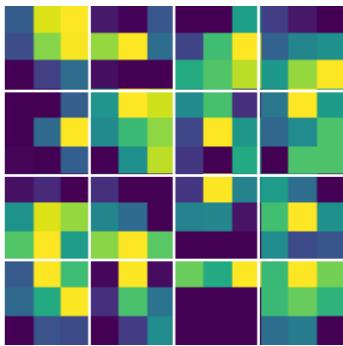
Feature Channels (showing 16/32)



Feature Channels (showing 16/32)



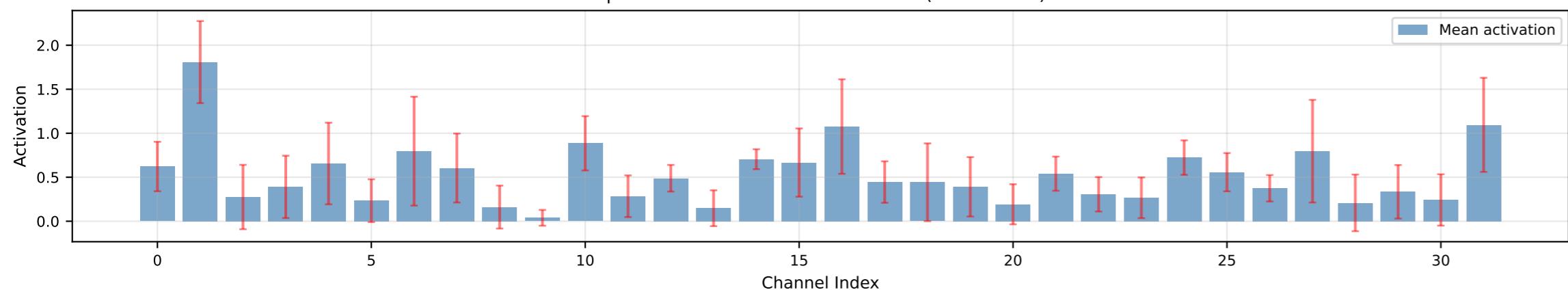
Feature Channels (showing 16/32)



DIAGNOSIS SUMMARY:

- GOOD SPARSITY
- △ LOW ENTROPY: Low information

Spatial Feature Channel Activations (mean ± std)



Spatial Location Similarity
(sample of 9 locations)

