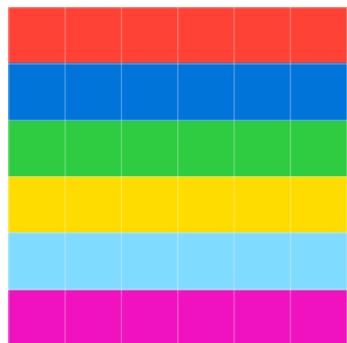
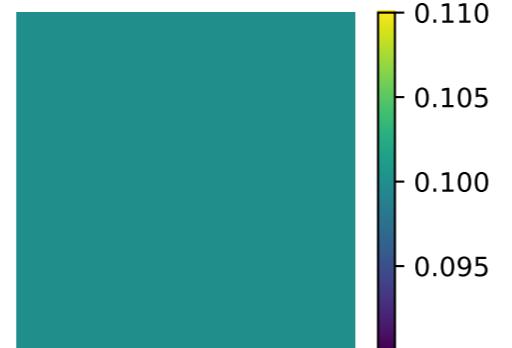


# CNN Feature Analysis: Grid 6x6

**Original Input**



One-Hot Embedding  
10 channels



Grid Size: 6x6 | CNN: 3 conv layers | Hidden Dim: 128

Layer Statistics:

EMBED: Ch=10, Size=(6, 6), Mean=0.100±0.082, Range=[0.00,1.00]  
Diversity=0.8333, Sparsity=0.900, Entropy=1.08

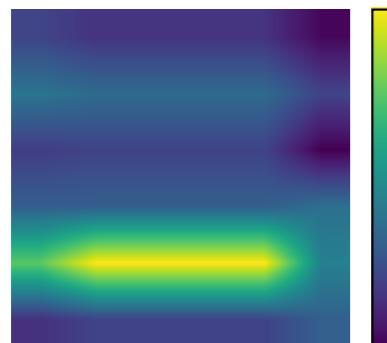
CONV1: Ch=32, Size=(6, 6), Mean=0.665±0.490, Range=[0.00,6.43]  
Diversity=0.6479, Sparsity=0.637, Entropy=2.18

CONV2: Ch=32, Size=(6, 6), Mean=0.412±0.318, Range=[0.00,6.43]  
Diversity=0.6859, Sparsity=0.666, Entropy=2.12

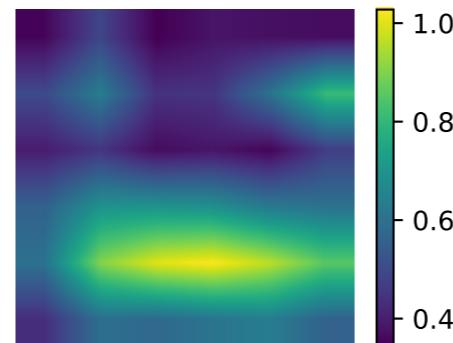
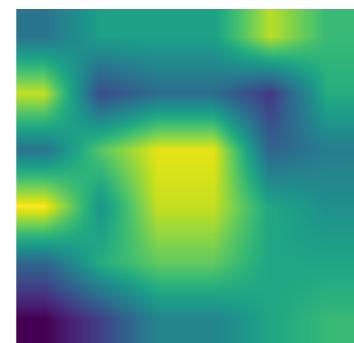
CONV3: Ch=32, Size=(6, 6), Mean=0.575±0.360, Range=[0.00,3.91]  
Diversity=0.4365, Sparsity=0.391, Entropy=2.80

SPATIAL\_FEATURES: Ch=32, Size=(6, 6), Mean=0.575±0.360, Range=[0.00,3.91]  
Diversity=0.4365, Sparsity=0.391, Entropy=2.80

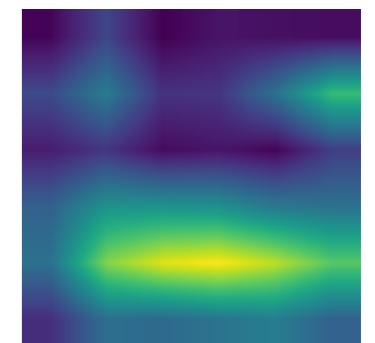
CONV1 Average  
1 channels



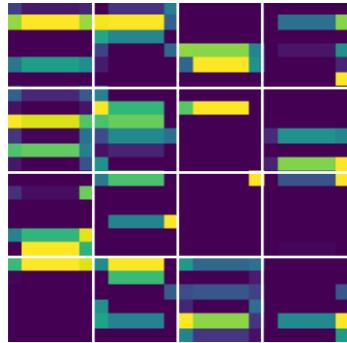
CONV2 Average  
1 channels



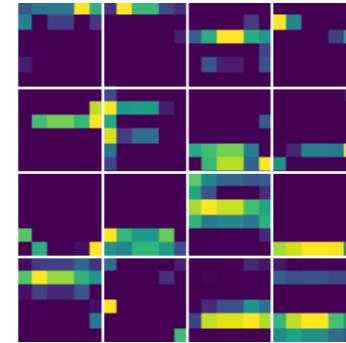
Spatial Features (avg)  
→ Input to Slot Attention



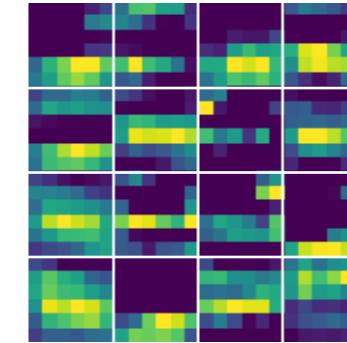
Feature Channels (showing 16/32)



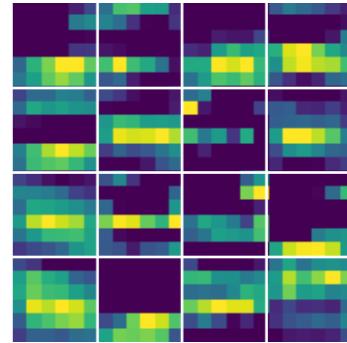
Feature Channels (showing 16/32)



Feature Channels (showing 16/32)



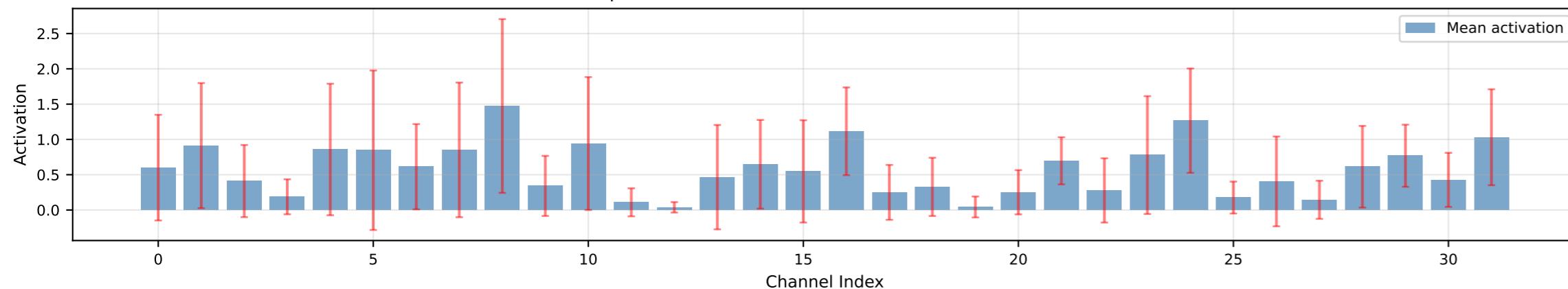
Feature Channels (showing 16/32)



## DIAGNOSIS SUMMARY:

- GOOD SPARSITY
- GOOD ENTROPY

Spatial Feature Channel Activations (mean ± std)



Spatial Location Similarity  
(sample of 36 locations)

