Normalized confusion matrix (gpt-4o-mini) action -0.45 0.02 0.01 0.03 0.03 0.02 0.16 0.05 0.00 0.05 0.14 0.03 0.02 aggression -0.09 0.73 0.00 0.00 0.05 0.00 0.05 0.00 0.05 0.00 0.05 0.00 0.00 entrance -0.00 0.00 0.00 0.00 0.87 0.03 0.00 0.00 0.00 0.05 0.00 0.05 0.00 exit -0.07 0.00 0.00 0.00 0.11 0.78 0.03 0.00 0.00 0.01 0.00 0.00 0.00 True label interaction -0.00 0.10 0.00 0.00 0.03 0.00 <mark>0.87</mark> 0.00 0.00 0.00 0.00 0.00 0.00 movement -0.17 0.03 0.00 0.00 0.06 0.17 0.19 0.36 0.00 0.00 0.00 0.00 0.03 music -0.01 0.00 0.00 0.05 0.02 0.00 0.01 0.00 0.83 0.04 0.02 0.02 0.01 narration -0.08 0.11 0.00 0.06 0.03 0.00 0.03 0.03 0.00 0.56 0.06 0.06 0.00 object -0.16 0.02 0.00 0.03 0.00 0.00 0.21 0.00 0.00 0.02 0.55 0.00 0.02 setting -0.05 0.00 0.00 0.00 0.02 0.02 0.04 0.09 0.00 0.07 0.00 0.72 0.00 action movement interaction MUSIC narration Predicted label