Normalized confusion matrix (gpt-4o-mini) action -0.40 0.01 0.00 0.03 0.00 0.02 0.23 0.08 0.00 0.06 0.12 0.03 0.01 delivery -0.06 0.00 0.02 0.67 0.02 0.02 0.08 0.00 0.02 0.11 0.00 0.00 0.02 exit -0.00 0.00 0.00 0.00 0.12 0.84 0.01 0.03 0.00 0.00 0.00 0.00 0.00 True label interaction -0.00 0.10 0.00 0.00 0.00 0.00 <mark>0.87</mark> 0.03 0.00 0.00 0.00 0.00 0.00 movement -0.08 0.03 0.00 0.00 0.00 0.17 0.17 0.56 0.00 0.00 0.00 0.00 0.00 object -0.05 0.02 0.00 0.00 0.00 0.00 0.19 0.00 0.00 0.03 0.71 0.00 0.00 setting -0.04 0.00 0.02 0.00 0.00 0.02 0.05 0.12 0.00 0.05 0.00 0.70 0.00 actionession movement interaction narration MUSIC Predicted label