Normalized confusion matrix (gpt-4o-mini) action -0.49 0.02 0.00 0.03 0.01 0.02 0.19 0.02 0.00 0.06 0.14 0.01 0.00 delivery -0.05 0.02 0.02 0.50 0.02 0.02 0.14 0.00 0.12 0.12 0.00 0.00 0.00 exit -0.00 0.00 0.00 0.00 0.10 0.82 0.07 0.01 0.00 0.00 0.00 0.00 0.00 True label interaction -0.00 0.23 0.00 0.00 0.03 0.00 0.71 0.03 0.00 0.00 0.00 0.00 0.00 movement -0.06 0.06 0.00 0.00 0.00 0.14 0.17 0.58 0.00 0.00 0.00 0.00 0.00 setting -0.02 0.00 0.02 0.00 0.00 0.02 0.05 0.16 0.00 0.04 0.00 0.70 0.00 action movement interaction narration MUSIC Predicted label