Normalized confusion matrix (gpt-4o-mini) action -0.62 0.02 0.00 0.01 0.01 0.01 0.11 0.08 0.01 0.04 0.04 0.03 0.01 delivery -0.12 0.00 0.02 0.47 0.03 0.02 0.17 0.03 0.05 0.06 0.00 0.02 0.02 entrance -0.03 0.00 0.00 0.00 0.76 0.05 0.03 0.00 0.03 0.00 0.08 0.00 exit -0.01 0.00 0.00 0.00 0.14 0.74 0.04 0.05 0.00 0.01 0.00 0.00 0.00 True label interaction -0.00 0.26 0.00 0.00 0.00 0.00 <mark>0.74</mark> 0.00 0.00 0.00 0.00 0.00 0.00 movement -0.08 0.00 0.00 0.00 0.06 0.17 0.11 0.53 0.00 0.03 0.00 0.00 0.03 music -0.01 0.00 0.02 0.00 0.01 0.00 0.01 0.00 0.88 0.02 0.01 0.03 0.01 narration -0.11 0.08 0.00 0.00 0.03 0.00 0.19 0.03 0.03 0.44 0.00 0.08 0.00 setting -0.04 0.00 0.02 0.00 0.02 0.04 0.00 0.12 0.00 0.04 0.00 0.74 0.00 toward -0.00 0.00 0.00 0.01 0.03 0.00 0.08 0.01 0.00 0.00 0.00 0.12 0.75 actionession movement interaction narration MUSIC

Predicted label