Normalized confusion matrix (gpt-4o) action -0.68 0.01 0.00 0.03 0.02 0.00 0.12 0.03 0.00 0.05 0.04 0.02 0.00 delivery -0.09 0.00 0.03 0.66 0.00 0.02 0.03 0.00 0.05 0.11 0.00 0.00 0.02 exit -0.04 0.00 0.00 0.00 0.10 0.68 0.03 0.11 0.00 0.04 0.00 0.00 0.00 True label interaction -0.00 0.03 0.00 0.00 0.00 0.00 <mark>0.97</mark> 0.00 0.00 0.00 0.00 0.00 0.00 movement -0.06 0.00 0.00 0.00 0.03 0.14 0.11 0.56 0.00 0.06 0.00 0.00 0.06 narration -0.11 0.00 0.00 0.03 0.00 0.00 0.03 0.00 0.00 0.75 0.03 0.06 0.00 object -0.18 0.02 0.00 0.02 0.00 0.00 0.15 0.00 0.00 0.00 0.63 0.02 0.00 setting -0.05 0.00 0.02 0.00 0.00 0.02 0.02 0.09 0.00 0.05 0.00 0.74 0.02 action adjoestion movement interaction narration Music

Predicted label