Normalized confusion matrix (gpt-4o) action -0.65 0.01 0.00 0.03 0.03 0.01 0.08 0.05 0.00 0.05 0.06 0.02 0.00 delivery -0.11 0.00 0.03 0.64 0.00 0.02 0.03 0.00 0.05 0.11 0.00 0.00 0.02 exit +0.03 0.00 0.00 0.00 0.10 0.73 0.01 0.11 0.00 0.03 0.00 0.00 0.00 True label interaction -0.03 0.06 0.00 0.00 0.00 0.00 <mark>0.90</mark> 0.00 0.00 0.00 0.00 0.00 0.00 movement -0.06 0.00 0.00 0.00 0.00 0.14 0.14 0.56 0.00 0.06 0.00 0.00 0.06 narration -0.06 0.00 0.00 0.03 0.00 0.00 0.00 0.03 0.00 0.81 0.03 0.06 0.00 setting -0.07 0.00 0.02 0.00 0.00 0.04 0.00 0.11 0.00 0.04 0.00 0.74 0.00 toward $+0.01\ 0.00\ 0.01\ 0.00\ 0.00\ 0.01\ 0.00\ 0.00\ 0.00\ 0.00\ 0.00\ 0.00$ action adjoestion movement interaction narration Music

Predicted label