## Normalized confusion matrix (gpt-4o) action -0.72 0.02 0.00 0.01 0.03 0.01 0.12 0.02 0.00 0.01 0.04 0.02 0.00 aggression -0.05 0.86 0.00 0.00 0.00 0.00 0.00 0.00 0.05 0.05 0.00 0.00 delivery -0.12 0.02 0.00 0.58 0.05 0.03 0.05 0.02 0.08 0.05 0.00 0.00 0.02 exit +0.01 0.01 0.00 0.00 0.12 0.82 0.01 0.01 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>0.90</mark> 0.00 0.00 0.00 0.00 0.00 0.00 movement -0.03 0.00 0.00 0.00 0.06 0.19 0.14 0.56 0.00 0.00 0.00 0.00 0.03 narration -0.14 0.08 0.00 0.03 0.00 0.03 0.06 0.08 0.00 0.47 0.03 0.08 0.00 object -0.10 0.02 0.00 0.00 0.00 0.00 0.15 0.00 0.00 0.02 0.73 0.00 0.00 setting -0.04 0.00 0.02 0.00 0.02 0.00 0.12 0.00 0.02 0.00 0.77 0.00 toward +0.01 0.01 0.00 0.01 0.02 0.00 0.04 0.01 0.00 0.00 0.00 0.02 0.87 action adjoestion

apaite

movement interaction narration MUSIC Predicted label