Normalized confusion matrix (gpt-4o-mini) action -0.29 0.03 0.01 0.04 0.06 0.01 0.26 0.07 0.00 0.05 0.16 0.01 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 -0.11 0.03 0.03 0.39 0.03 0.02 0.14 0.05 0.08 0.09 0.00 0.02 0.02 entrance -0.00 0.00 0.00 0.00 0.89 0.05 0.00 0.00 0.00 0.00 0.00 0.05 0.00 exit -0.08 0.03 0.00 0.01 0.15 0.67 0.04 0.01 0.00 0.00 0.00 0.00 0.00 True label interaction -0.00 0.16 0.00 0.00 0.00 0.00 <mark>0.81</mark> 0.00 0.00 0.00 0.03 0.00 0.00 movement -0.03 0.00 0.00 0.00 0.06 0.19 0.17 0.53 0.00 0.00 0.03 0.00 0.00 music -0.02 0.01 0.09 0.07 0.02 0.00 0.02 0.01 0.68 0.03 0.01 0.03 0.01 narration -0.11 0.03 0.00 0.00 0.03 0.00 0.06 0.03 0.11 0.47 0.08 0.08 0.00 object -0.06 0.03 0.00 0.16 0.00 0.00 0.13 0.00 0.00 0.02 0.60 0.00 0.00 setting -0.02 0.00 0.02 0.00 0.02 0.04 0.00 0.18 0.00 0.02 0.00 0.72 0.00 toward -0.01 0.01 0.11 0.02 0.14 0.00 0.11 0.01 0.00 0.00 0.00 0.04 0.54 action movement interaction narration Music apaite Predicted label