Normalized confusion matrix (gpt-4o-mini) action -0.64 0.02 0.00 0.03 0.02 0.01 0.16 0.05 0.01 0.01 0.01 0.02 0.01 aggression -0.09 0.86 0.00 0.00 0.00 0.00 0.00 0.00 0.05 0.00 0.00 0.00 delivery -0.11 0.02 0.02 0.44 0.02 0.02 0.19 0.03 0.09 0.08 0.00 0.00 0.00 entrance -0.03 0.00 0.00 0.00 0.87 0.05 0.00 0.00 0.00 0.00 0.00 0.05 0.00 exit -0.00 0.01 0.00 0.00 0.15 0.81 0.03 0.00 0.00 0.00 0.00 0.00 0.00 True label interaction -0.00 0.19 0.00 0.00 0.00 0.00 <mark>0.81</mark> 0.00 0.00 0.00 0.00 0.00 0.00 movement -0.08 0.00 0.00 0.00 0.03 0.17 0.19 0.53 0.00 0.00 0.00 0.00 0.00 music -0.01 0.01 0.12 0.03 0.02 0.01 0.02 0.00 0.68 0.03 0.03 0.05 0.01 narration -0.17 0.06 0.00 0.00 0.00 0.00 0.19 0.03 0.03 0.42 0.03 0.08 0.00 object -0.29 0.02 0.00 0.24 0.00 0.00 0.19 0.02 0.00 0.02 0.23 0.00 0.00 setting -0.05 0.00 0.02 0.00 0.02 0.04 0.02 0.14 0.00 0.04 0.00 0.68 0.00 action movement interaction narration aparte MUSIC Predicted label