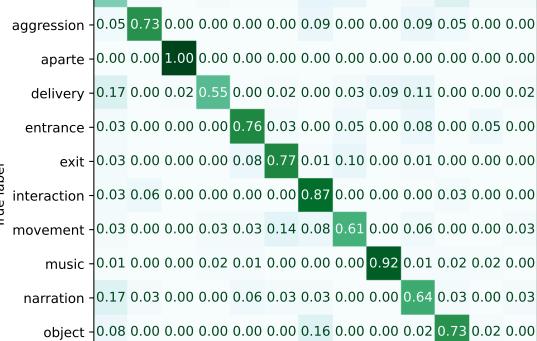
Normalized confusion matrix (mistral-large-2407) action -0.45 0.02 0.00 0.03 0.01 0.12 0.10 0.00 0.04 0.17 0.03 0.03



True label movement -0.03 0.00 0.00 0.03 0.03 0.14 0.08 0.61 0.00 0.06 0.00 0.00 0.03 object -0.08 0.00 0.00 0.00 0.00 0.16 0.00 0.00 0.02 0.73 0.02 0.00 setting -0.09 0.00 0.00 0.00 0.02 0.00 0.00 0.14 0.00 0.05 0.00 0.70 0.00 action movement interaction MUSIC narration apaite delivery Predicted label