Hyperparameter tuning: The algorithm uses a single hyperparameter α to control the heterogeneity of distances among elements that are clustered together. We used 10% of the labeled data to tune the hyperparameter, searching for optimal values on both sides of the default value suggested by prior work (α=1). Observing two peaks (A and B), we inspected the values of the hyperparameter at a finer level of granularity near these peaks (bottom charts). We selected two hyperparameter values (-0.04 and 0.475) from the analysis.

AUXILIARY MATERIAL: HYPERPARAMETER TUNING

As the above figure shows, searching on both sides of the default hyperparameter value (α=1) suggested by prior work [13], we found two peaks (A and B) for the aggregate F-Score of scale and cluster recognition. We further checked the values at a finer level of granularity near both the peaks (bottom charts). Based on the performance plots, we selected the value of -0.04 for peak A and 0.475 for peak B. We note that the range near peak B (α=0.45-0.5) is more stable, i.e., there is no sharp increase or decrease in performance near that value.