

The table presents macro F1 scores and their arithmetic mean for selected configurations in each model family.

- Unless indicated otherwise, all models have been trained for 3 epochs.
- $n_\pi$  denotes the number of random directions.
- $\text{blur} = \sqrt{\varepsilon}$ , where  $\varepsilon$  is the parameter in the entropy-regularised 1-Wasserstein cost.
- `style_dim` corresponds to the style dimension embedding.
- Intuitively,  $\tau$  penalises creating or destroying mass.
- Ensemble results: UB – Unbalanced Optimal Transport, B – Balanced OT, CL – Contrastive Learning, ep. – epoch.
- The multiplier indicates the voting weight in the ensemble.