

Appendix A. Appendix

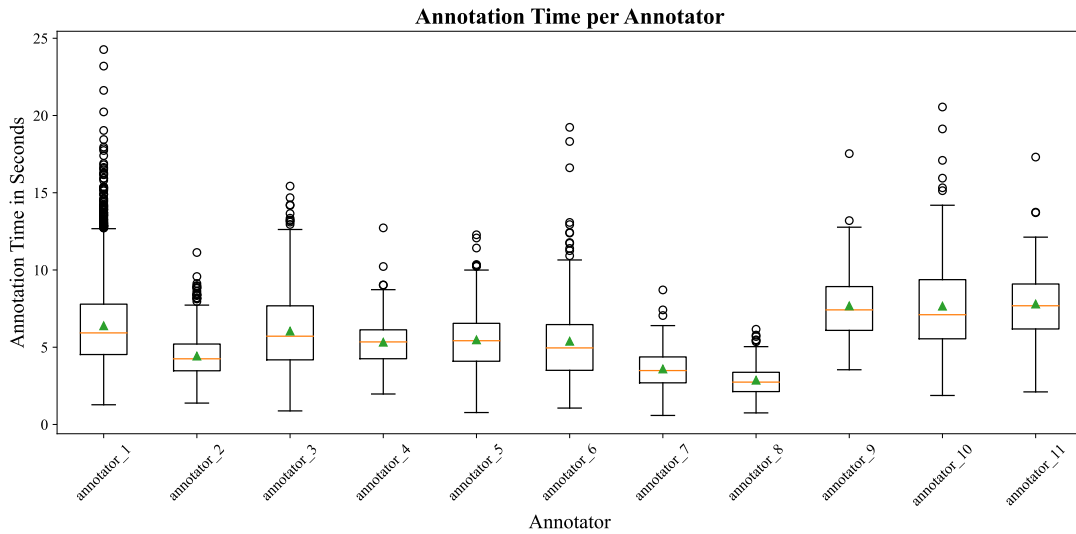


Figure A.1: Boxplot of the annotation time per human annotator for all datasets aggregated. The green triangles visualize the mean values.

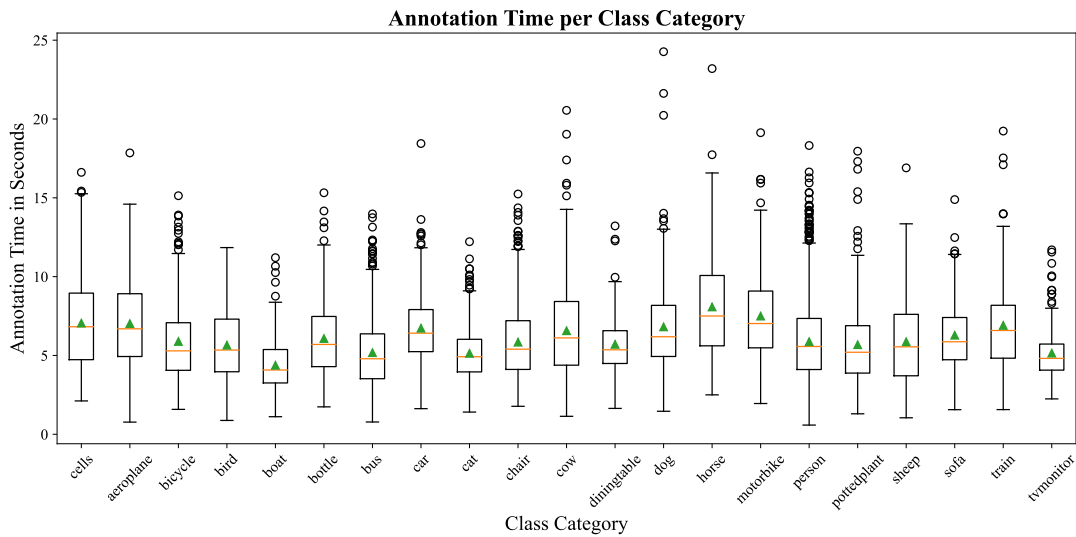


Figure A.2: Boxplot of the annotation time per class category for all datasets aggregated. The green triangles visualize the mean values.

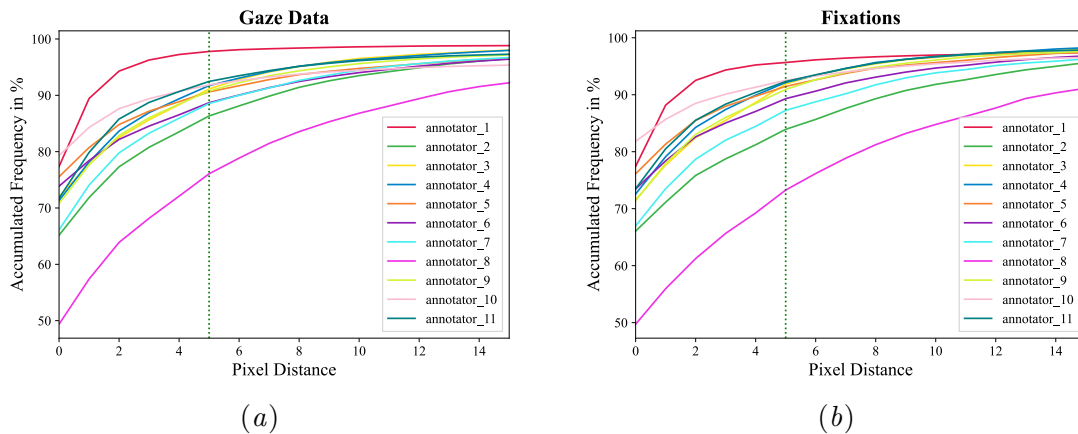


Figure A.3: (a) Raw Gaze Data and (b) Fixations grouped by their distance to the ground truth mask of all datasets grouped per human annotator.

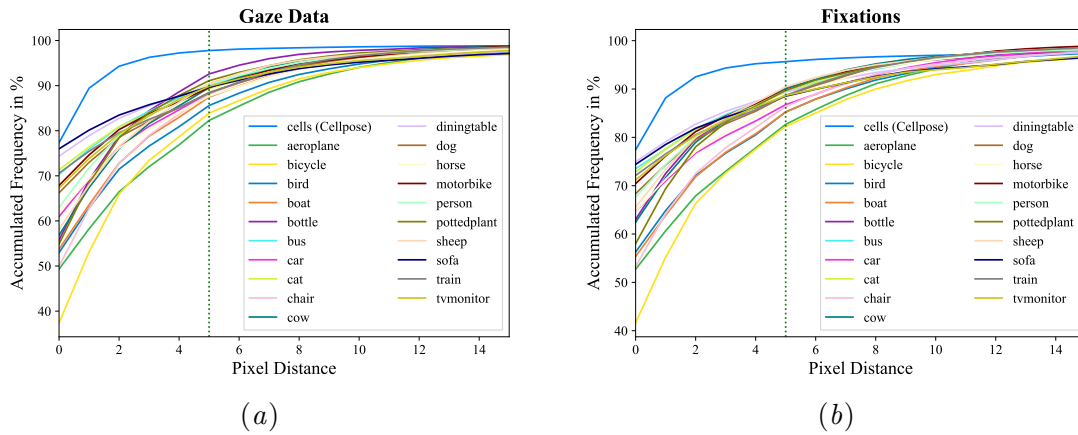


Figure A.4: (a) Raw Gaze Data and (b) Fixations grouped by their distance to the ground truth mask of all datasets grouped per class category.

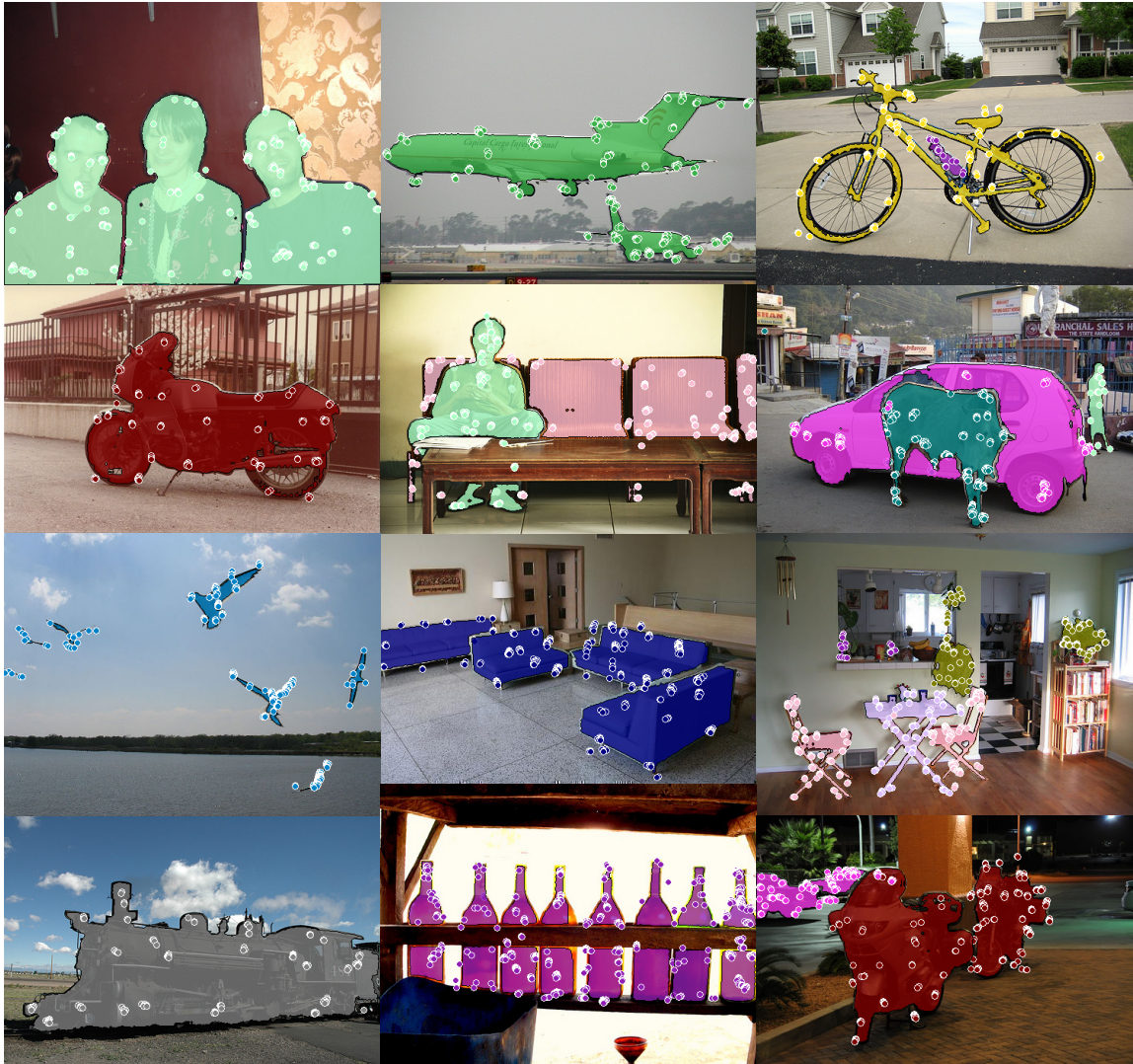


Figure A.5: Additional example results for fixation-based prompt segmentation.