

Annotation KPIs Benchmark Study

Key Facts

- 157h to annotate 1h of Egocentric Video with Ramblr's AI-powered annotation pipeline
- Quality metrics: 0.94 IOU and 0.90 F1 score achieved
- 97% of 109,370 frames annotated automatically resulting in an avg. annotation time of 5.2s per frame
- Diverse dataset with 14 different scene types
- Open vocabulary with 529 unique object instances
- Accurate mask and category annotations for all context-relevant objects on every frame

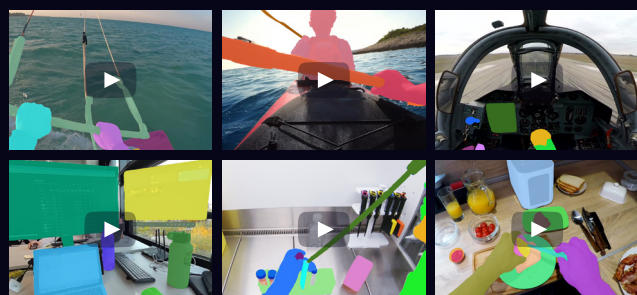
Test Dataset

- Collection devices: Meta Quest 3, GoPro, Vuzix Blade, Project Aria
- Datasets: Collected by Ramblr, Ego4D, Aria Pilot Dataset

	Video duration	Unique objects
Total	1h	529
Max per video	5min	10

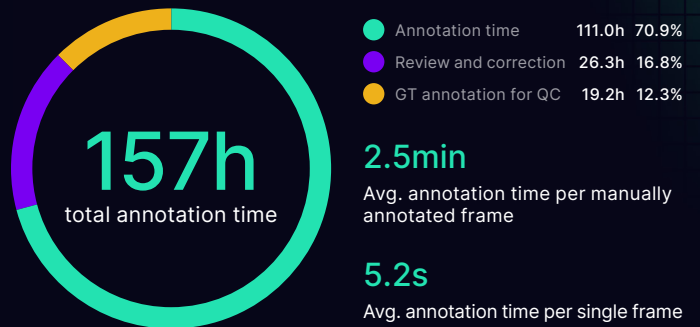
* Frame rate: 30 fps

Video Duration / Scene Type (% of total dataset)



Annotation Time

1 Hour of Egocentric Video



Video frames	Auto-annotated frames	Manually annotated frames
109,370	105,568 (97%)	3,802 (3%)

Quality

Average IOU and F1 Score for Dataset



Annotations

- Temporal consistent segmentation of object instances
- Multi-object tracking for all context relevant objects
- AI-assisted annotation guidance to minimize subjective interpretation

Annotate Context Relevant Objects

Open vocabulary: All objects the ego interacts with; Ramblr's AI-models detect hand-object interactions and gaze signals to provide annotation guidance

Closed vocabulary: Objects always considered to be relevant e.g., mobile phones and laptops. Ramblr's custom trained detection model provides annotation guidance

Do you want to hear more about Ramblr?
Reach out to us at info@ramblr.ai