

Aggregating Crowdsourced Image Segmentations

Doris Jung-Lin Lee¹, Akash Das Sarma², Aditya Parameswaran¹

¹ University of Illinois, ² Facebook Inc.



Research Overview

Goals:

- Precise segmentations around a semantic object
- Characterization of common worker errors in crowdsourced segmentation

Contributions:

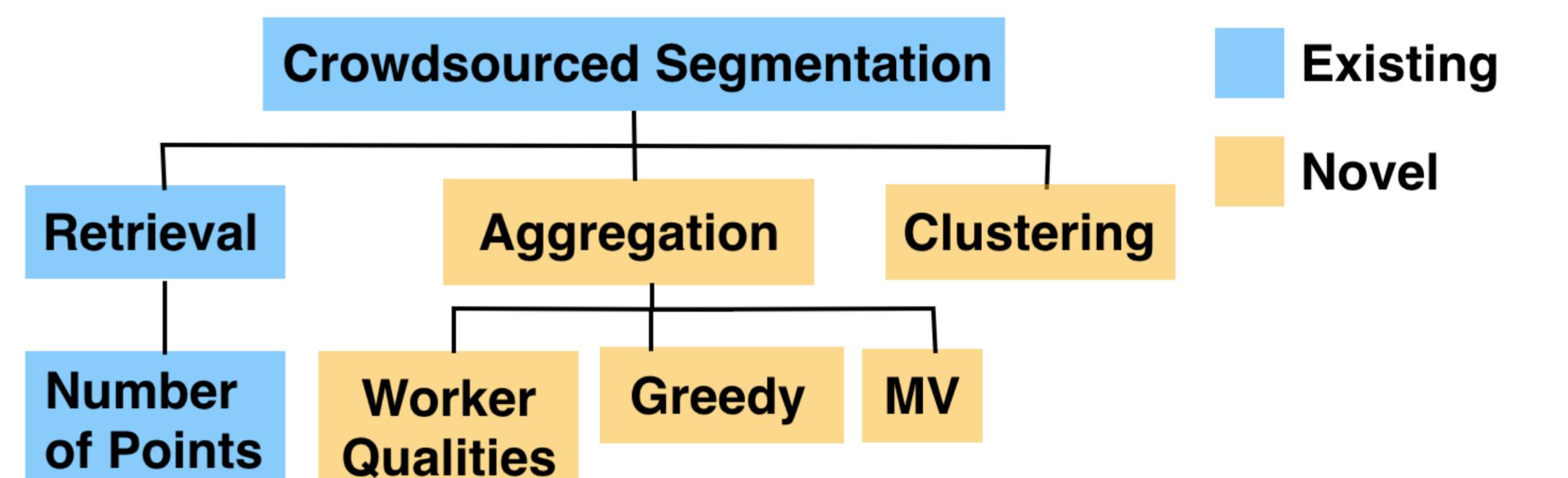
- Clustering for capturing workers with different semantic perspectives
- Novel class of aggregation-based algorithms using tile data representation



Problem Description

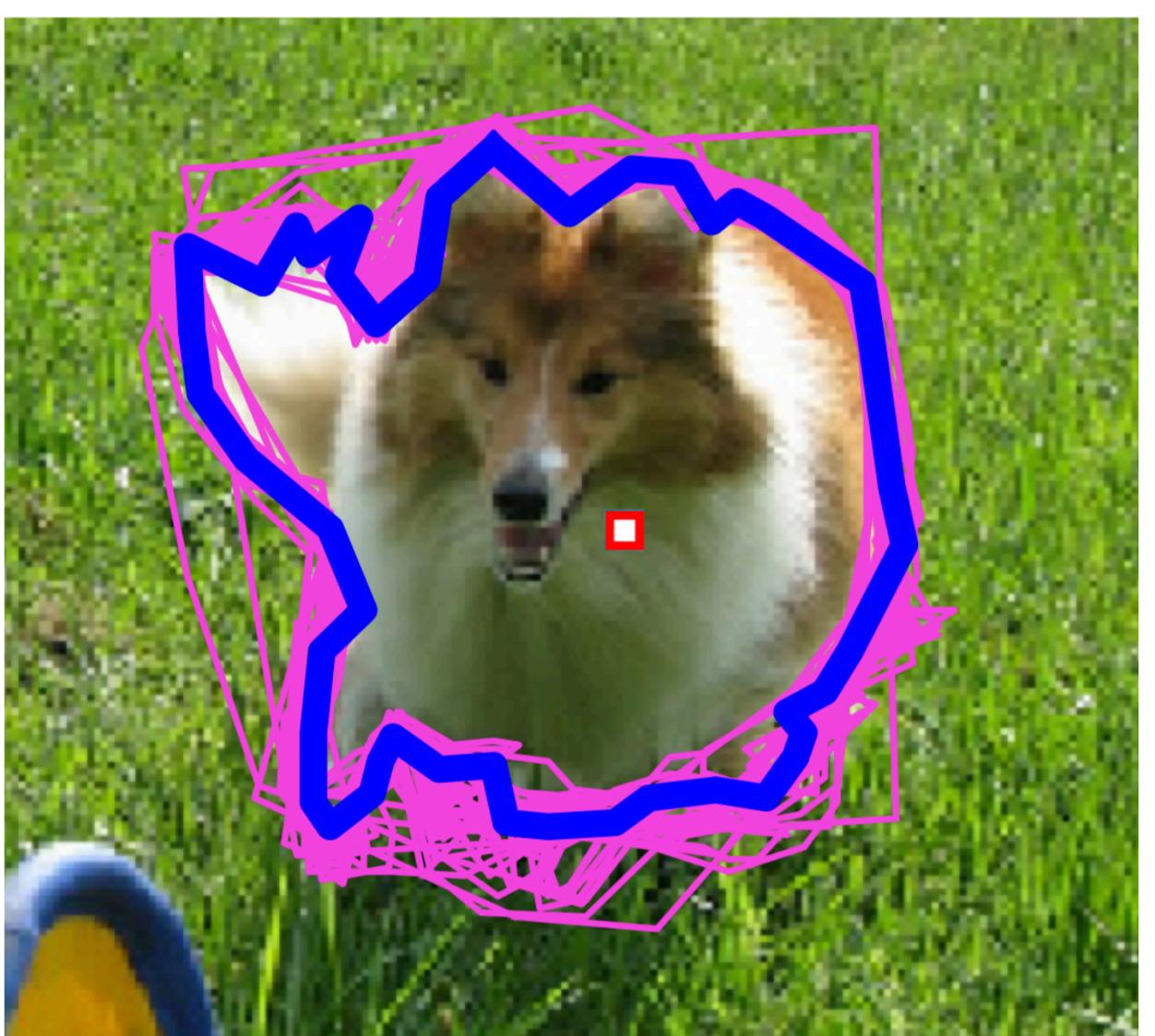
- Worker responses often noisy → quality evaluation
- Retrieval-based: pick “best” based on heuristics
- Aggregation-based: combine multiple segmentations

Taxonomy of quality evaluation algorithms



Dataset Description & Error Analysis

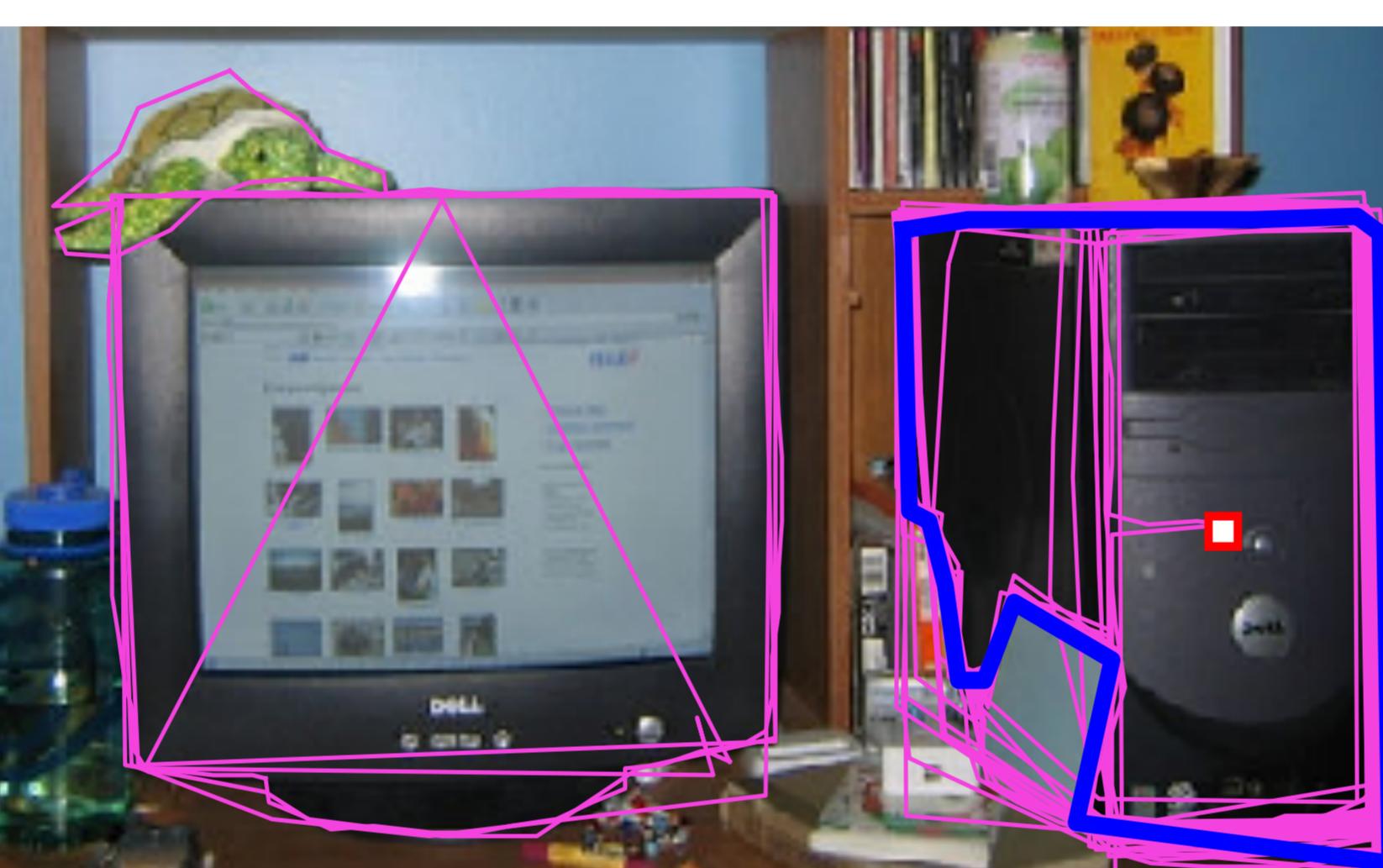
Boundary Imprecision [dog]



Semantic Ambiguity [vase]



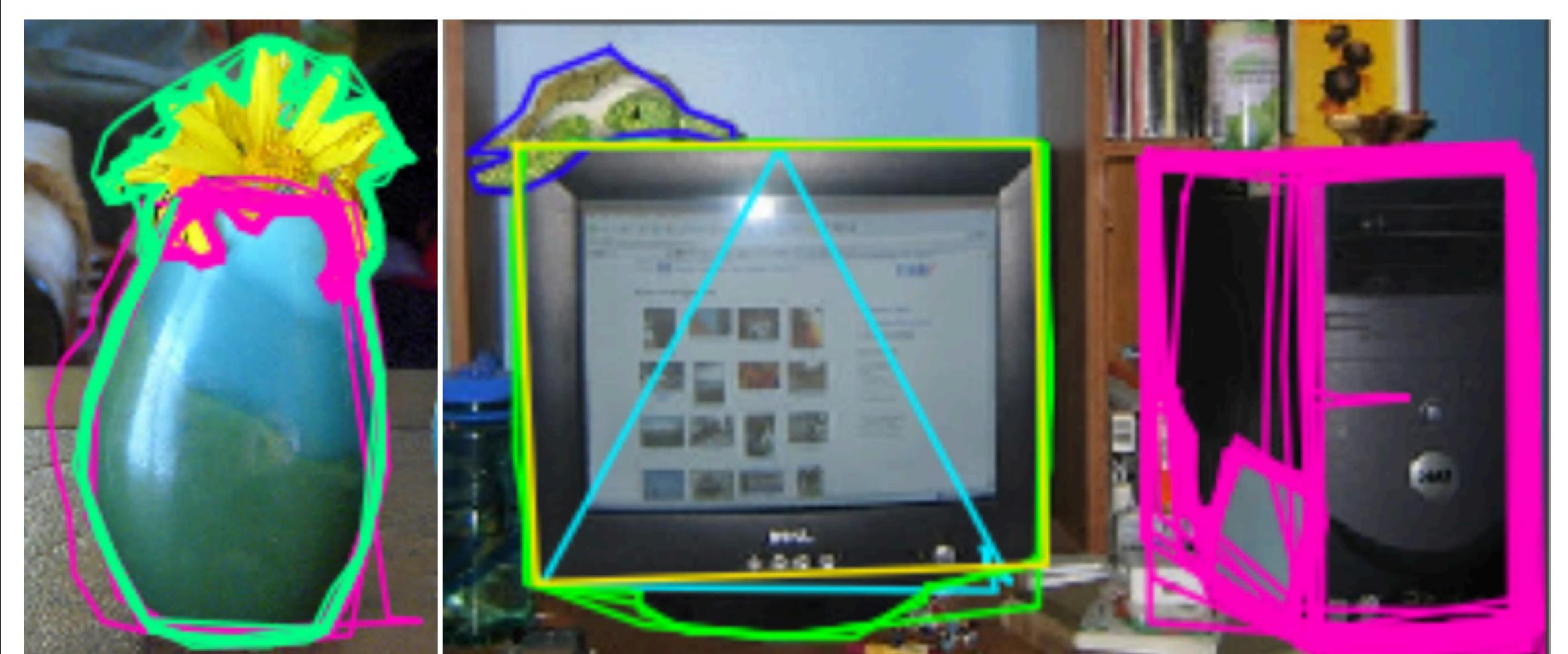
Semantic Error [computer]



- Dataset:** 46 objects from the MSCOCO, total 1840 segmentations
- Semantic ambiguity and errors:** wrong subject or portions of objects segmented
- Prior work largely focusses on addressing **bounding imprecision**.

Perspective Resolution

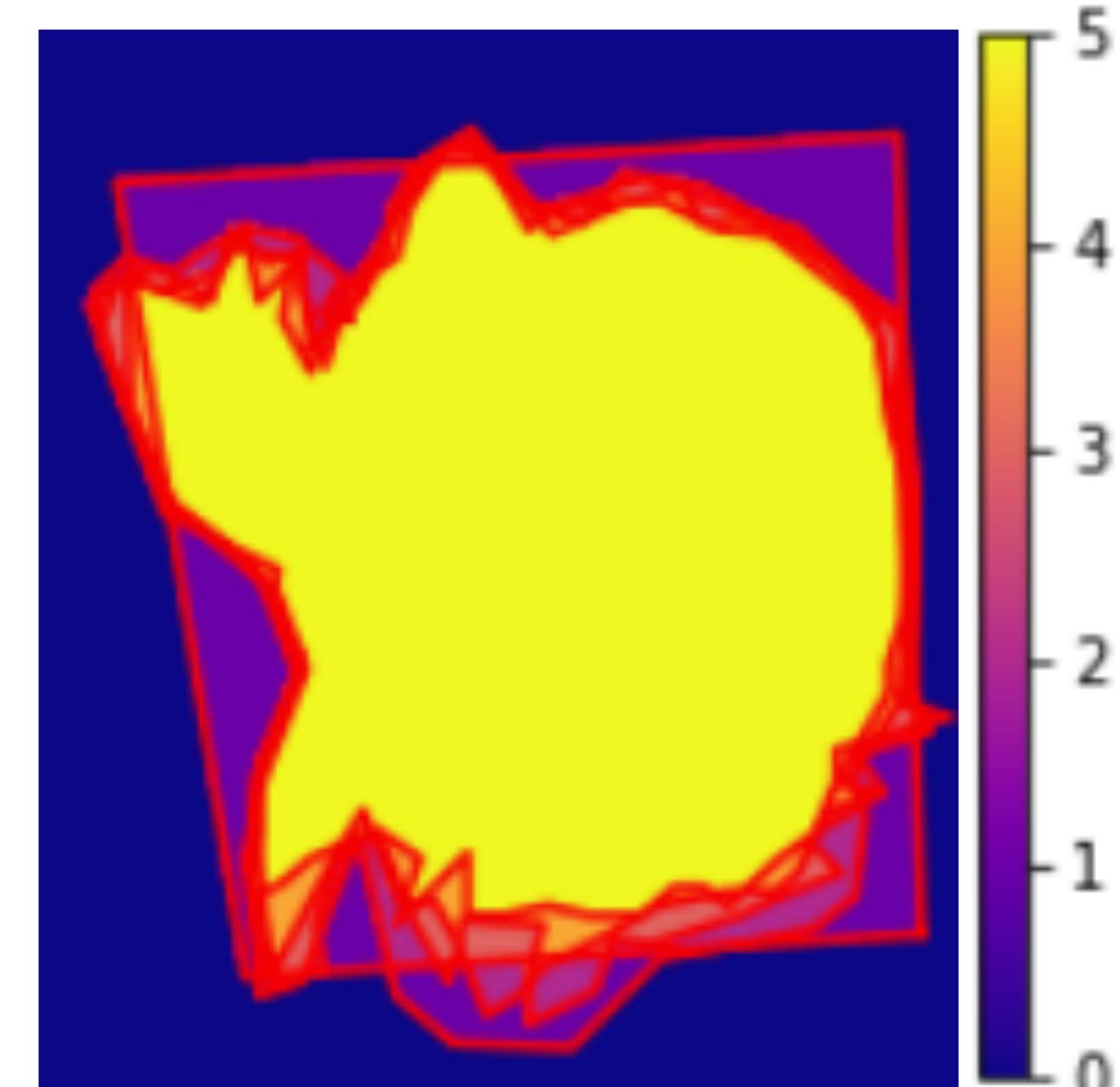
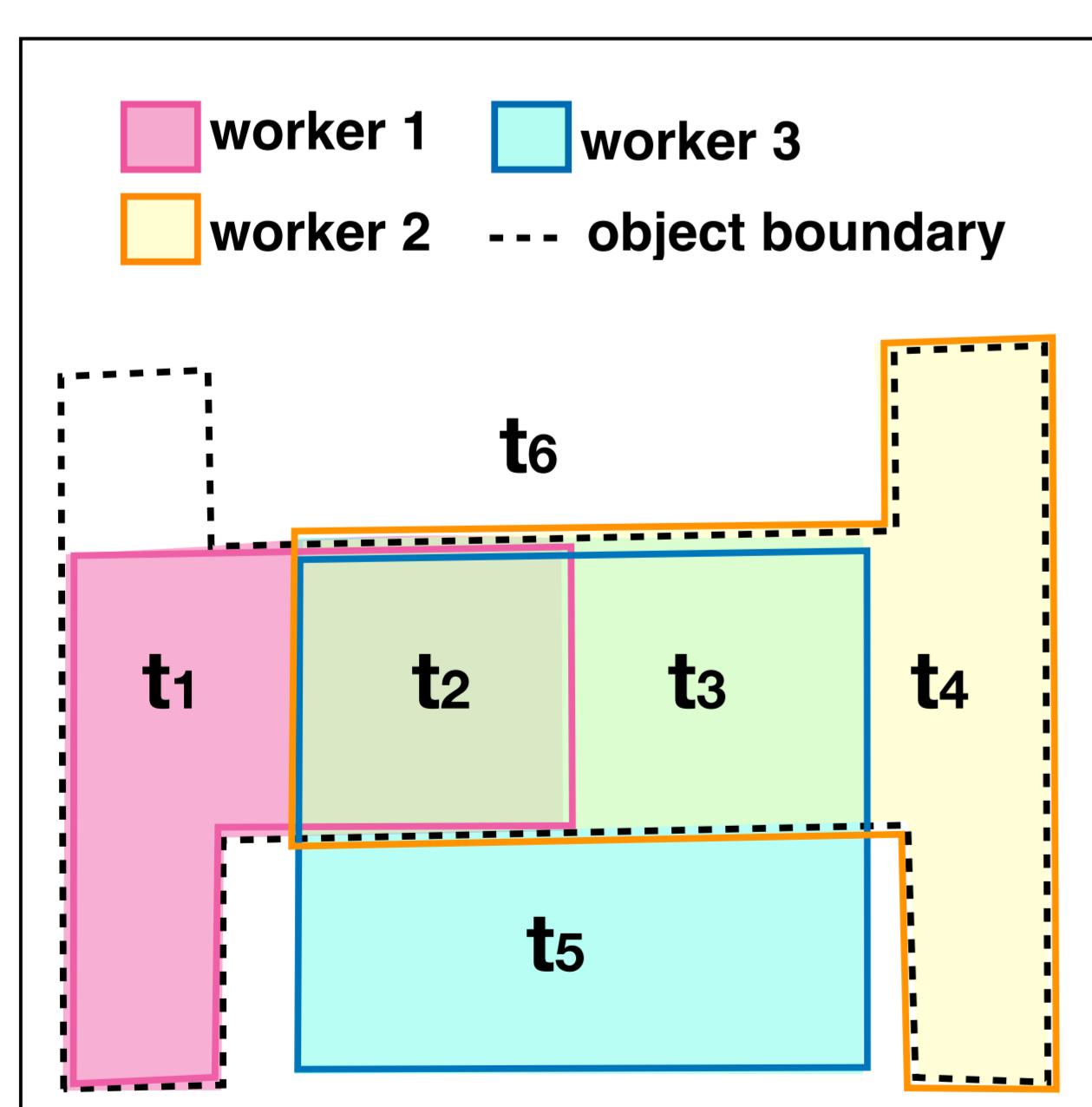
- Spectral clustering on pairwise Jaccard matrix
- Preprocess by retaining largest cluster
- Improves Jaccard score of algorithms up to 5.92%



Fixing Boundary Imperfections

Tile Data Representation

- Tiles: non-overlapping unit of worker segmentations overlaid on each other
- Inference on tiles → aggregation across multiple worker segmentations
- Worker segmentation → “boolean votes” on tiles

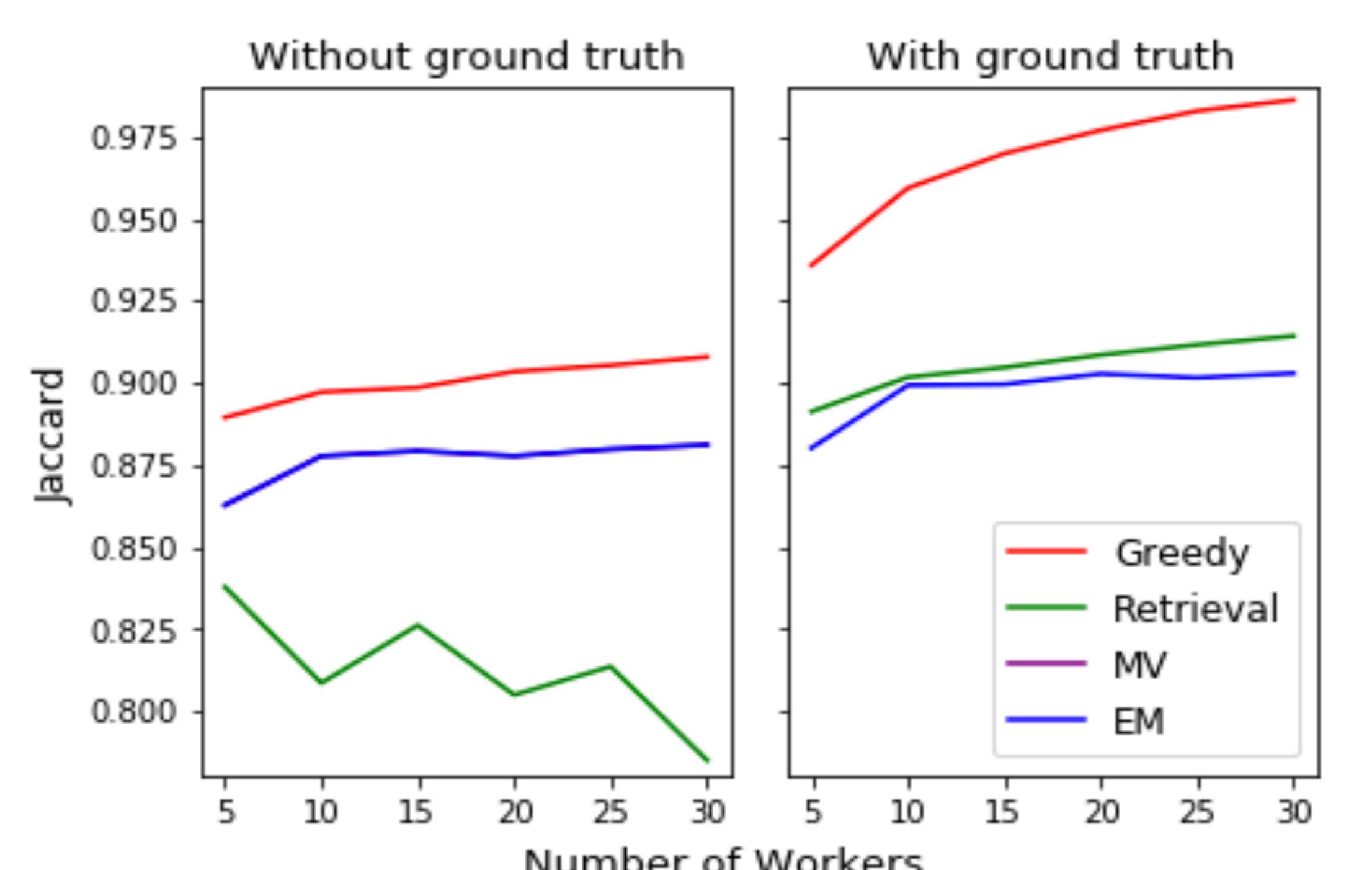


Algorithms

- Majority vote:** include tile if picked by > 50% workers
- Expectation Maximization:** infer likelihood that a tile is part of the ground truth segmentation, while simultaneously estimating hidden worker qualities
- Greedy Tile Picking:** pick tiles in descending order of tiles’ estimated ratios of overlap area with the ground truth
- Baseline:** Retrieving segmentation with highest number of control points

Experimental Evaluation

- Variants which makes use of ground truth information estimates algorithmic upper bound
- Performance scaling as # of workers increases



Conclusion

Contribution:

Semantic perspective resolution through clustering and novel tile-based aggregation algorithms

Future Work:

- Worker qualities good indicators of actual segmentation accuracy, future work on improving inference algorithms
- Incorporating vision signals in tile-based inference