A QA-Gated Ensemble Framework for Robust AI

David Číž[1], João Carias[1], Monika Weissová[1], Petra Bobíková[2]
[1]IT4Innovations, VSB - Technical University of Ostrava, Czech Republic | [2]Asseco Central Europe, a.s, Slovakia

Motivation

- Current “Silver Truth“ (Consensus of Models) generation relies on simple averaging, hiding errors.
- High-stakes biomedical AI requires an auditable score for every prediction.
- EU regulatory alignment demands explainable and robust systems.

Alignment with AI Ethics

- **Explainability:** The QA Referee provides transparent confidence metric, acting as an automated auditor, filtering out low-quality outliers.
- **Bias Mitigation:** With the QA selection, the Ensemble model guides the generation of high quality outputs.
- **Robustness:** Prevents “hallucinations“ from corrupting the final dataset, essential for reliable industrial AI.

Bridging the Trust Gap for Researchers

QA & Ensemble fusion architecture

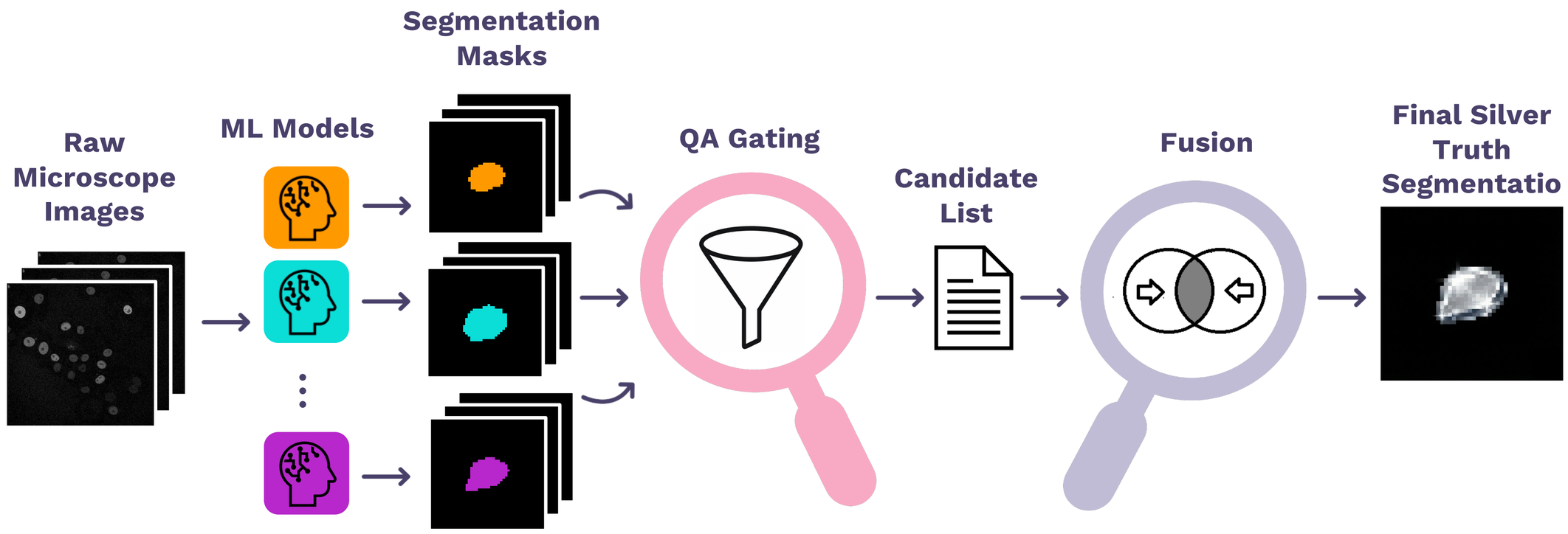


Figure 1: The End-to-End QA-Gated Ensemble Architecture.
Multiple ML base learners generate candidate masks, which are filtered by the QA Referee before the fusion step.

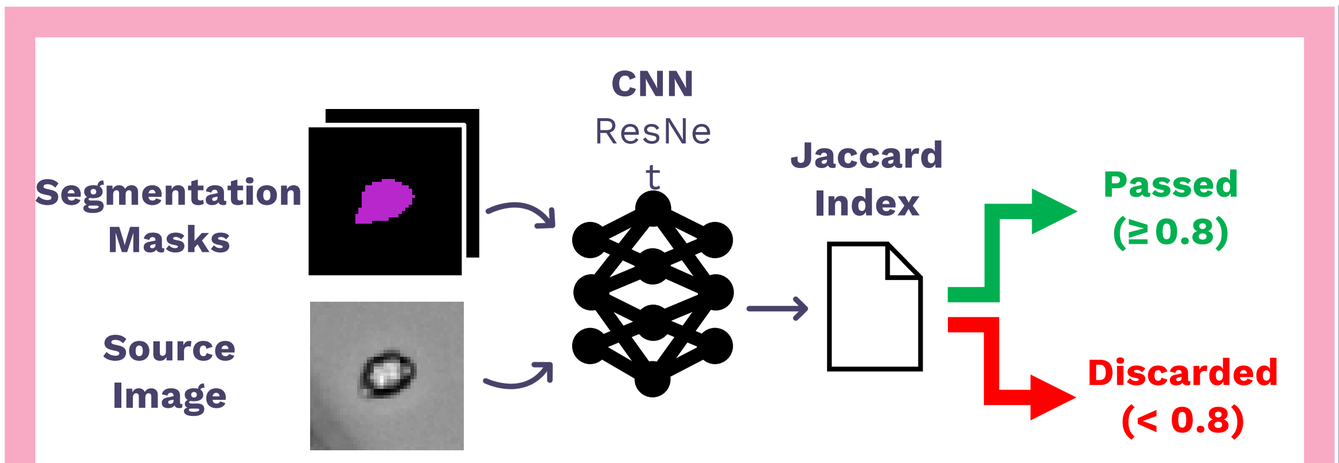


Figure 2: The QA Referee. A CNN classifier estimates the Jaccard Index to filter low-quality segmentations.

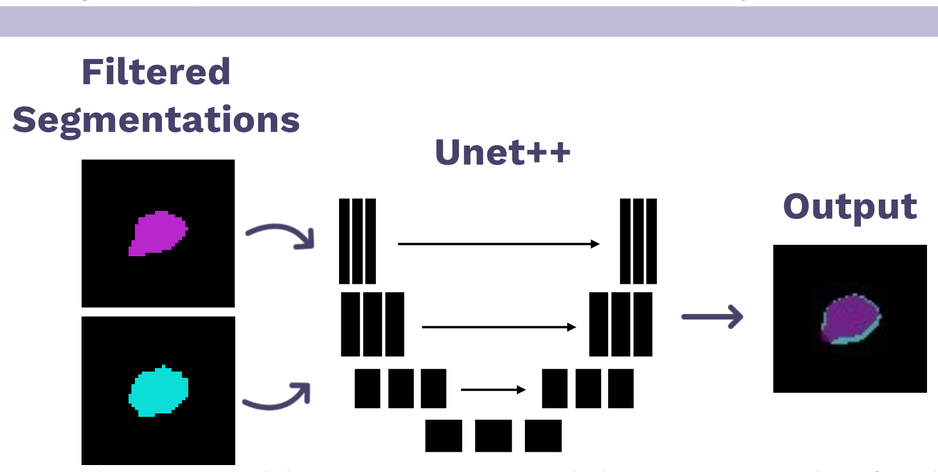
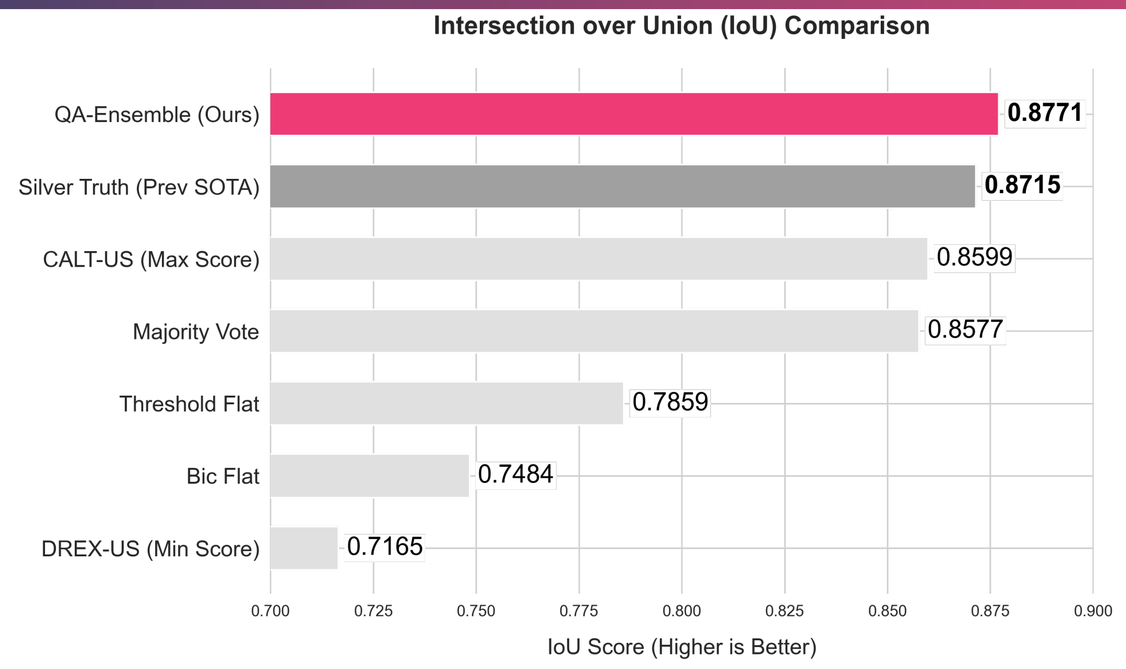


Figure 3: The Ensemble. A Unet++ model generates the final output using the provided segmentations.

Results



Graph 1: Performance benchmarking. Our QA-Ensemble sets a new SOTA IoU of 0.8771, surpassing both consensus averaging (Silver Truth) and majority voting.

Conclusion

We propose an instance-level QA framework that audits every segmentation, achieving new SOTA (+0.6% IoU). This method allows the **generation** of **robust** high-fidelity 'Silver Truth' datasets, **improving the reliability of the trained models** for cell tracking and segmentation challenges.
Limitations: The framework relies on the **diversity of base learners**; if all competitors fail simultaneously (e.g., on Out-of-Distribution data), the Referee, lacking valid options, defaults to the **highest-scoring available mask**.

https://innovaite.sk

Funded by the EU NextGenerationEU through the Recovery and Resilience Plan for Slovakia under the project No. 09I02-03-V01-00029.

InnovAite SLOVAKIA

Funded by the European Union