B39 (Fri-AM)

KING'S College LONDON | University of BRISTOL

MICCAI 2025 — 28th International Conference on Medical Image Computing and Computer Assisted Intervention, 23-27 September 2025, Daejeon Convention Center — Daejeon REPUBLIC OF KOREA

Motion-Boundary-Driven Unsupervised Surgical Instrument Segmentation in Low-Quality Optical Flow

Yang Liu[1], Peiran Wu[2], Jiayu Huo[1], Gongyu Zhang, zhean Yuan, Christos Bergeles[1], Rachel Sparks[1], Prokar Dasgupta[1], Alejandro Granados[1], and Sebastien Ourselin[1]

[1] King's College London
[2] University of Bristol

Introduction

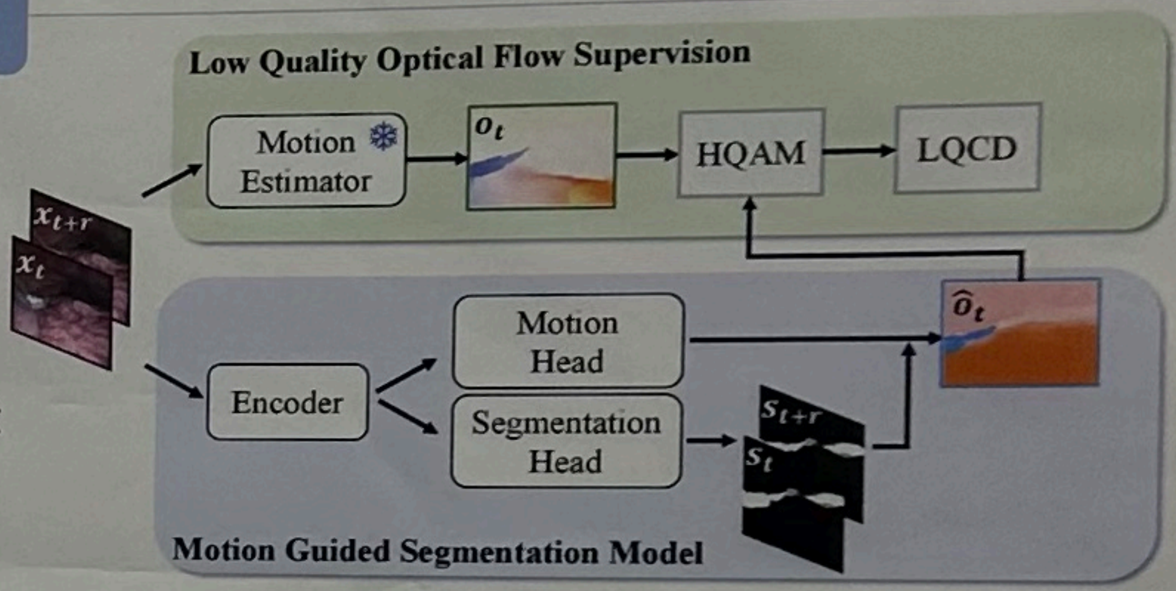
Background:
Instrument segmentation is essential in robotic-assisted surgery, supporting intraoperative guidance and downstream tasks such as workflow recognition and tracking. Existing deep learning methods achieve strong performance but rely heavily on costly manual annotations, limiting scalability in clinical settings. We propose an **unsupervised segmentation framework** that eliminates the need for manual labels.

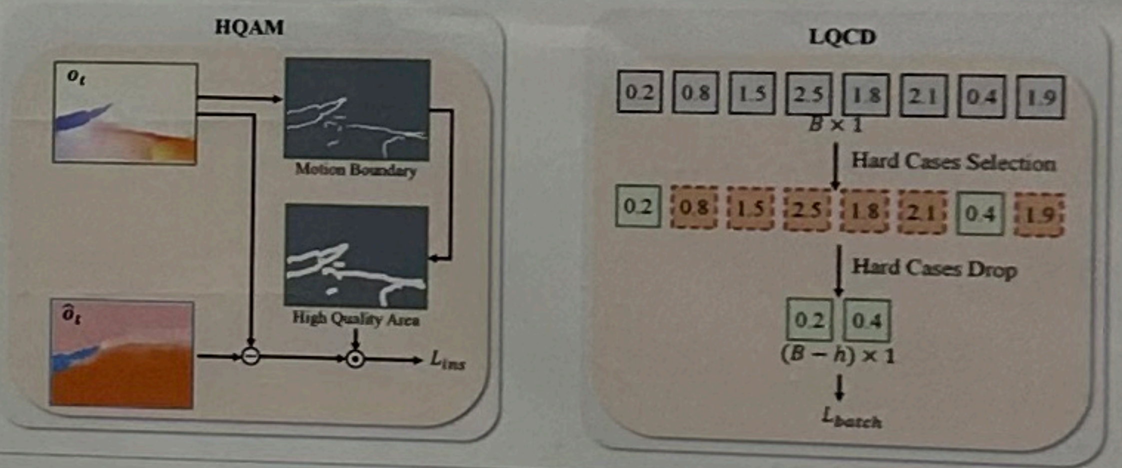
Motivation:
Optical Flow Limitation: Surgical videos often suffer from low-quality flow due to dark regions, abrupt motions, and stationary instruments.
Boundary Challenge: Blurred boundaries in surgical scenes reduce segmentation accuracy.
Subtle Movements: Small or slow instrument motions are easily missed.

Method

We propose a motion-boundary-driven unsupervised surgical instrument segmentation framework. No manual annotations are required.

- **Backbone:** Based on RCF, an end-to-end unsupervised segmentation model.
- **Challenge:** Surgical optical flow is often unreliable due to low light, abrupt motion, or stationary instruments.
- **Our Solutions:**
  1. **HQAM (High-Quality Area Matching):** Focuses supervision on reliable motion boundaries, improving segmentation when instrument interiors have weak signals.
  2. **LQCD (Low-Quality Cases Drop):** Removes globally low-quality frames (hard cases) to prevent error propagation.
  3. **Variable Frame Rate Strategy:** Uses variable frame intervals to better capture subtle or slow instrument movements.





Experiment

Table 2. The effects of different components in method.

Function			Metric
LQCD	HQAM	Variable	mIoU(%)
			46.09
✓			47.15
✓	✓		74.47
✓	✓	✓	75.07

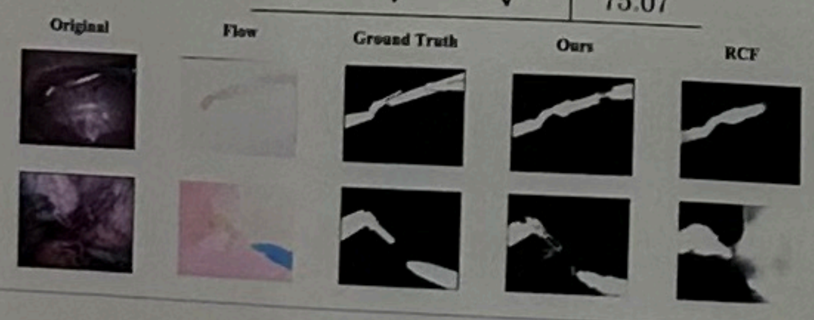


Table 1. Comparison with SOTA supervised methods and unsupervised methods. Mean IoU (%) is reported. And the results of prior works are quoted.

Annotation	*Shape-priors*	Stage	Method	EndoVis 2017VOS	EndoVis 2017Challenge
100%	No	1	TernausNet [18]	89.06	83.60
100%	No	1	MF-TAPNet [6]	89.61	87.56
0%	Yes	1	FUN-SIS(Stage1)	40.08	37.03
0%	Yes	2	FUN-SIS(Stage2)	74.78	68.31
0%	Yes	3	FUN-SIS(Stage3) [16]	83.77	76.25
0%	No	1	AGSD [9]	71.47	67.85
0%	No	1	RCF(Stage1)	46.09	46.17
0%	No	2	RCF(Stage2) [7]	49.18	49.19
0%	No	1	RCF(Stage1) + **Ours**	**75.07**	**72.07**

Parameters		Metric
α	d	mIoU(%)
	7	68.34
$\frac{\pi}{12}$	3	66.04
	1	67.81
	7	68.74
$\frac{\pi}{6}$	3	66.03
	1	66.21
	7	**70.66**
$\frac{\pi}{3}$	3	67.03
	1	64.74

Conclusion

We propose an unsupervised surgical instrument segmentation framework that integrates HQAM, LQCD, and variable frame rates to overcome low-quality optical flow. Our method achieves strong performance, reducing the reliance on manual annotations. With its **plug-and-play** design, the approach can be extended to other motion-driven tasks (e.g., unsupervised depth estimation), offering a promising step toward scalable and robust clinical applications.

We are recruiting PhD students! *Interested candidates are welcome to get in touch.*