JOHNS HOPKINS BIOMEDICAL ENGINEERING

DURR LAB

NEURAL INFORMATION PROCESSING SYSTEMS

SAM2Flow: Interactive Optical Flow Estimation with Dual Memory for *in vivo* Microcirculation Analysis

Luojie Huang, Ryan Zhang, Marisa M. Morakis, Michaela Taylor-Williams,
Gregory N. McKay, Nicholas J. Durr ✉
Department of Biomedical Engineering, Johns Hopkins University, Baltimore, USA

lhuang48@jhu.edu

Poster Session: Wed 3Dec 4:30-7:30 pm at Exhibit Hall C, D, E

Microvascular Flow Estimation

JOHNS HOPKINS
BIOMEDICAL ENGINEERING
durr.jhu.edu #2

Clinical Impact

In vivo quantitative measurement of blood flow within capillaries has broad clinical applications:

- **Cardiovascular and Critical Care Monitoring**: shock and hypoxia
- **Hematologic and Vascular Disorders**: sickle cell disease, microthrombosis
- **Inflammatory and Infectious Diseases**: neuroinflammation and sepsis
- **Oncology**: tumor angiogenesis

Current Methods:

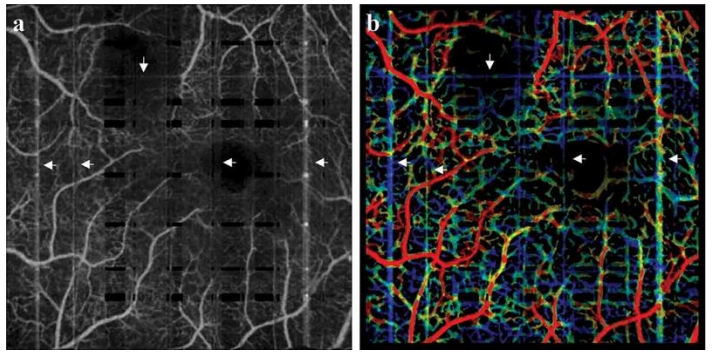
- Laser Doppler Flowmetry (LDF)
- Optical Coherence Tomography Angiography (OCTA)
- Sidestream Dark Field Microscopy (SDF)
- Laser Speckle Contrast Imaging (LSCI)

Limitations:

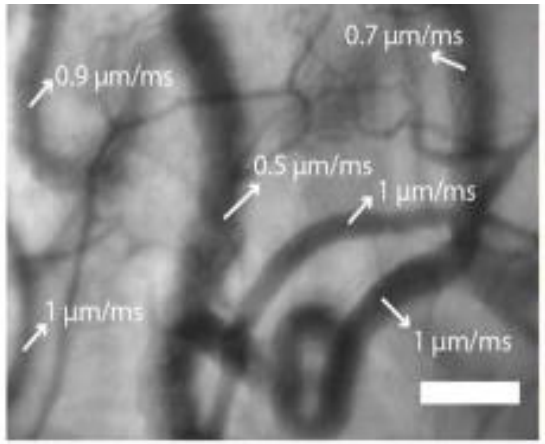
- Imaging modalities: *limited spatial (cellular) or temporal resolution;*
- Susceptible to *motion artifacts*;
- Output measurements: *manual estimations* (SDF) or *relative flow index* (LDF, LSCI)

LDF *[1]*

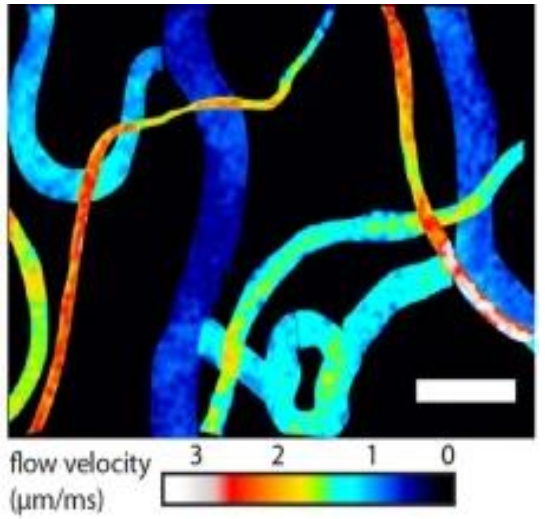
OCTA *[2]*



SDF *[3]*



LSCI *[3]*



[1] OxyFlo™ Laser Doppler Flowmetry
[2] Arya, M., et al. *Expert review of medical devices, 15.23 (2018)*
[3] Nadort, A., et al. *Scientific reports*, *6*(1), 25258 (2016).

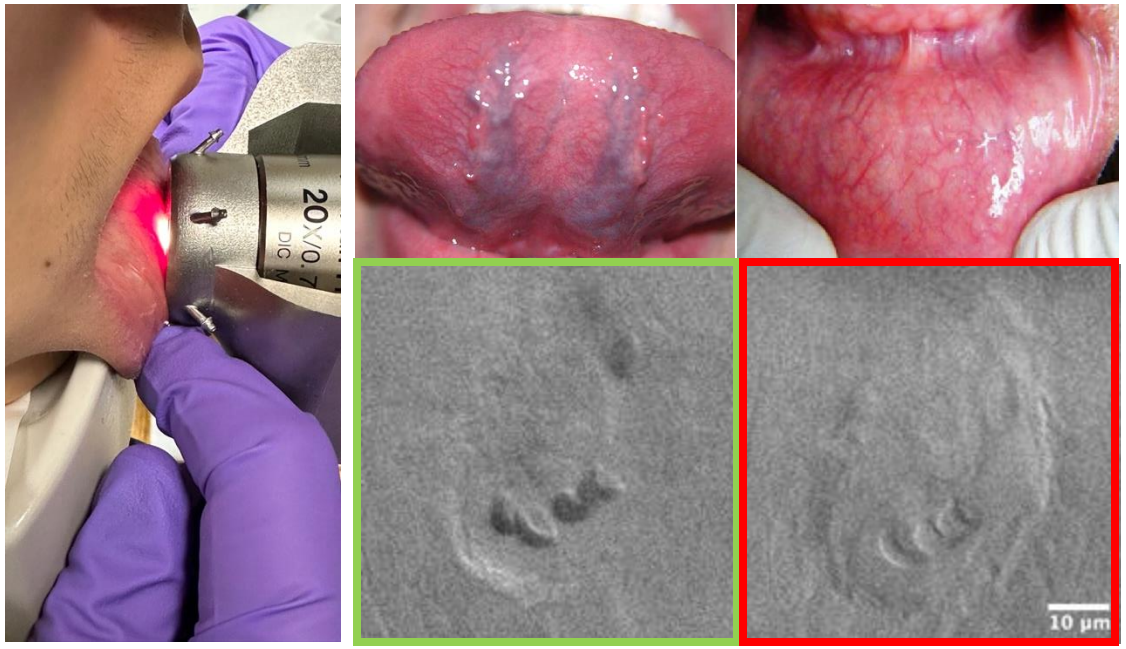
Microvascular Flow Estimation

JOHNS HOPKINS
BIOMEDICAL ENGINEERING
durr.jhu.edu #3

Proposed Method:

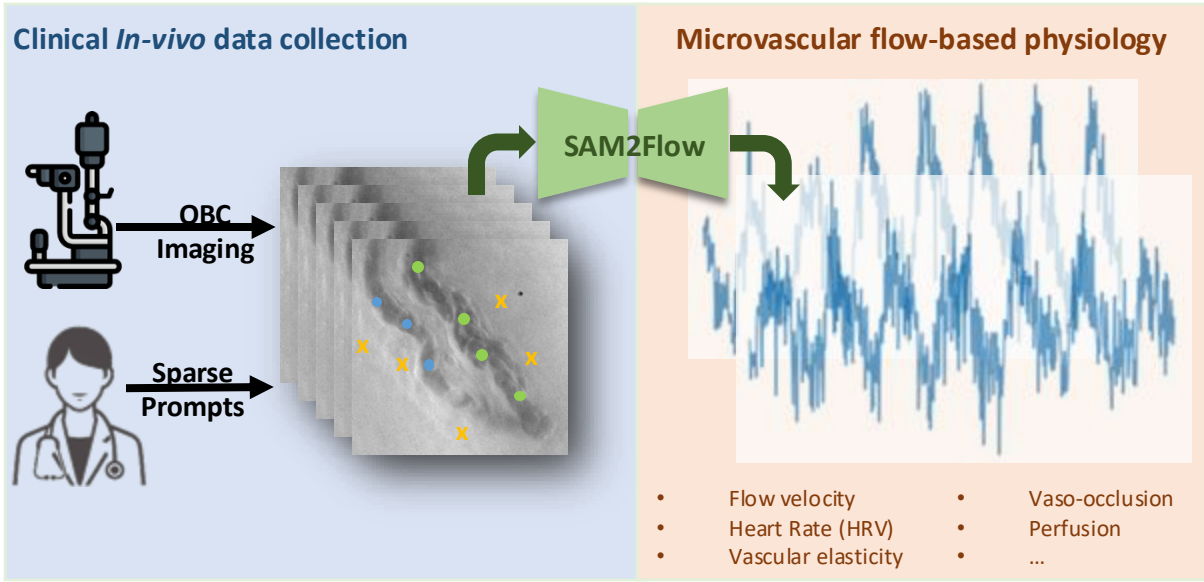
Imaging modality: Oblique back-illumination microscopy (OBM)

- *Enhanced Contrast*
  - Illumination: Green & Red
  - Contrast: absorption(G) & phase(G, R)
- *Non-invasive*
  - Imaging site: Oral mucosa (inner lip)
  - Imaging depth: ~100 µm
- *Fast*
  - Imaging speed: 200 fps
  - FoV: 1024x1536 (176 µm x280 µm, 20x)



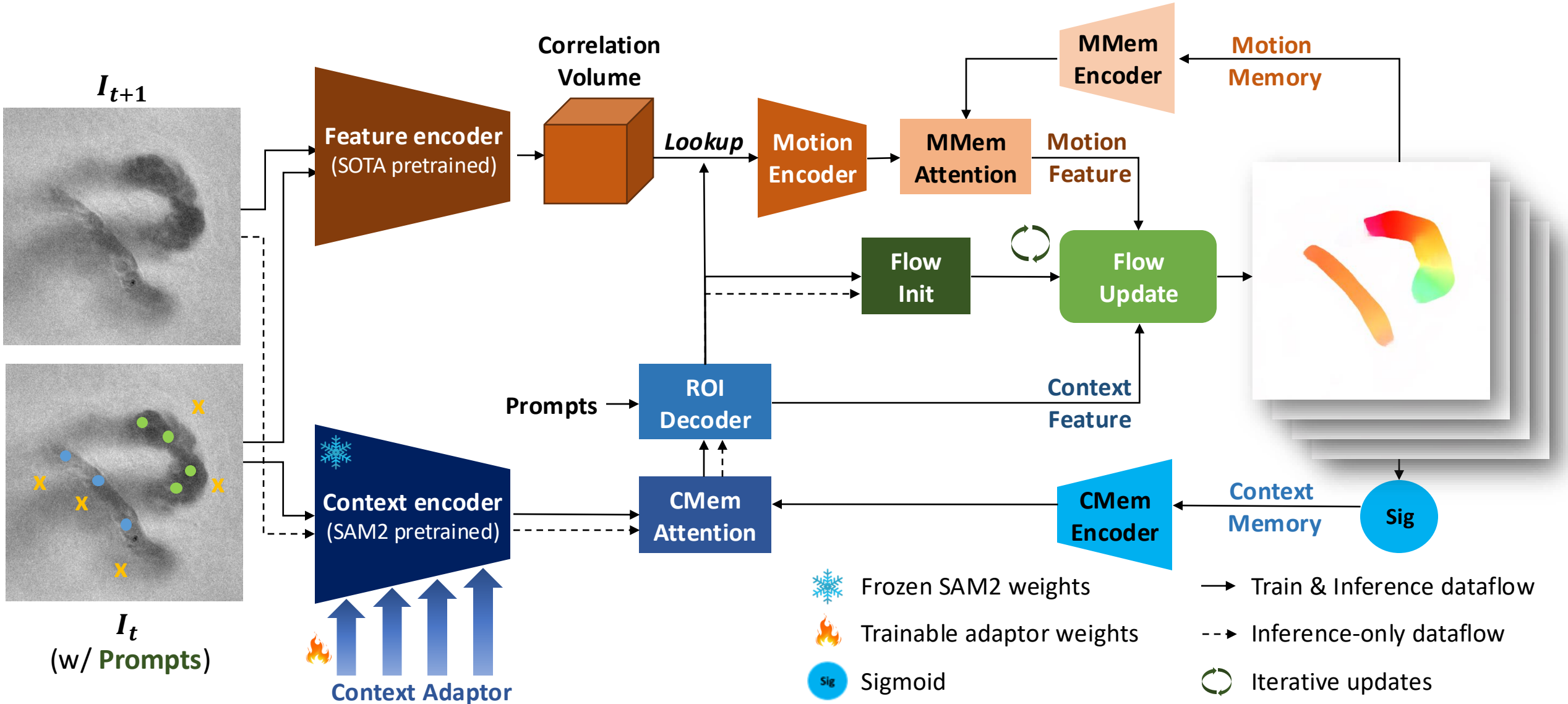
Flow Estimation Algorithm: SAM2Flow

- Interactive ROI Specification: sparse point prompts to SAM-based context branch (**SAM2-L**);
- Accurate Foreground Estimation: focused optical flow refinement within ROI;
- Long-term Stability: optical flow estimation over Long OBM sequences with a dual(context and motion) memory mechanism;
- Fast Inference: efficient estimation matching high-speed video recording.



SAM2Flow

- **Motion memory**: ensures temporally stable flow outputs over a long range.
- **Context memory**: propagates user-defined ROIs across frames, increasing efficiency and suppressing background noise;

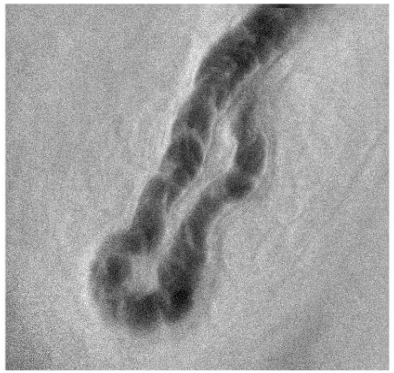


Microvascular Flow Dataset

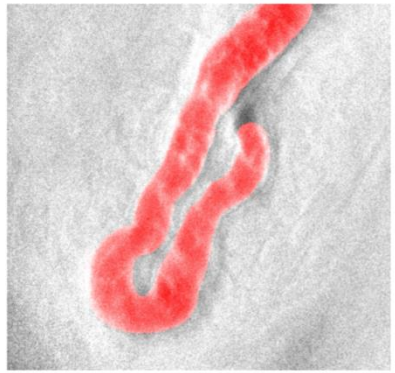
Data Collection:
- 15 healthy volunteers
- Inner lower lip surface
- 200Hz, 512x512x1 (Green Channel)

Flow Annotation: semi-automatic
- Flow segmentation
- Centerline extraction
- Spatio-temporal diagram Estimation
- Optical Flow mapping

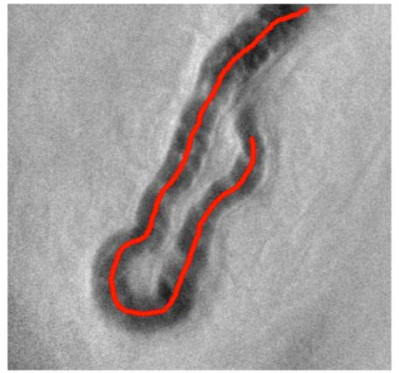
Microvascular Flow Dataset:
- Vessel diversity: singular vessels, crossing vessels, overlapping vessels
- Flow diversity: straight single-profile flow, laminar flow, turbulent flow
- Long videos: 4091 frames/video



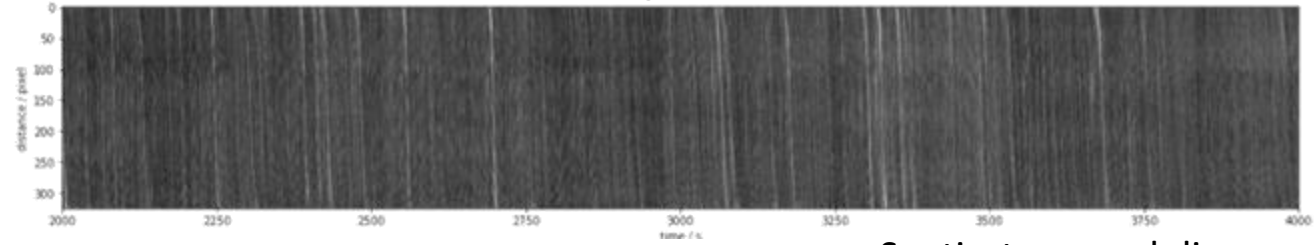
Stabilized video



Flow segmentation



Centerline extraction



Spatio-temporal diagram

Dataset Overview

Dataset	Videos #	Flowmaps #	Avg video length (flowmaps/video)
Sintel[51]	23	1,041	45
Spring[52]	37	10,000	270
KITTI 2015[54]	400	1,600	4
FlyingChairs[55]	N/A	22,872	N/A
Microvascular	75	306,800	4,091

Experiment: Microvascular Flow Estimation

SAM2Flow achieves **SOTA foreground optical flow estimation** performance on the microvascular flow dataset.

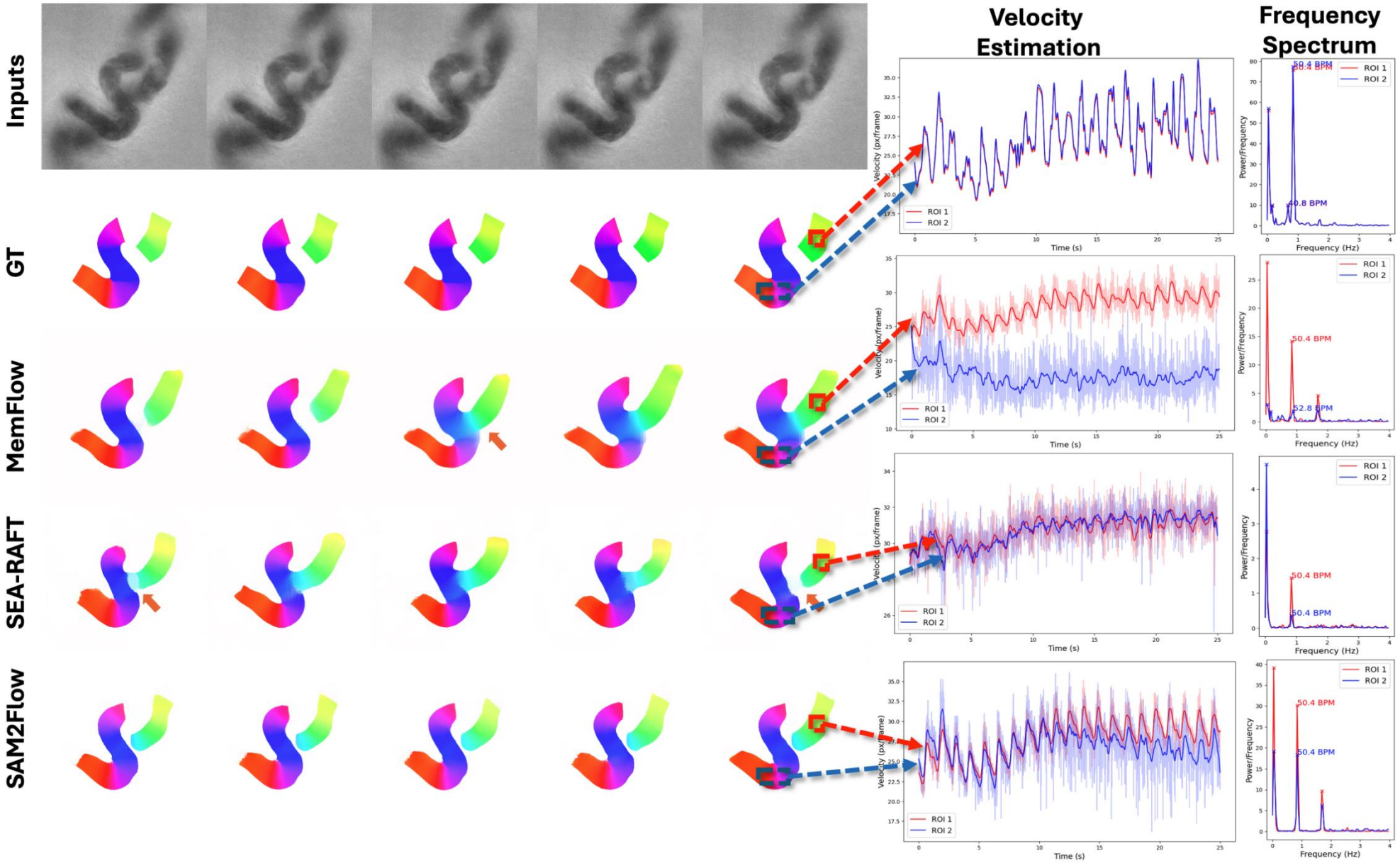
- **Accurate foreground flow estimation**: 12% improvement in Foreground EPE (FEPE);
- **Fast inference**: over 20 FPS as a multi-frame OF model;

Model	Whole Image				Foreground				Speed
	EPE↓	1px↑	3px↑	5px↑	FEPE↓	5px↑	10px↑	15px↑	mspf↓
RAFT [16]	3.18 (2.61)	0.86	0.89	0.91	27.73 (24.79)	0.39	0.52	0.56	51.48
GMA [17]	3.22 (3.66)	0.87	0.89	0.91	28.34 (26.82)	0.38	0.54	0.58	<u>43.66</u>
SEA-RAFT[46]	<u>1.28</u> (1.03)	**0.88**	<u>0.92</u>	<u>0.94</u>	<u>6.60</u> (<u>5.47</u>)	**0.69**	**0.86**	<u>0.91</u>	**21.14**
FlowFormer++ [44]	1.72 (1.38)	**0.88**	0.91	0.93	10.89 (9.28)	0.60	0.78	0.84	133.95
VideoFlow_BOF$^{(MF)}$ [18]	3.28 (2.51)	0.86	0.87	0.88	28.16 (26.64)	0.15	0.32	0.41	112.67
MemFlow$^{(MF)}$ [20]	1.79 (1.40)	**0.88**	0.91	0.93	12.47 (10.23)	0.58	0.74	0.80	43.98
StreamFlow$^{(MF)}$ [19]	1.43 (<u>1.02</u>)	**0.88**	0.90	0.93	10.13 (8.36)	0.49	0.74	0.84	60.07
SAM2Flow$^{(MF)}$	**1.14** (**0.92**)	**0.88**	**0.93**	**0.96**	**5.84** (**4.86**)	<u>0.66</u>	**0.86**	**0.93**	48.78

* $^{(MF)}$ indicates multi-frames optical flow models; best performance is **highlighted**, while second best performance is <u>underlined</u>; **EPE** & **FEPE**: Mean(Standard Deviation, SD); **mspf**: milliseconds per frame.

Physiology Analysis

- **ROI-centric prediction**: separate flow estimation for specified vessels;
- **Long-range stability**: stable velocity plots extracted from optical flow estimations over 5000 frames;
- **Clinically relevant estimation:** accurate heart rate estimation.



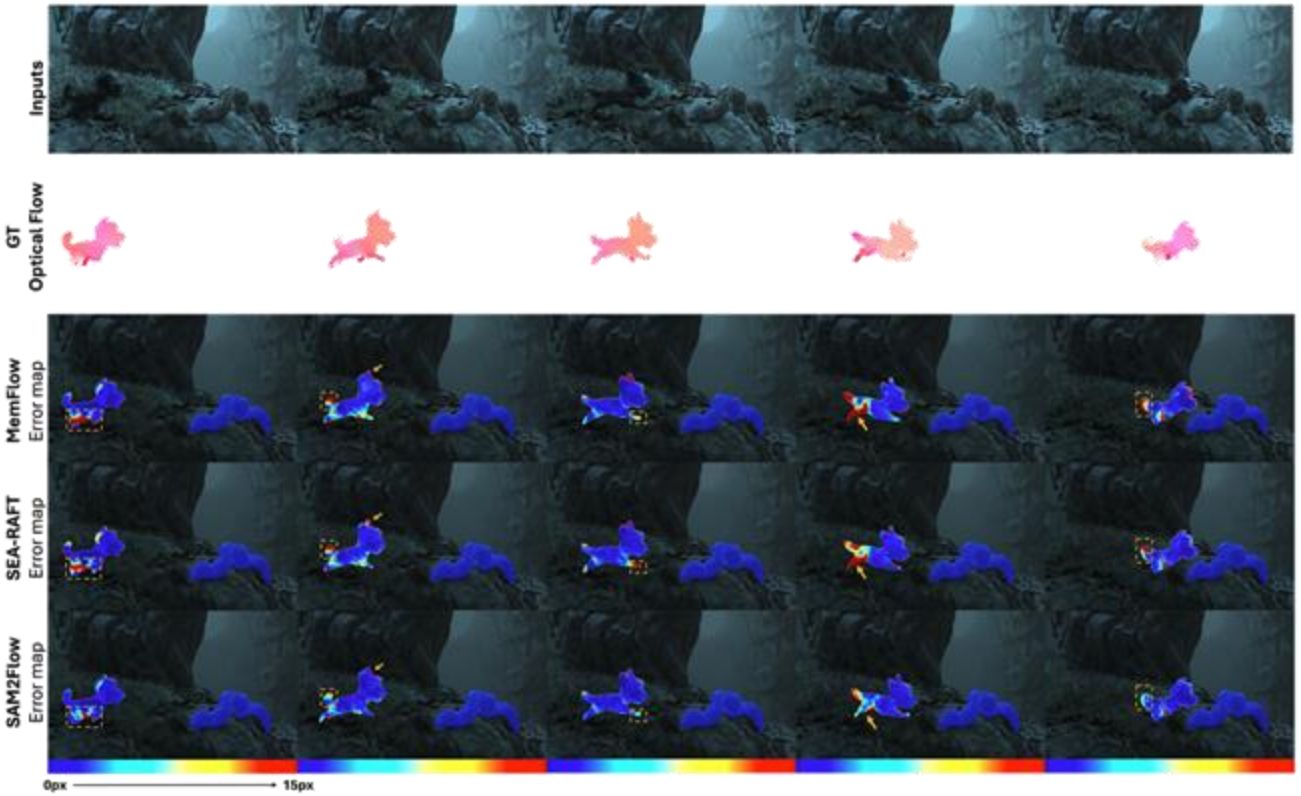
Experiment: Foreground OF Estimation

SAM2Flow demonstrates **impressive** foreground optical flow estimation on public benchmarks of synthetic videos, including **Sintel** and **Spring**.

- SOTA optical flow estimation within the ROI with the **most prominent motions**;
- **Higher accuracy at object contours** due to the awareness of ROI segmentation.

Model	Sintel-FEPE	
	Clean↓	Final↓
RAFT [16]	5.21 (9.30)	5.47 (10.12)
GMA [17]	4.65 (7.45)	5.14 (8.17)
SEA-RAFT[46]	3.26 (7.54)	4.08 (8.94)
MemFlow$^{(MF)}$ [20]	3.77 (5.82)	4.27 (7.09)
StreamFlow$^{(MF)}$ [19]	4.06 (**5.37**)	4.43 (**5.66**)
SAM2Flow$^{(MF)}$	**3.17** (6.89)	**3.39** (5.97)

Model	Spring-Foreground				
	FEPE↓	1px (%)↑	Fl (%)↓	Fl-epe↓	WAUC↑
RAFT [16]	2.25 (6.97)	74.25	9.87	7.98 (8.98)	74.76
GMA [17]	2.17 (6.30)	76.21	9.08	7.67 (7.72)	79.67
SEA-RAFT[46]	1.45 (5.61)	86.32	**5.18**	8.18 (10.85)	83.85
MemFlow$^{(MF)}$ [20]	1.56 (7.27)	86.49	7.24	8.82 (12.72)	83.54
StreamFlow$^{(MF)}$ [19]	1.54 (7.23)	85.16	5.82	7.74 (10.16)	82.51
SAM2Flow$^{(MF)}$	**1.23 (4.21)**	**87.13**	5.29	**7.30 (5.29)**	**84.57**



Conclusion

We introduce **SAM2Flow**, an interactive ROI-centric optical flow estimation model that enables:

- **User-specified ROIs** through sparse point prompts;
- **Long-range temporal consistency** with motion & context memory attention mechanism;
- **Meaningful *in vivo* microvascular flow estimation** from long OBM videos;
- **Superior generalizability** proven on public datasets.

Limitations:

Flow estimation could be affected by **imperfect ROI detection** from the context branch:

- **Over segmentation**. Motion encoder and iterative flow updates further refine and delineate the motion boundary.
- **Transient frames drop or incomplete ROI**. The motion memory module provides redundancy against occasional failures by maintaining consistency with prior memories.
- **Missed/incomplete ROI over an extended period**. Failed ROI detection over consecutive frames might cause flow estimation failure. (**Solution**: additional or correctional user prompts.)

GitHub Project Page:

Future Works:

- **Efficiency.** More compact backbones, such as SAM2-B+ and SAM2-S.
- **Flow diversity**. SAM2Flow on microvascular flow with conditions, such as sickle cell and sepsis.

Acknowledgement: This work was supported in part by research funding from the Gates Foundation (INV-006005).