Memory-Augmented SAM2 for Training-Free Surgical Video Segmentation

Ming Yin, Fu Wang, Xujiong Ye, Yanda Meng, Zeyu Fu*

{my417, Z.Fu}@exeter.ac.uk

Department of Computer Science, University of Exeter, UK

University of Exeter

MICCAI2025

28th International Conference on Medical Image Computing and Computer Assisted Intervention

23-27 September 2025

Daejeon Convention Center

Daejeon REPUBLIC OF KOREA

Introduction

➢ Significance

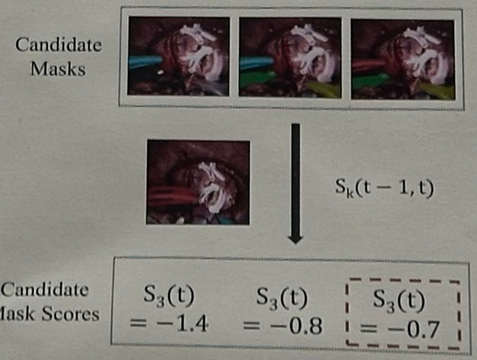
- A critical component of medical imaging and surgical scene analysis;
- Achieve pixel-level localization and contour extraction of instruments.

➢ Key Challenges

Time Flow →

Object Displacement

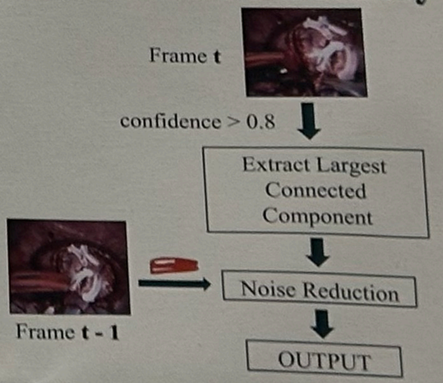
➢ Context-Aware Memory



Training-Free Tree Memory

- Tree-structured memory
- No additional training is required

➢ Occlusion-Resilient Memory



Candidate Generation

- Generate multiple candidate masks per frame