IMT Atlantique
Bretagne-Pays de la Loire
École Mines-Télécom

From SIAMOT to XMEM

Group 10:

Carlos ARGUILAR
Franco Martin DI MARIA
Renzo MORALES

SOMMAIRE

IMT Atlantique
Bretagne-Pays de la Loire
École Mines-Télécom

CHAPTER 1

SiamMOT: Siamese Multi-Object Tracking

Bing Shuai, Andrew Berneshawi, Xinyu Li, Davide Modolo, Joseph Tighe

IMT Atlantique
Bretagne-Pays de la Loire
École Mines-Télécom

~~SiamMOT: Sia racking~~

Problems with implementations requirements:

- Compatibility of the current cuda version with the machine
- Compatibility issues with the dependencies
- Downgrade Cuda's version in school computers

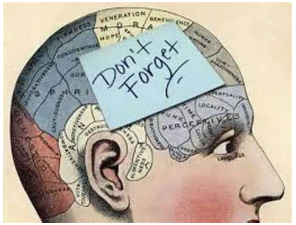
CHAPTER 2

XMem: Long-Term Video Object Segmentation with an Atkinson-Shiffrin Memory Model Ho

Kei Cheng and Alexander G. Schwing University of Illinois Urbana-Champaign {hokeikc2, aschwing}@illinois.edu

IMT Atlantique
Bretagne-Pays de la Loire
École Mines-Télécom

XMEM is a video object segmentation architecture inspired by Atkinson-Shiffrin memory model

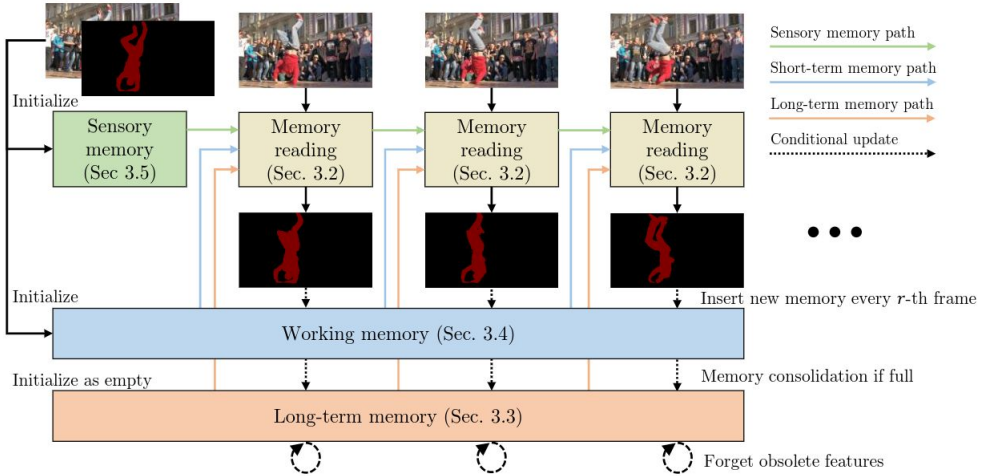


Wait a minute… What is Atkinson-Shiffrin memory model?

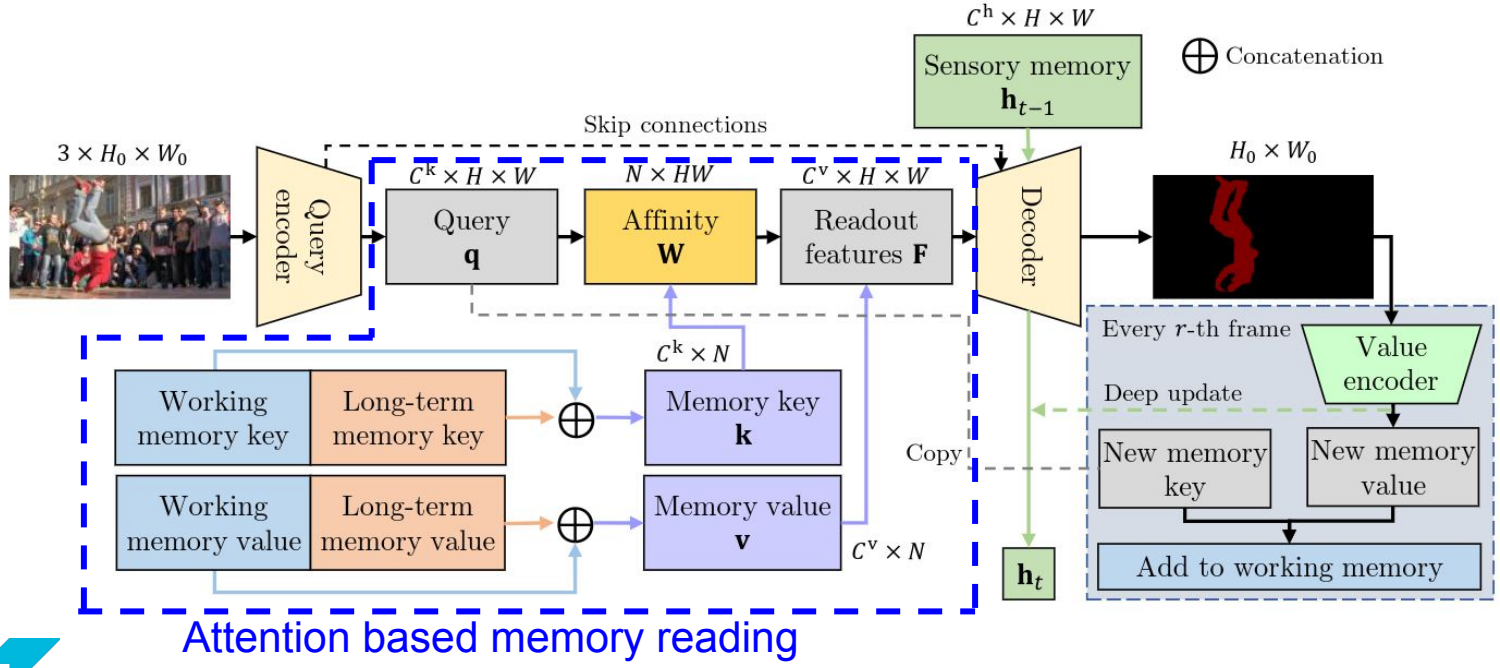
Sensory Memory	Short-Term memory	Long-Term Memory
⬇	⬇	⬇
Sensory registers	Temporary storage Working memory	Permanent storage

Atkinson-Shiffrin Memory Model:

- Sensory Memory: Updated every frame (short-term low level information)
- Working Memory: High resolution features
- Long-term Memory: Compact and sustained



Memory Reading:



Application in the challenge:

Sequence of images

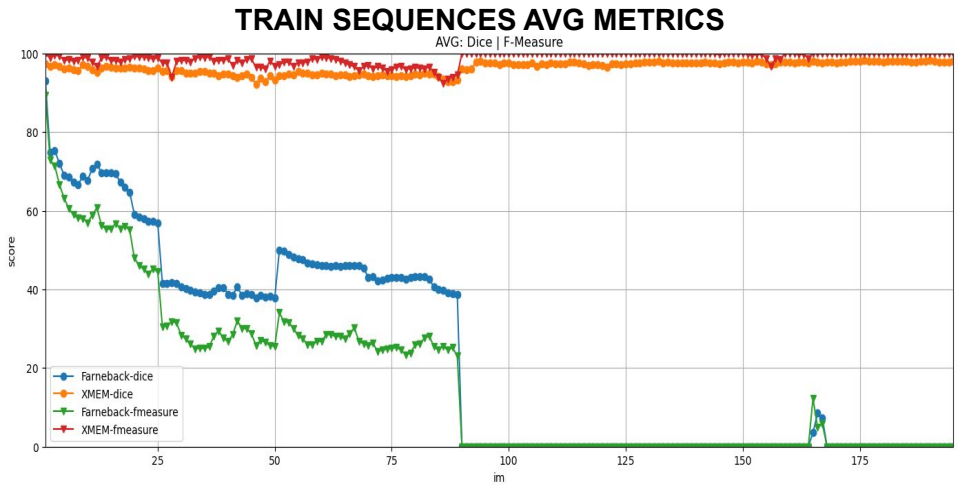
Ground truth segmentation mask for 1st frame

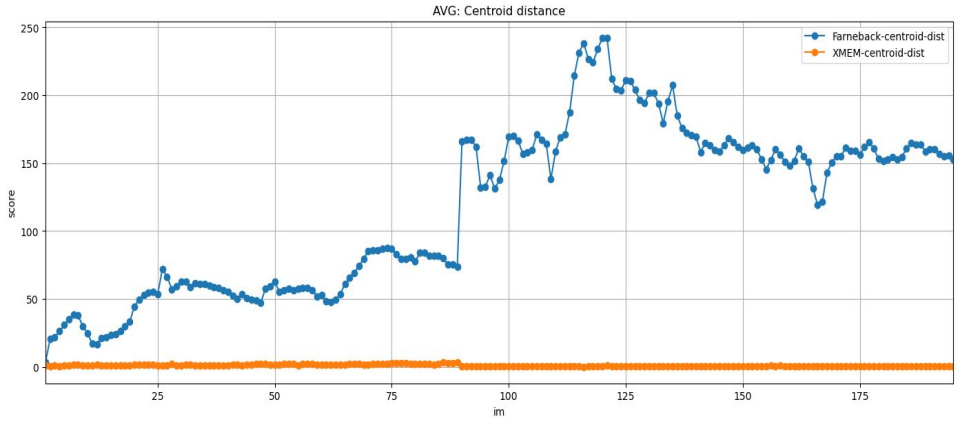
Sequence of images with object segmentation

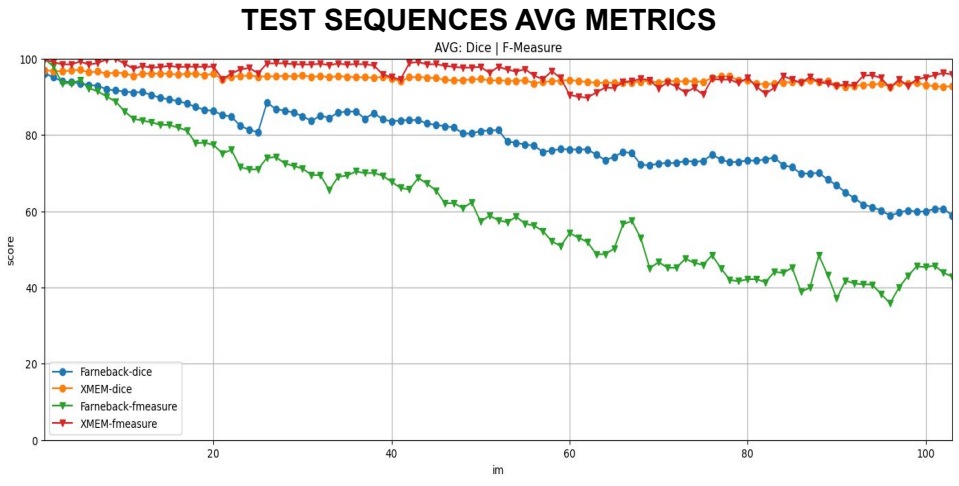
CHAPTER 3
RESULTS

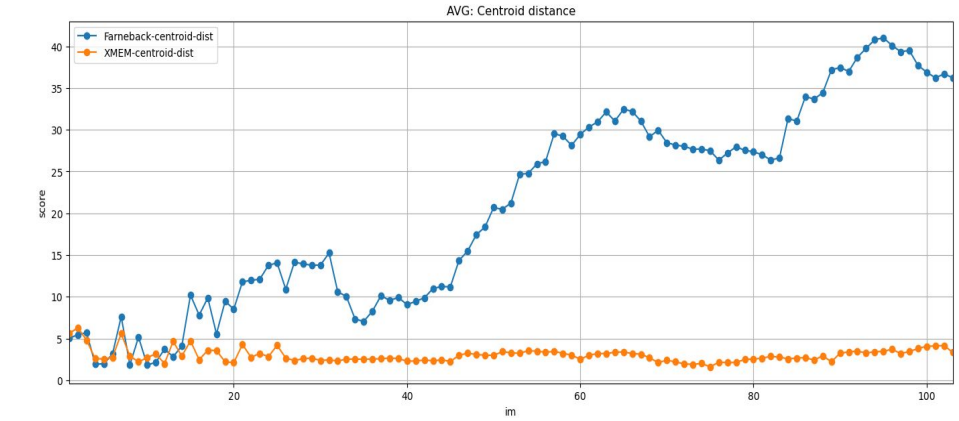
IMT Atlantique
Bretagne-Pays de la Loire
École Mines-Télécom

3.1 Metrics

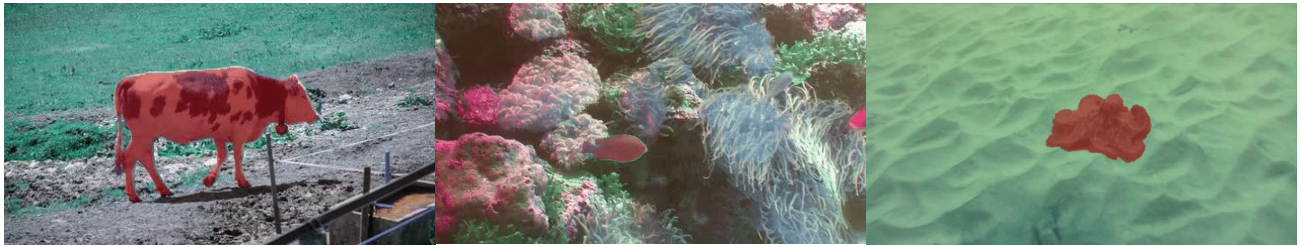

TRAIN SEQUENCES AVG METRICS
AVG: Dice | F-Measure
score
im
Farneback-dice
XMEM-dice
Farneback-fmeasure
XMEM-fmeasure


AVG: Centroid distance
score
im
Farneback-centroid-dist
XMEM-centroid-dist


TEST SEQUENCES AVG METRICS
AVG: Dice | F-Measure
score
im
Farneback-dice
XMEM-dice
Farneback-fmeasure
XMEM-fmeasure


AVG: Centroid distance
score
im
Farneback-centroid-dist
XMEM-centroid-dist

3.2 Video Trackings



Execution time: 1.99s

Execution time: 0.54s

Execution time: 0.71s

CHAPTER 4
Conclusion

IMT Atlantique
Bretagne-Pays de la Loire
École Mines-Télécom

Important features :

- Multi-store feature memory model : Sensory, Working and Long-Term memory
- Video multiple object segmentation
- Excellent performance with minimal GPU memory usage (< 6 GB) for both long and short videos.
- Good step toward accessible VOS on mobile devices

Thanks for your attention.

IMT Atlantique
Bretagne-Pays de la Loire
École Mines-Télécom