Interactive Video Object Segmentation via Shift Modules

Yuk Heo, Yeong Jun Koh, and Chang-Su Kim

Introduction

Introduction

- Interactive Video Object Segmentation
  - Assumes that an user gives iterative refinement inputs
  - Segment the object in the video iteratively

Introduction

- Scenario of DAVIS challenge 2019
  I. Scribbles are given for each target objects in a single frame
  II. Segment the video objects

Input scribbles

Introduction

- Scenario of DAVIS challenge 2019
  I. Scribbles are given for each target objects in a single frame
  II. Segment the video objects

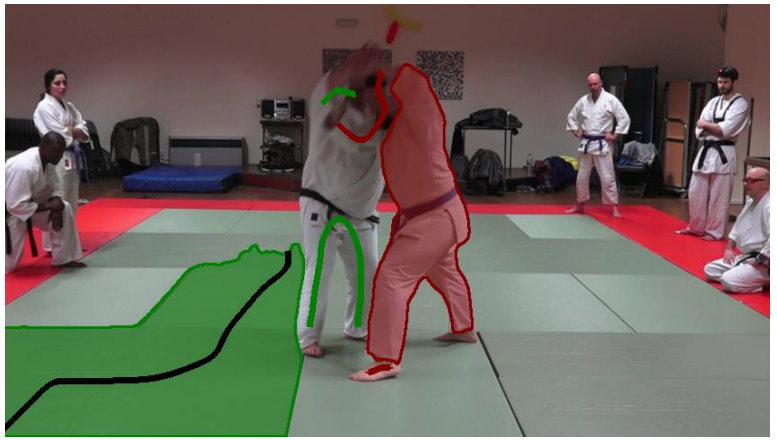
First round

Input scribbles

Introduction

- Scenario of DAVIS challenge 2019
  I. Scribbles are given for each target objects in a single frame
  II. Segment the video objects
  III. From the worst segmented frame, scribbles for refinement are given
  IV. Segment the video objects

Input of the first round



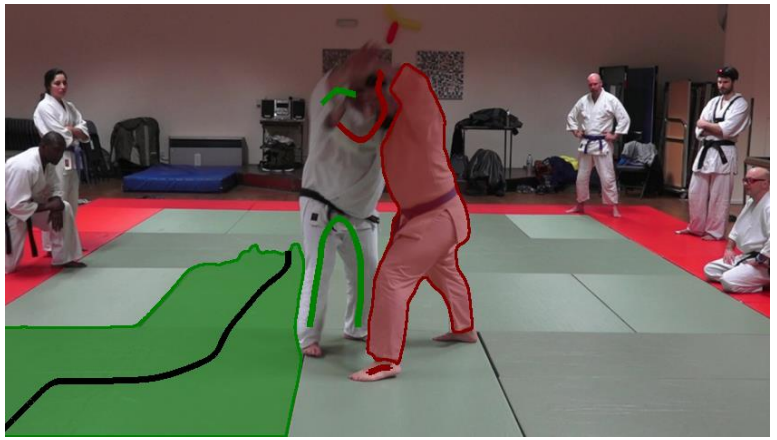
Input of the rest rounds

Introduction

- Scenario of DAVIS challenge 2019
  I. Scribbles are given for each target objects in a single frame
  II. Segment the video objects
  III. From the worst segmented frame, scribbles for refinement are given
  IV. Segment the video objects

Second to eighth round

Input of the first round

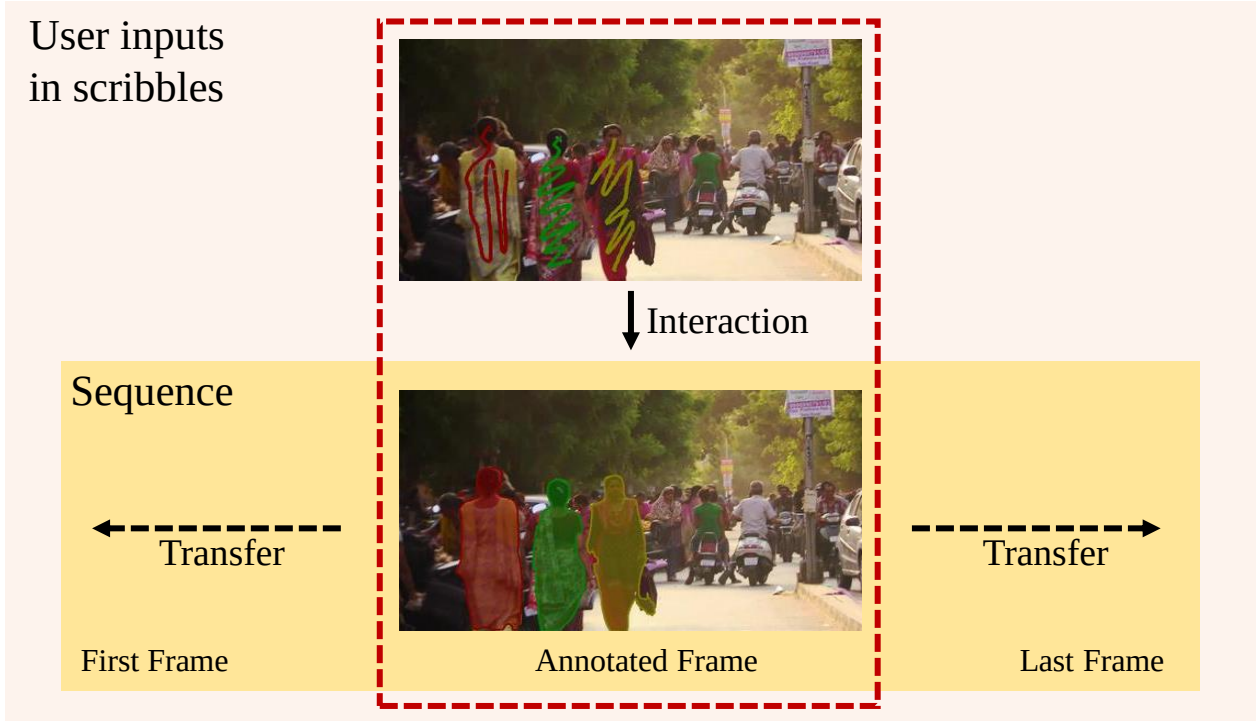


Input of the rest rounds

Proposed Algorithm

Dividing networks

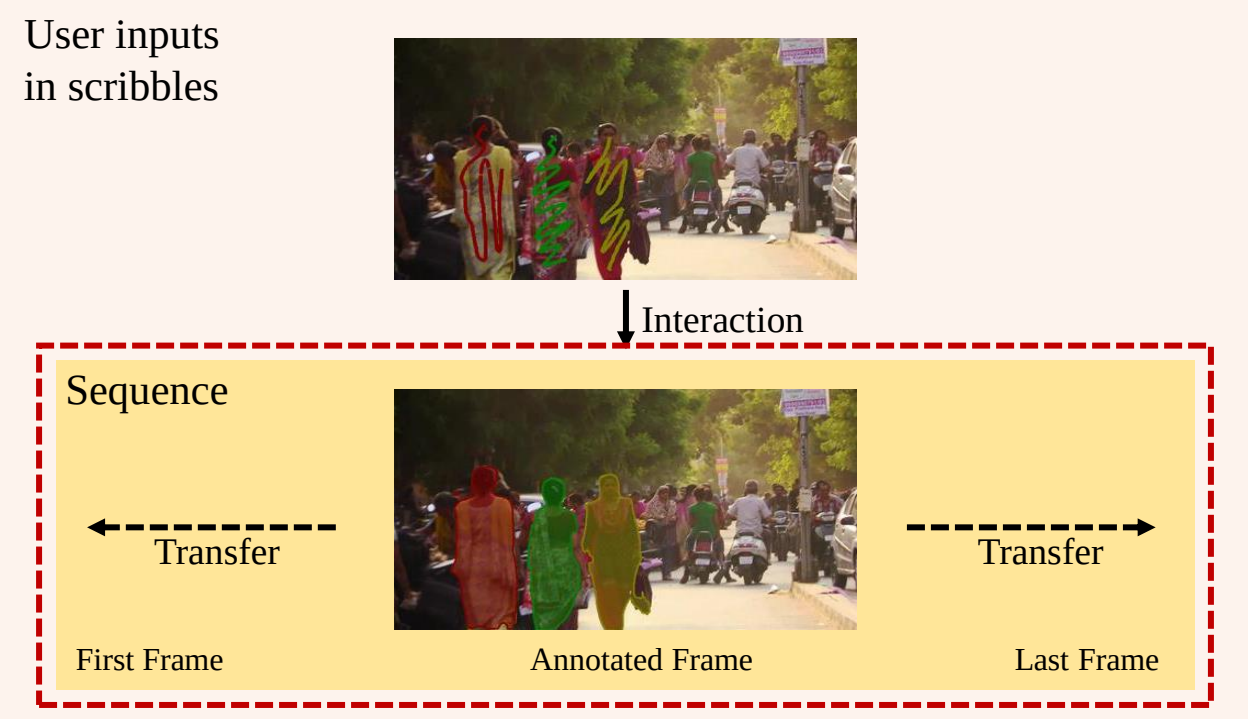
- Adopted the framework from [1]
  - Network for interaction
  - Network for transfer



[1] S. W. Oh, et al. Fast user-guided video object segmentation by interaction-and-propagation networks. In CVPR, 2019.

Dividing networks

- Adopted the framework from [1]
  - Network for interaction
  - Network for transfer

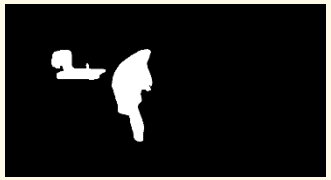


[1] S. W. Oh, et al. Fast user-guided video object segmentation by interaction-and-propagation networks. In CVPR, 2019.

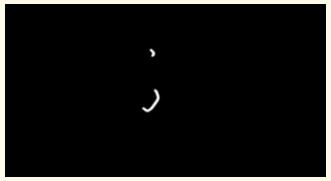
Network for interaction

- Network for interaction in the annotated frame

Annotated frame



Previous round mask



Positive scribbles

Negative scribbles

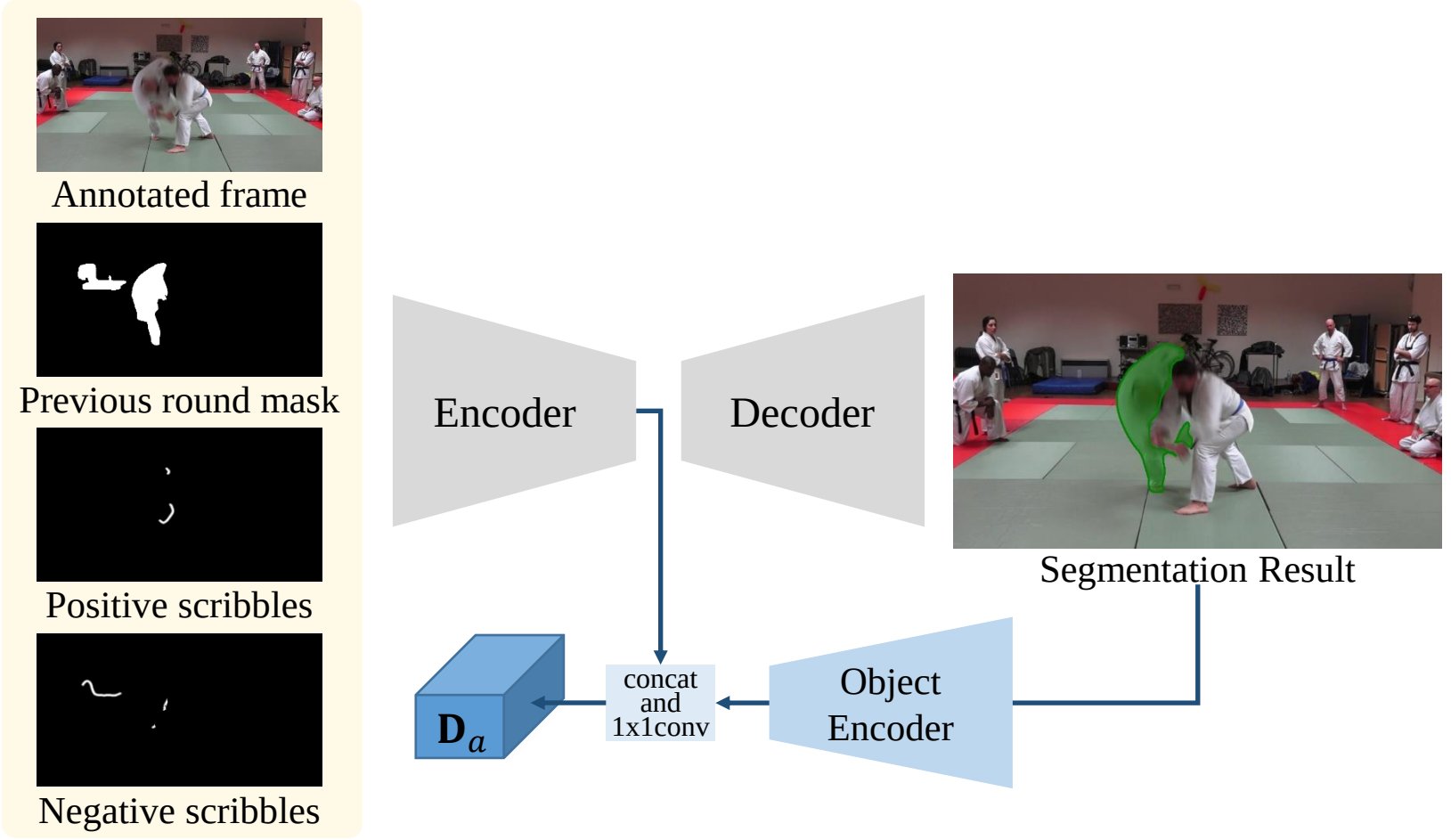
Encoder

Decoder

Segmentation Result

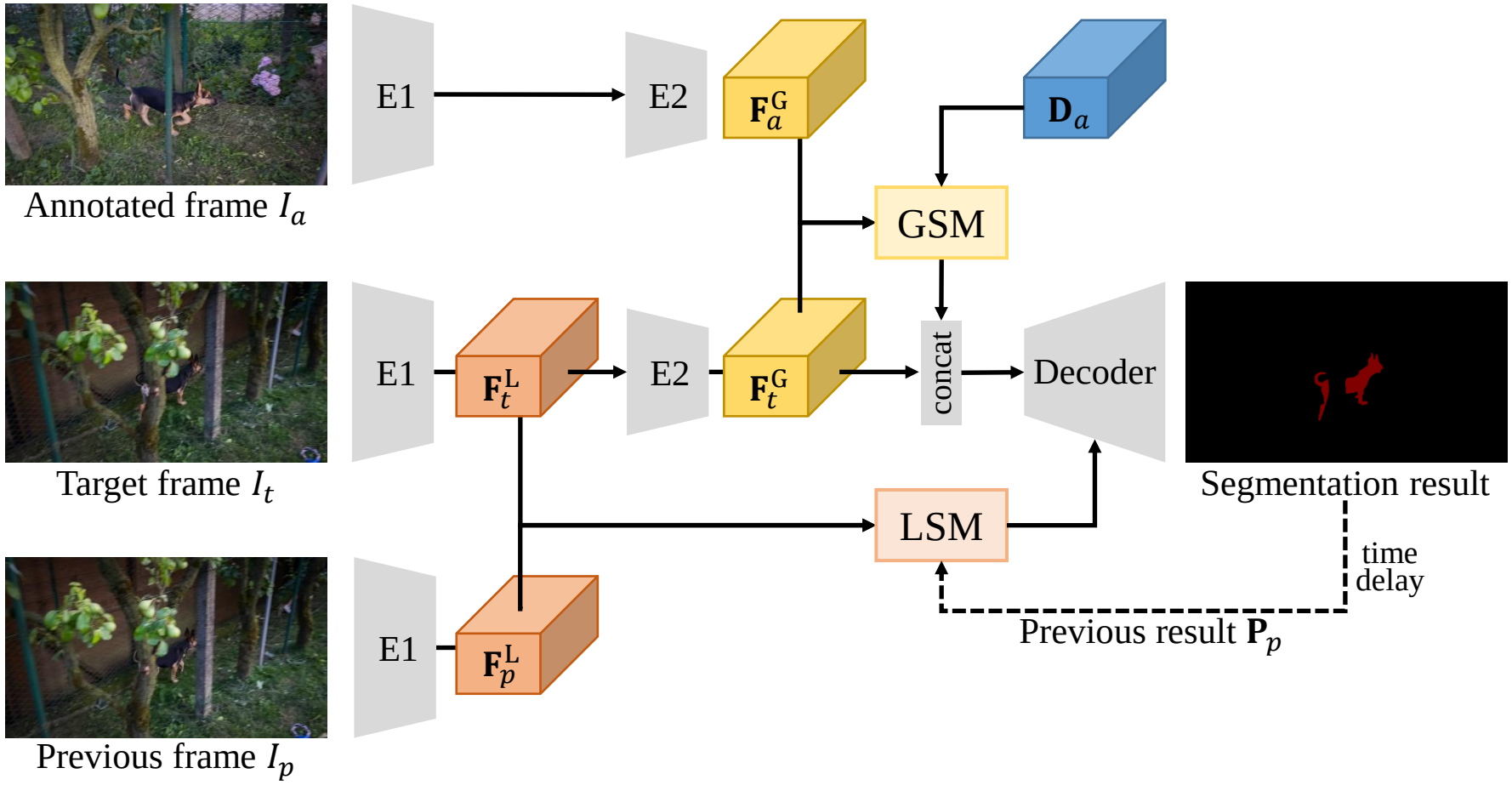
Network for interaction

- Network for interaction in the annotated frame



Network for transfer

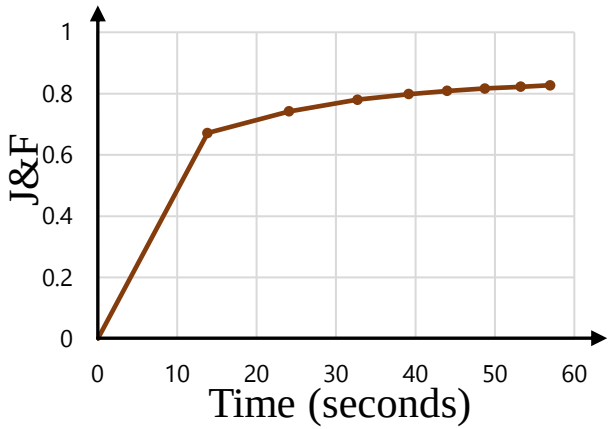
- Network architecture

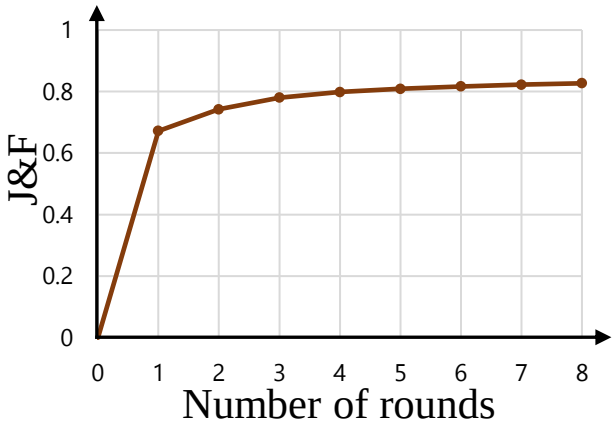


Experimental result

Experiment result

- DAVIS2017 validation set
  - Quantative result





Method	AUC-J	J@60s	AUC-J&F	J&F@60s
Proposed	**0.771**	**0.789**	**0.809**	**0.827**

Experiment result

- DAVIS2017 validation set
  - Qualitative result

Q&A

Thank you for your attention.