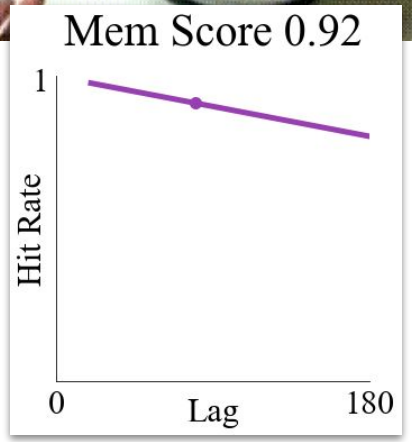
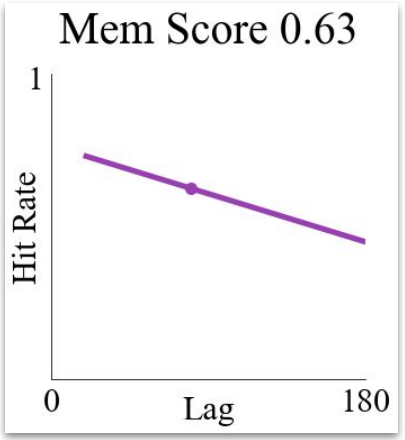
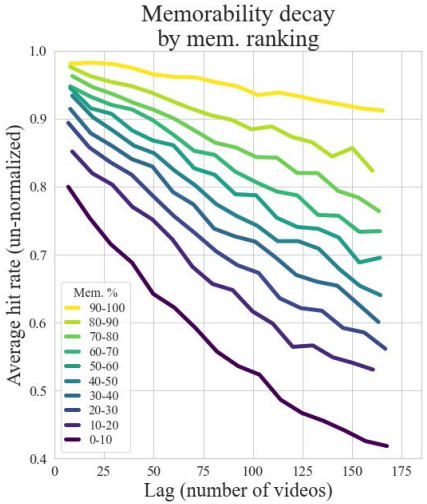

Mem Score 0.92
1
Hit Rate
0
Lag
180

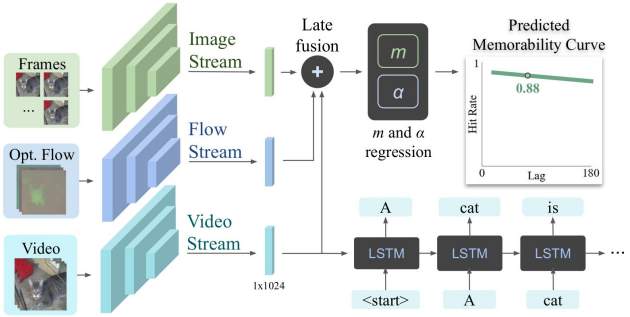

Mem Score 0.63
1
Hit Rate
0
Lag
180

Our Contributions

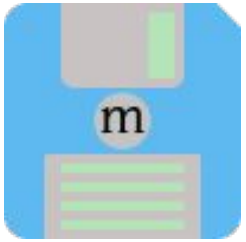
Memento10k



Modeling
Memorability Decay



SemanticMemNet



Multimodal Memorability

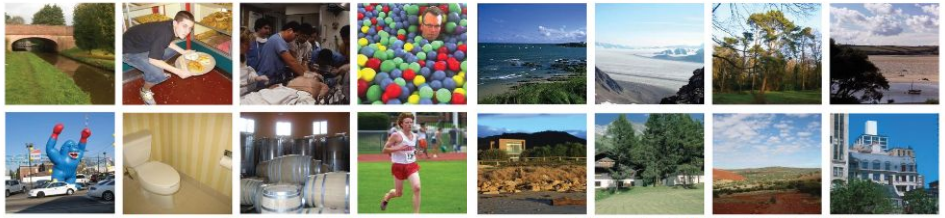
Modeling Effects of Semantics and Decay on Video Memorability

Anelise Newman*, Camilo Fosco*, Vincent Casser, Allen Lee,
Barry McNamara, and Aude Oliva

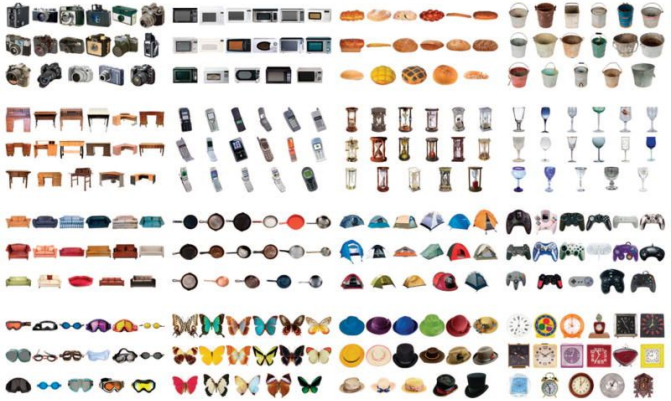
Code, data, and models available at

Memento.csail.mit.edu

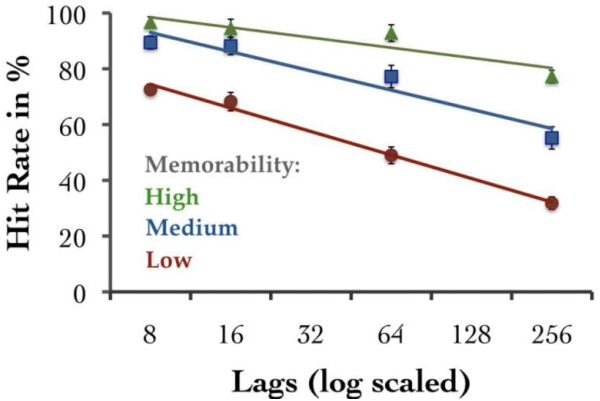
Related Work: Memorability in Cognitive Science



Isola 2011, CVPR



Konkle 2010, Journal of Experimental Psychology



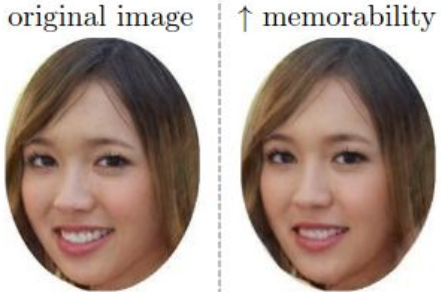
Vo 2017

Isola 2011, CVPR

Space for webcam

Related Work: Memorability in Computer Vision

Isola 2011, CVPR

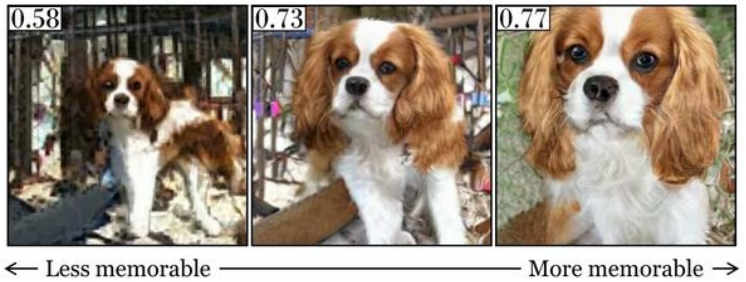


Khosla 2013, ICCV

Borkin 2013, InfoVis

LaMem

Khosla 2015, ICCV

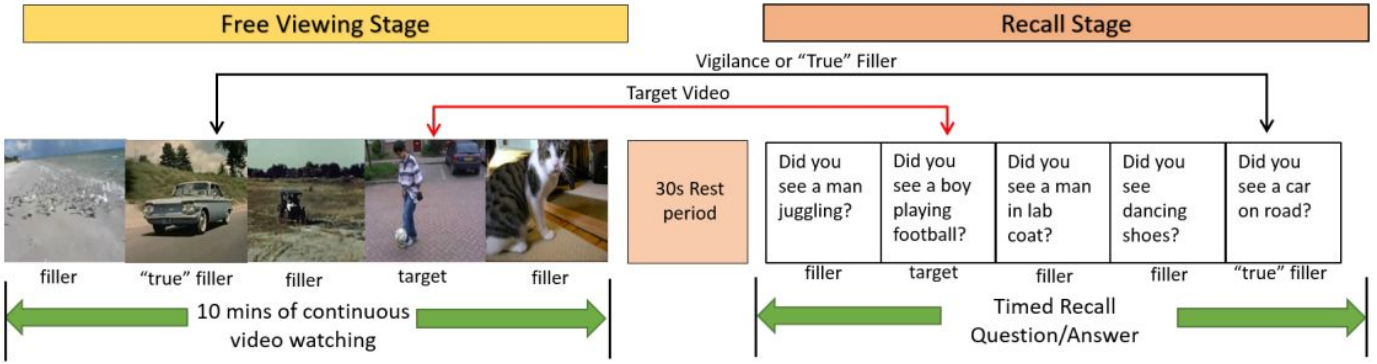


Goetschalckx 2019, ICCV

Related Work: Video Memorability

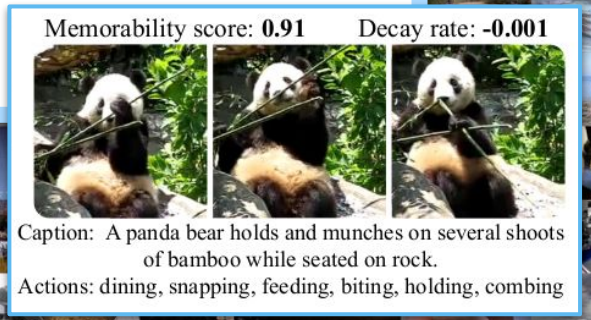
Cohendet 2019, ICCV

Shekhar 2017, ICCV

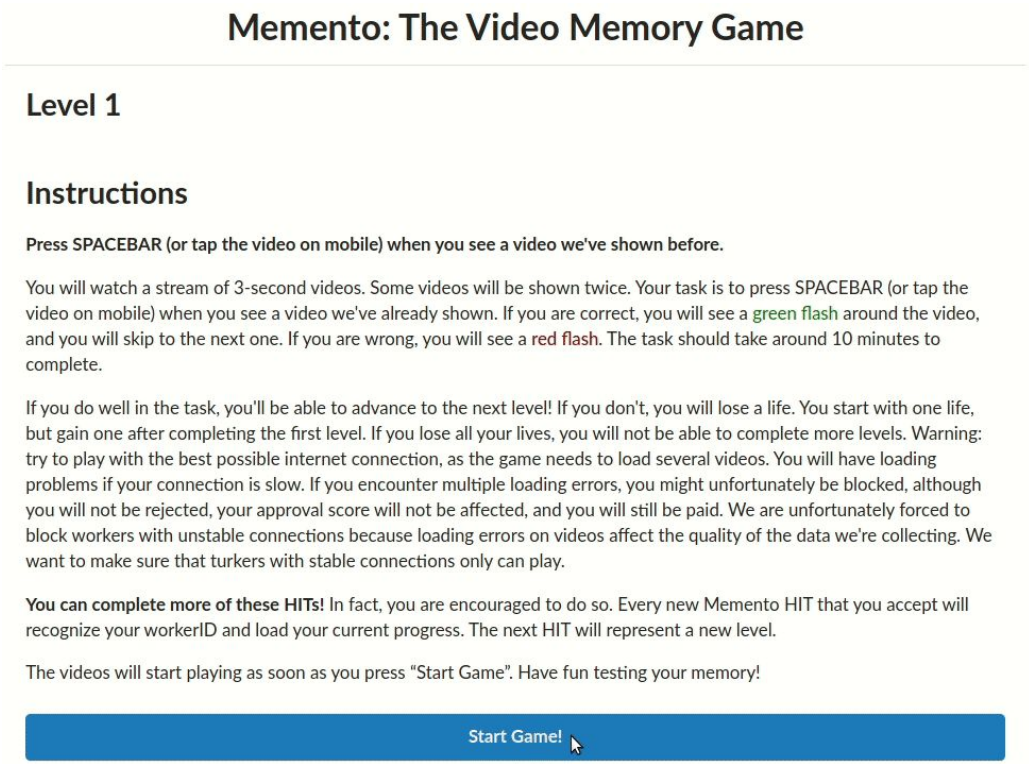


Memento10k

- 10k videos
- Nearly 1M human responses
- Semantic annotations: captions and action categories
- Variety of delays



The Memento Game



The Memento Game

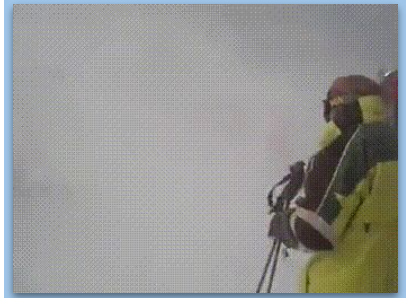
Hit Rate = # correct responses / # total responses

High Memorability



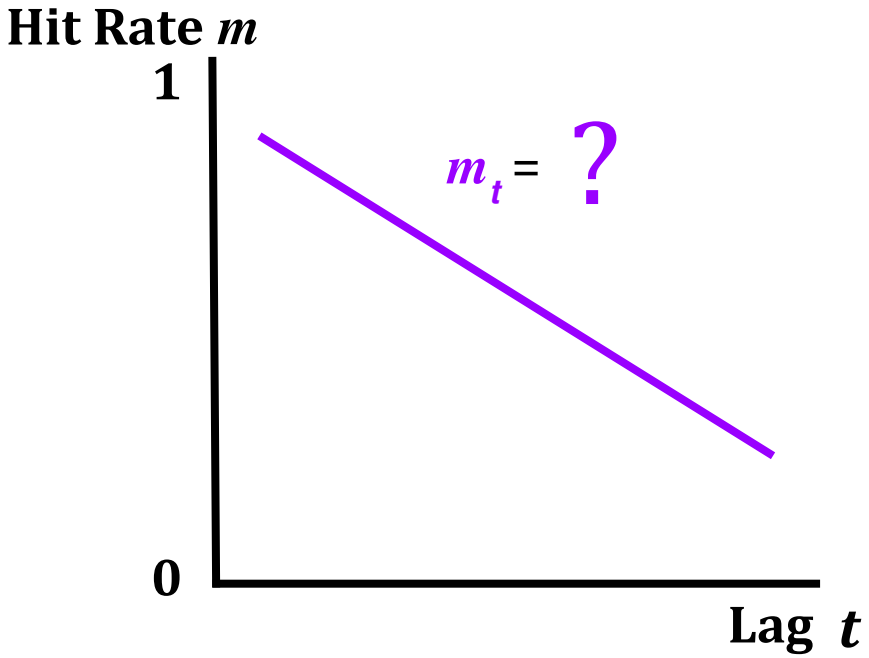



Low Memorability

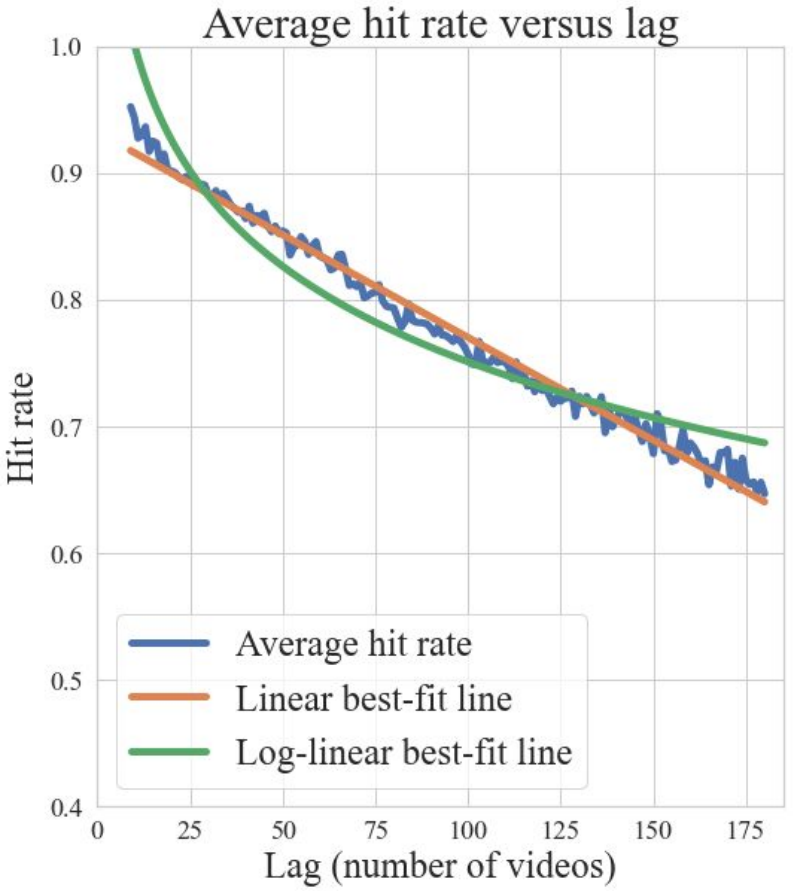





Measuring Memory Decay



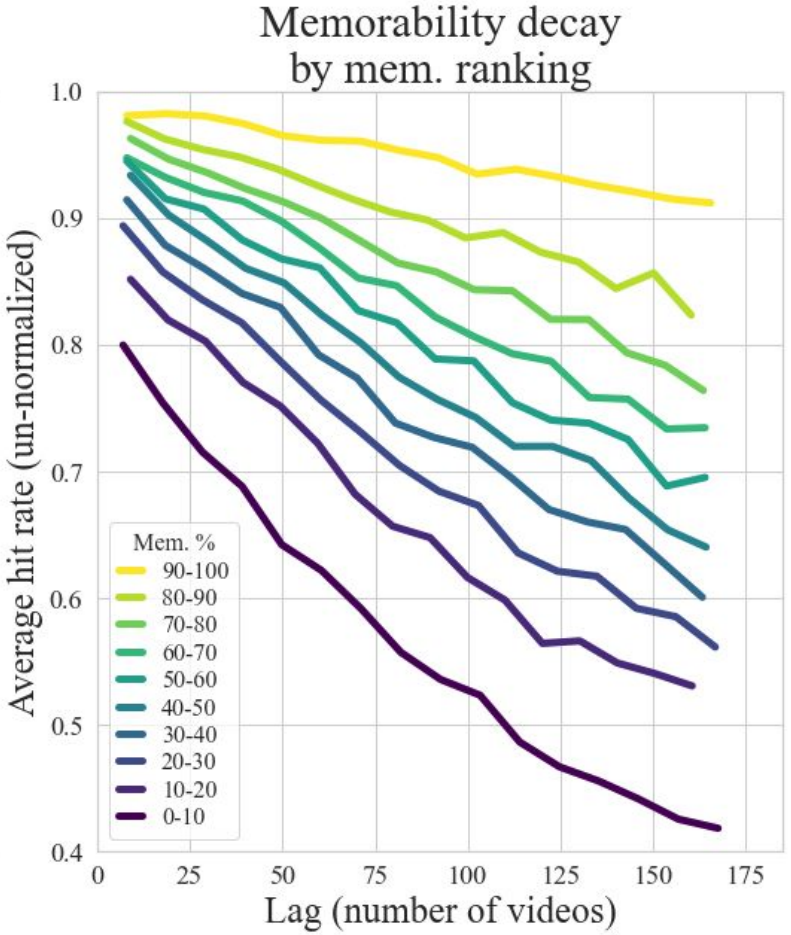
Linear Decay



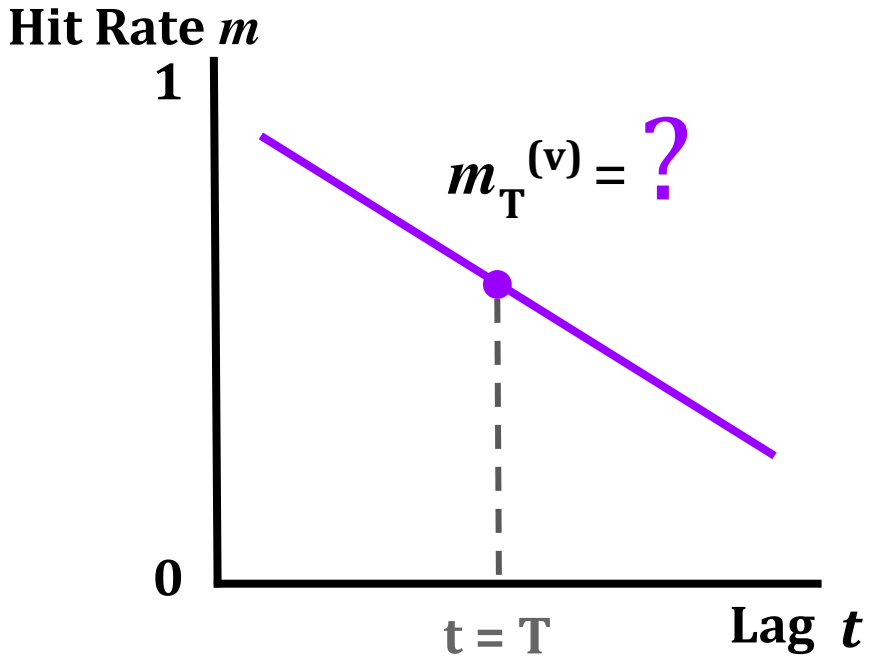
r = −0.953

r = −0.991

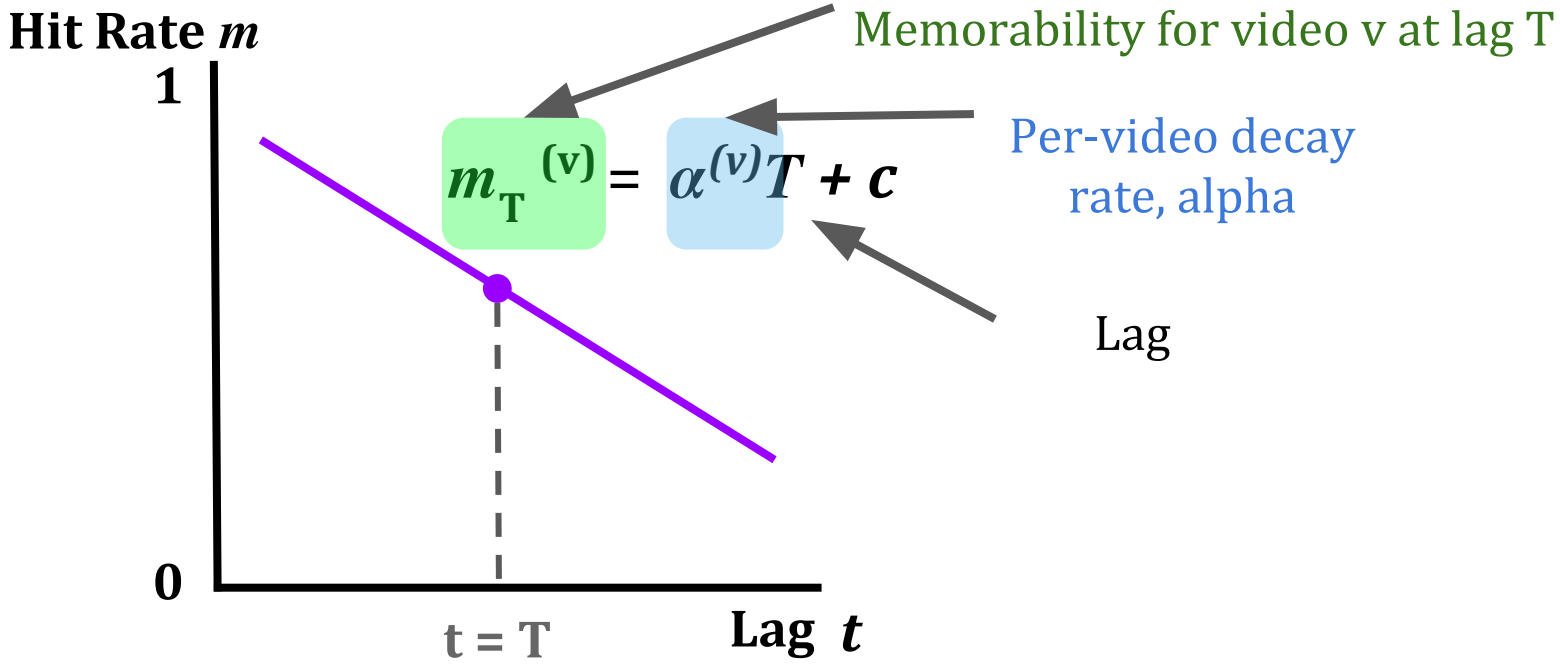
Different rates of decay



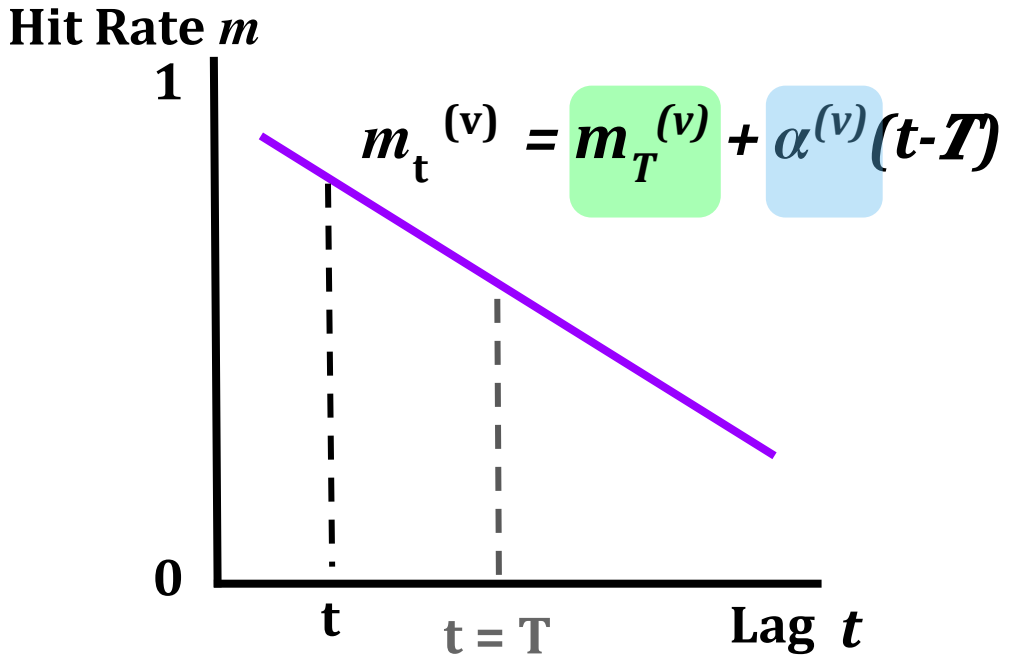
Calculating Memorability Scores



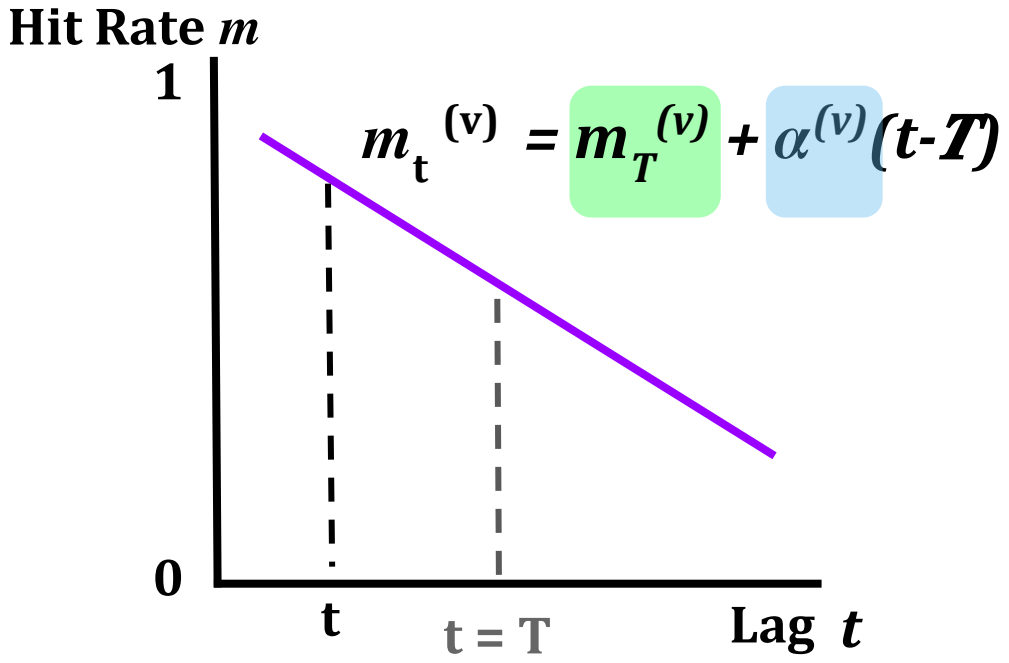
Calculating Memorability Scores



Calculating Memorability Scores



Calculating Memorability Scores



Define:

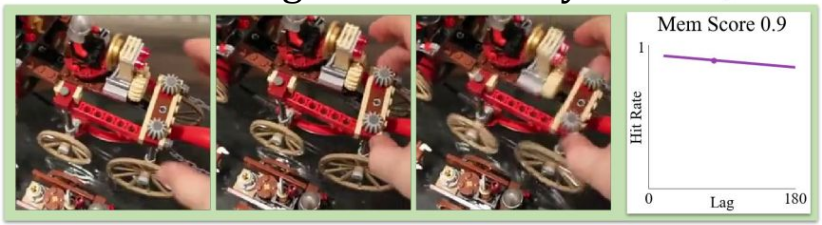
T = 80,

Memorability score

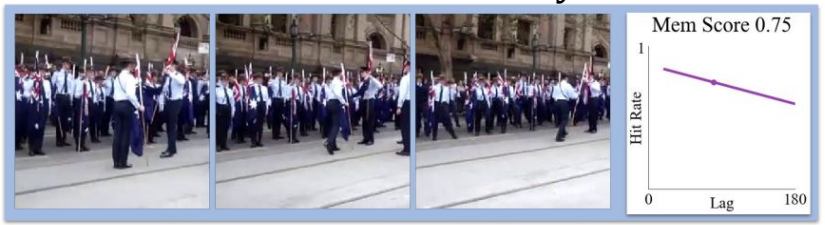
= m_{80}

Modeling Video Memorability

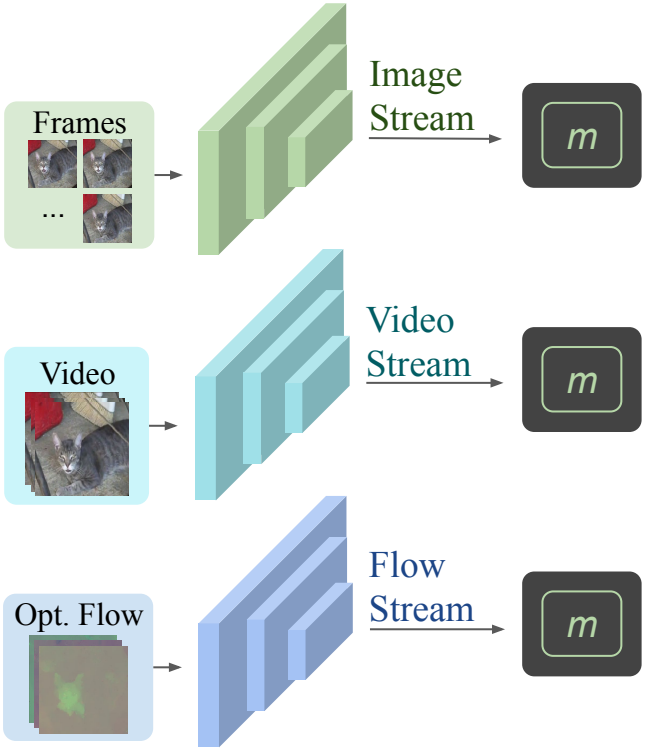
High Memorability



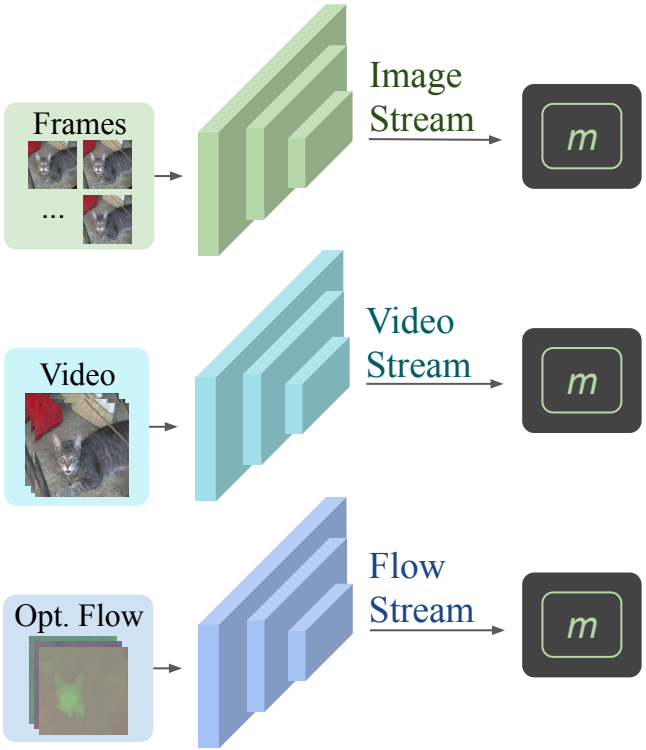
Low Memorability



Modeling Visual Features

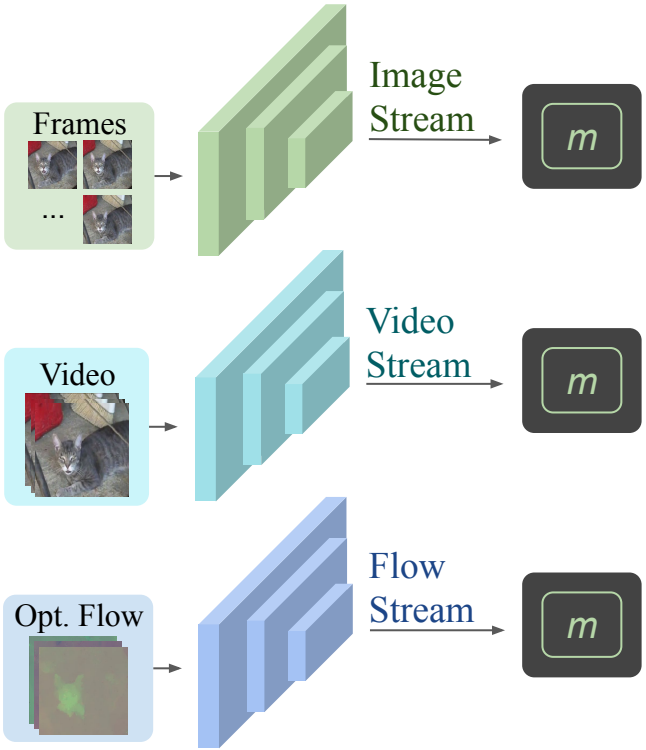


Modeling Visual Features



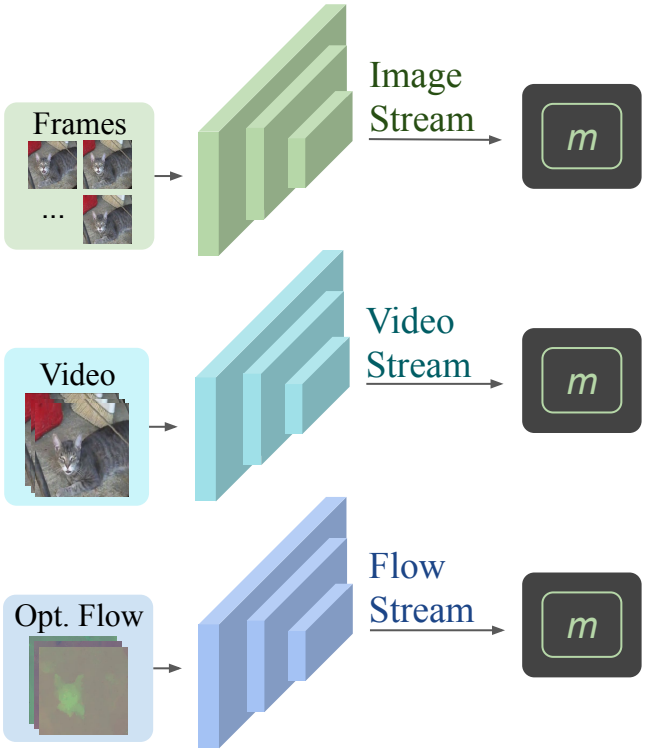
Approach	RC -- Memento10k test set	RC -- VideoMem val set
Human consistency	0.730	0.616
Frames	0.595	0.527
Video	0.596	0.492
Flow	0.579	0.425
Frames + Video + Flow	**0.659**	**0.555**

Modeling Visual Features



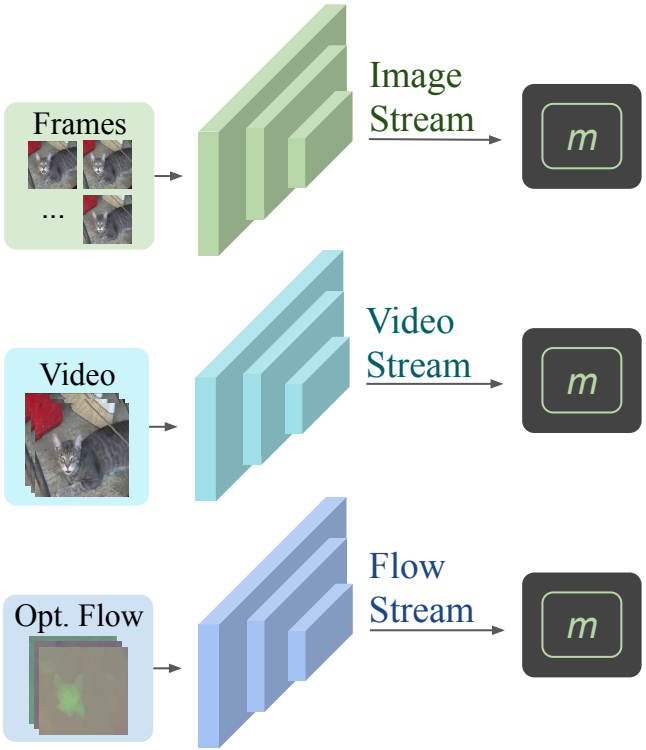
Approach	RC -- Memento10k test set	RC -- VideoMem val set
Human consistency	0.730	0.616
Frames	0.595	0.527
Video	0.596	0.492
Flow	0.579	0.425
Frames + Video + Flow	**0.659**	**0.555**

Modeling Visual Features



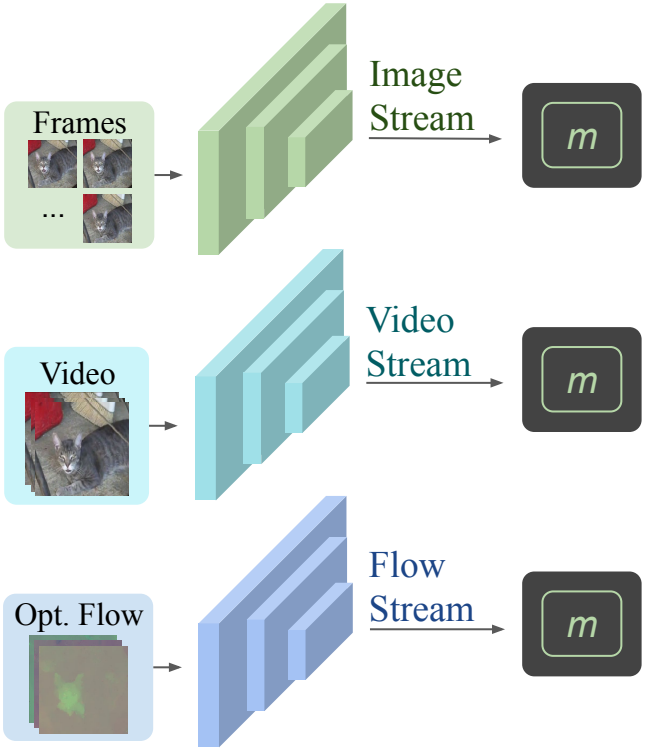
Approach	RC -- Memento10k test set	RC -- VideoMem val set
Human consistency	0.730	0.616
Frames	0.595	0.527
Video	0.596	0.492
Flow	0.579	0.425
Frames + Video + Flow	**0.659**	**0.555**

Modeling Visual Features

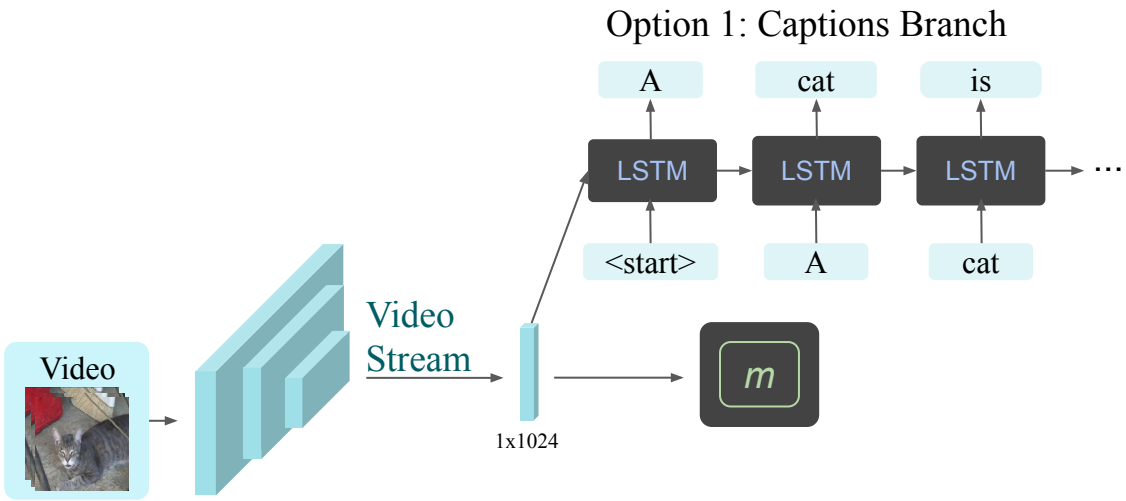


Approach	RC -- Memento10k test set	RC -- VideoMem val set
Human consistency	0.730	0.616
Frames	0.595	0.527
Video	0.596	0.492
Flow	0.579	0.425
Frames + Video + Flow	**0.659**	**0.555**

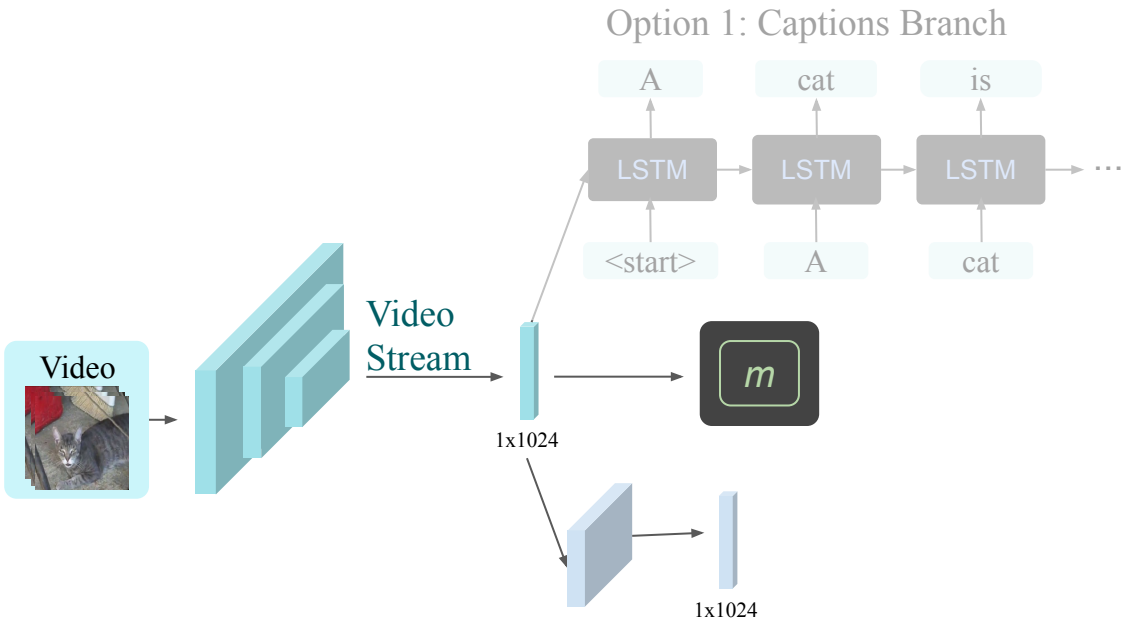
Modeling Semantic Features



Modeling Semantic Features

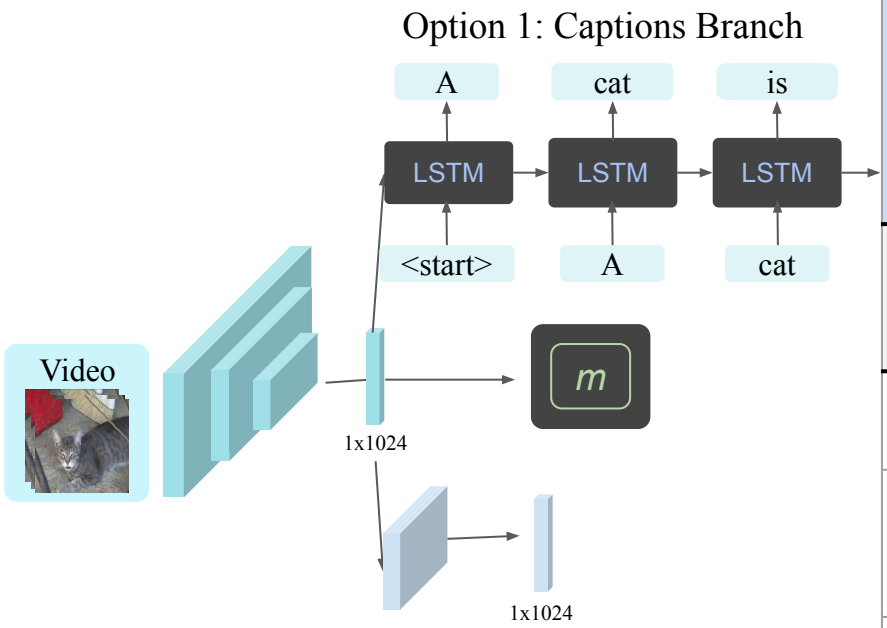


Modeling Semantic Features



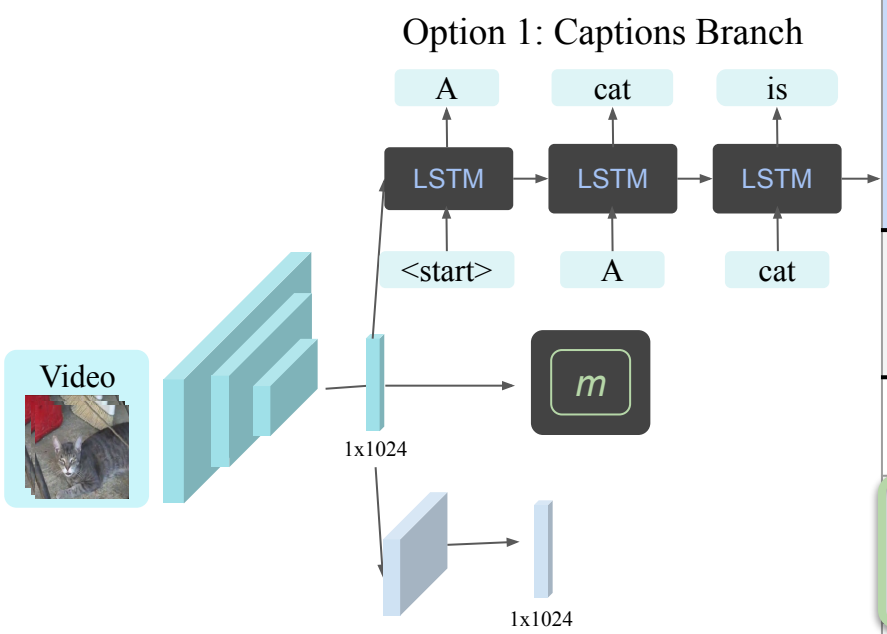
Option 2: Sentence Embedding

Modeling Semantic Features



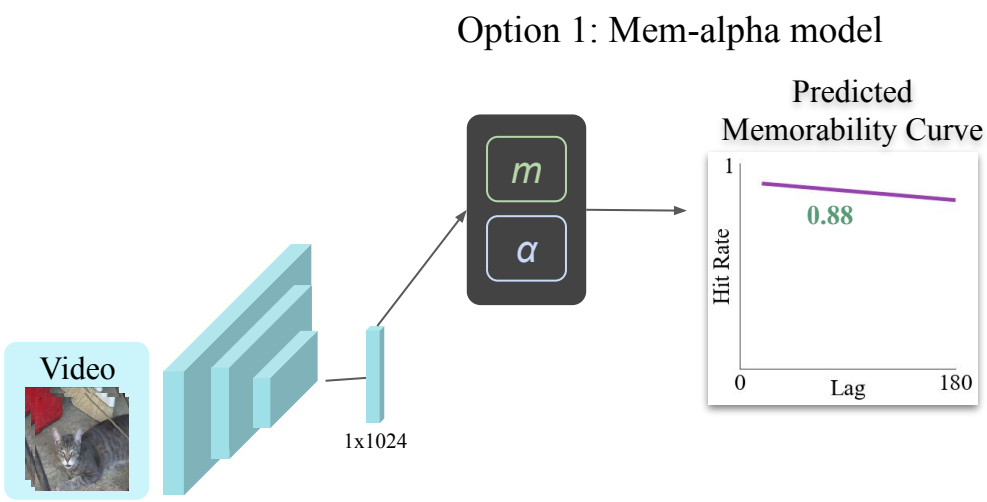
Approach	RC -- Memento 10k test set	RC -- VideoMem val set
Human consistency	0.730	0.616
Video	0.596	0.492
Video + Captions	**0.602**	**0.512**
Video + Triplet Loss	0.599	--

Modeling Semantic Features

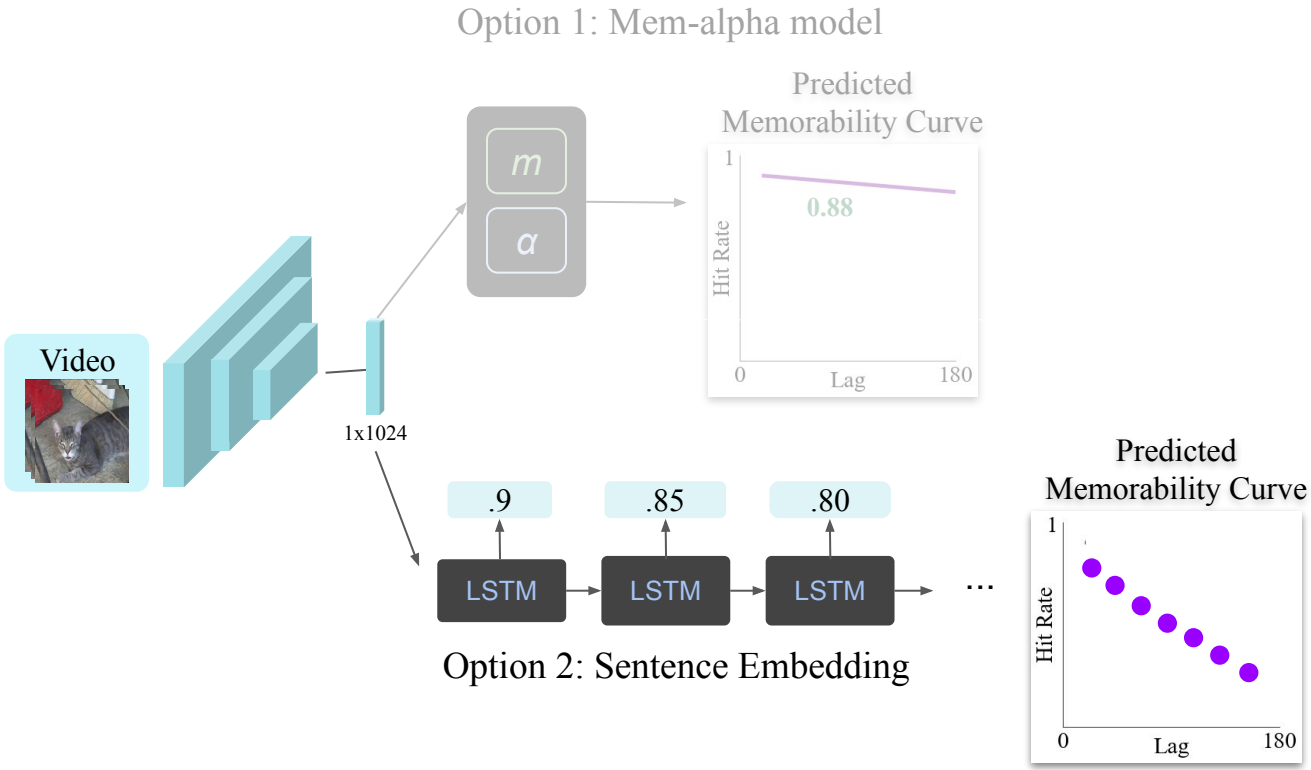


Approach	RC -- Memento 10k test set	RC -- VideoMem val set
Human consistency	0.730	0.616
Video	0.596	0.492
Video + Captions	**0.602**	**0.512**
Video + Triplet Loss	0.599	--

Modeling Memorability Decay

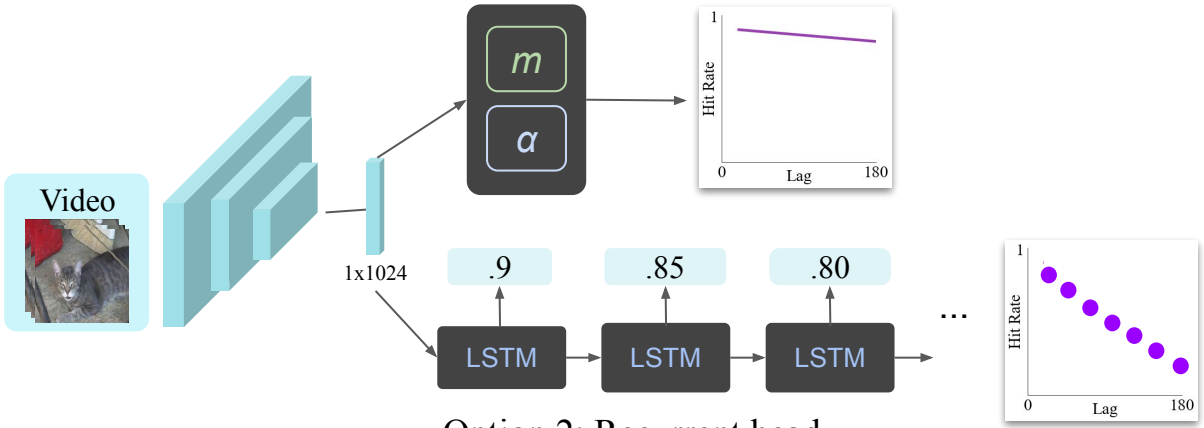


Modeling Memorability Decay



Modeling Memorability Decay

Option 1: Mem-alpha model

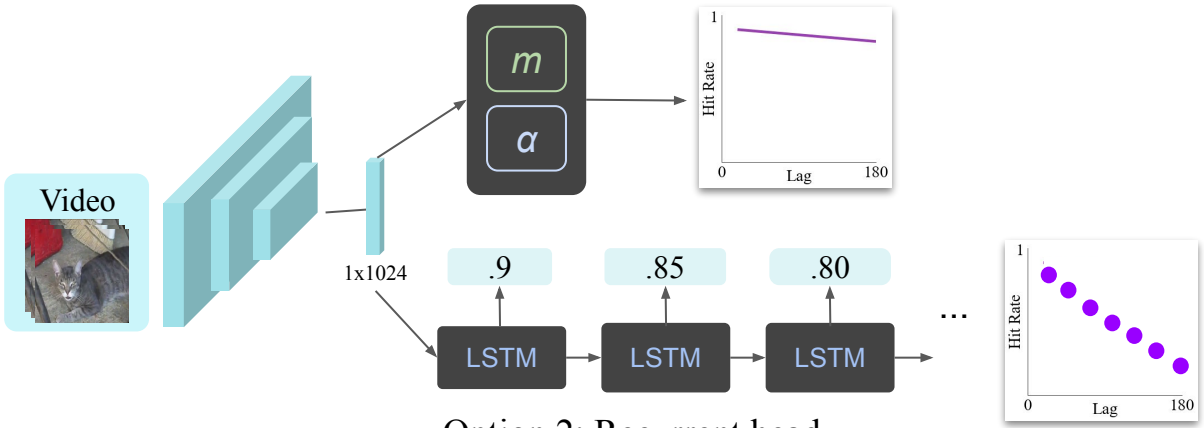


Option 2: Recurrent head

Approach	RC	R^2 t=40	t=80	t=160
Mem-alpha	**0.604**	0.146	0.227	0.121
Recurrent head	0.599	**0.298**	**0.364**	**0.219**

Modeling Memorability Decay

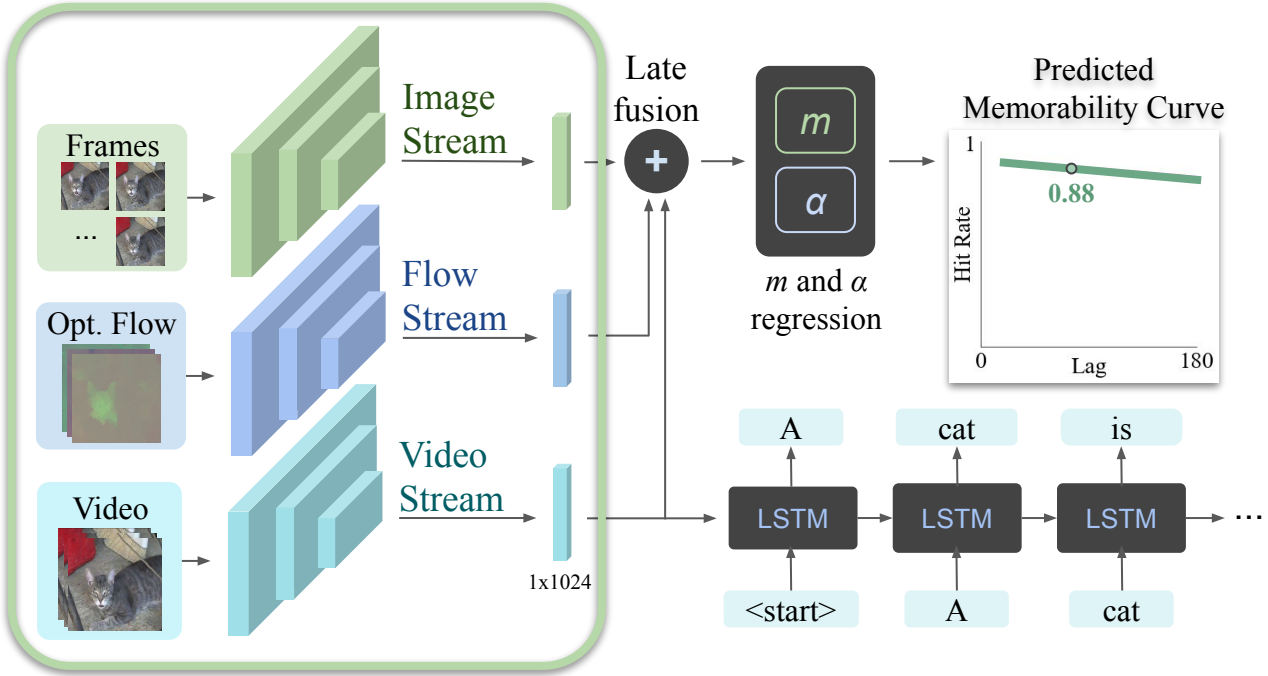
Option 1: Mem-alpha model



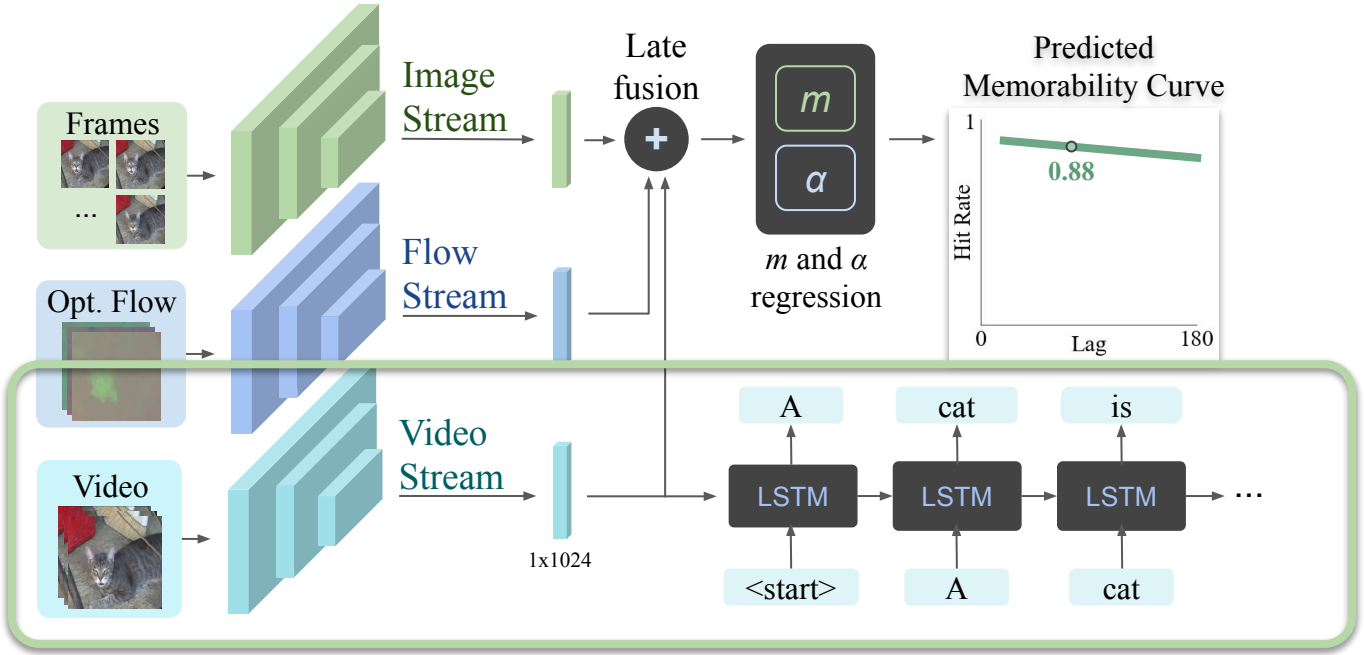
Option 2: Recurrent head

Approach	RC	R^2 t=40	t=80	t=160
Mem-alpha	**0.604**	0.146	0.227	0.121
Recurrent head	0.599	**0.298**	**0.364**	**0.219**

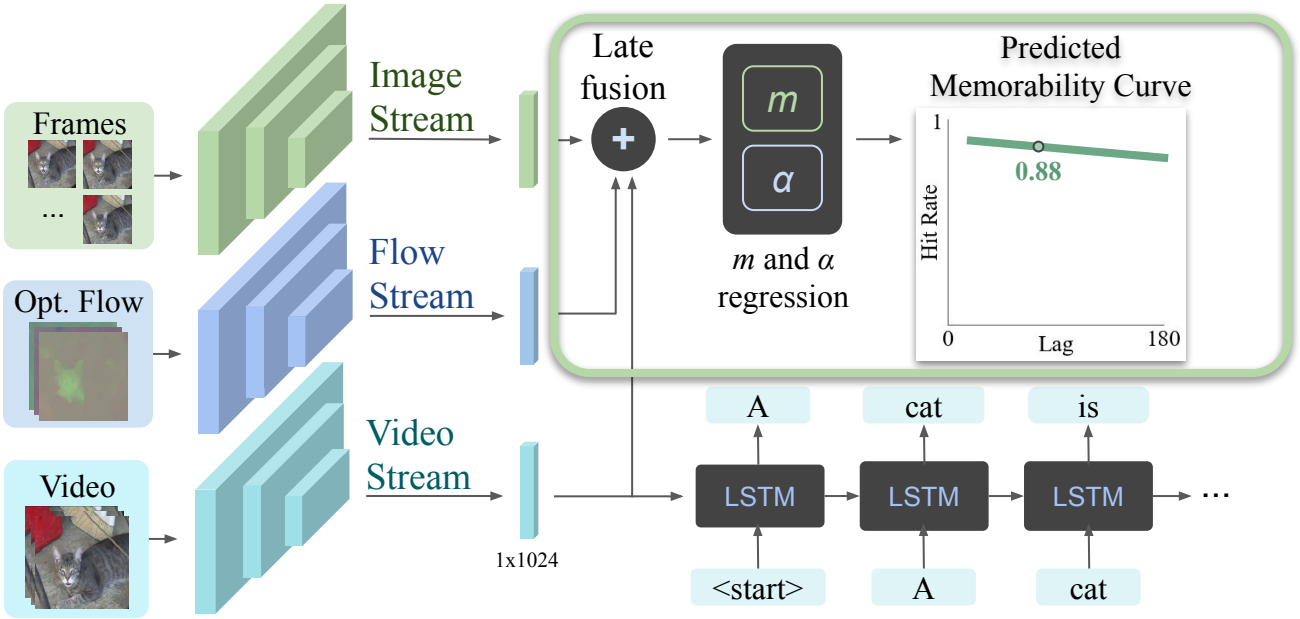
SemanticMemNet



SemanticMemNet



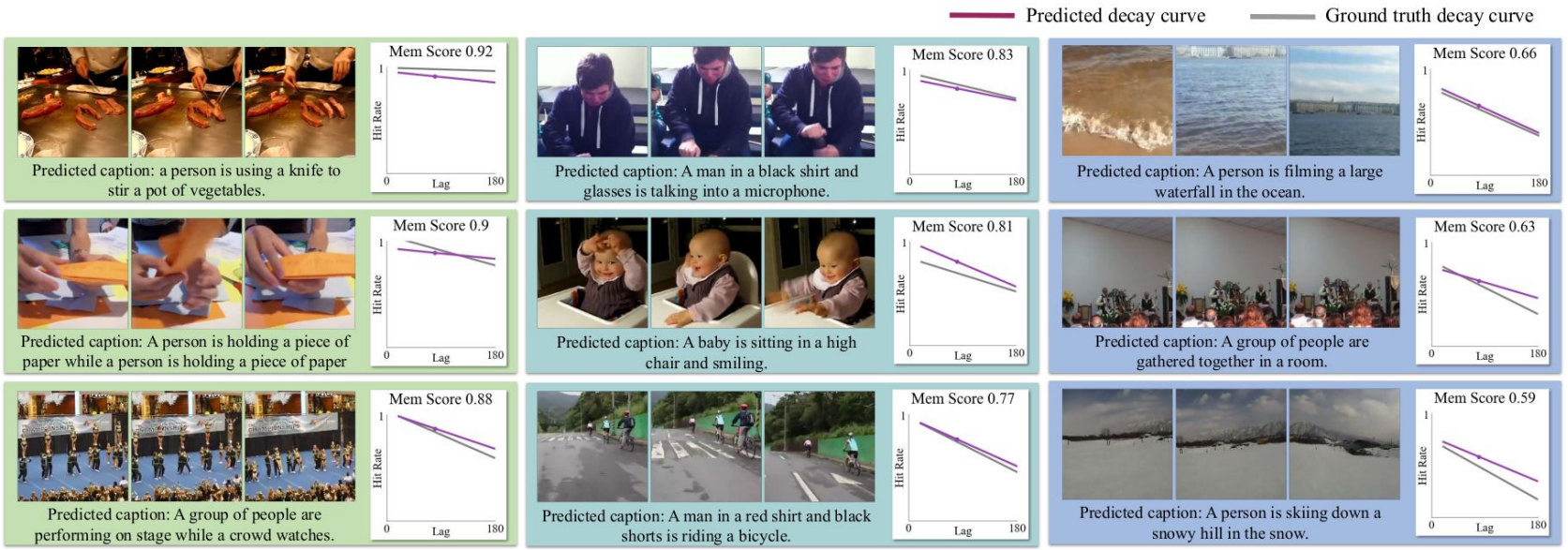
SemanticMemNet



SemanticMemNet Results

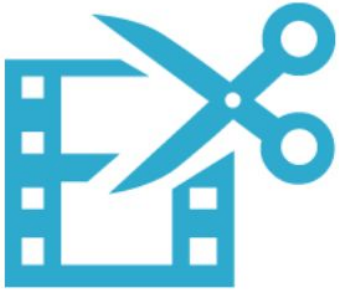
Approach	RC - Memento10k (test set)	RC - VideoMem (validation set)
Human consistency	0.730	0.616
MemNet Baseline [30]	0.485	0.425
Feature extraction + regression (as in [39])*	0.615	0.427
Cohendet et al. (ResNet3D) [13]	0.574	0.508
Cohendet et al. (Semantic)[13]	0.552	0.503
SemanticMemNet	**0.663**	**0.556**

SemanticMemNet Results



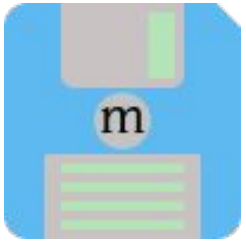
Future Directions

Education /
Content Creation



Video Editing /
Summarization

Video Selection



Multimodal Memorability

Modeling Effects of Semantics and Decay on Video Memorability

Anelise Newman*, Camilo Fosco*, Vincent Casser, Allen Lee,
Barry McNamara, and Aude Oliva

Code, data, and models available at

Memento.csail.mit.edu