ComputerVisionLab
Seoul National University

AIM 2020 Challenge on Video Temporal Super-Resolution

Sanghyun Son, Jaerin Lee, Seungjun Nah, †Radu Timofte, Kyoung Mu Lee

Department of ECE, ASRI, SNU, Korea
†CVL, ETH Zurich

August 28th, 2020

CVL Computer Vision Lab

ECCV'20 ONLINE 23-28 AUGUST 2020

AIM 2020

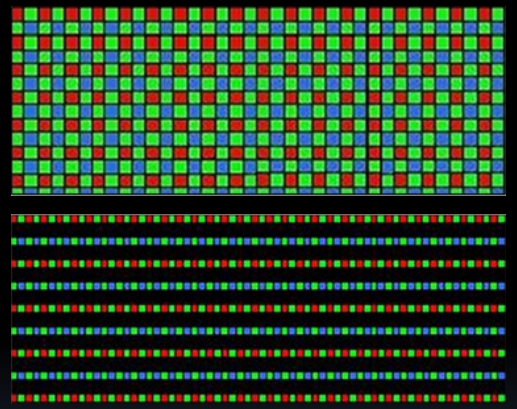

15 fps video
60 fps video
banila co.
Foam Machine
QX-1000

Recording High-Speed Videos

- **Hardware requirements**
  - Fast processor
  - Large & fast storage

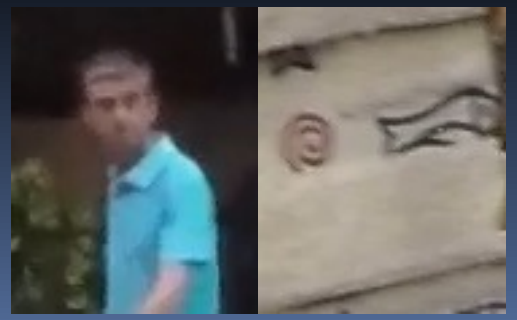
Large & Fast storage

- **Visual quality degradation**
  - Low effective resolution (line skips)
  - High compression
  - Short exposure & high noise



Bayer sensor full pattern

Line skipping at readout

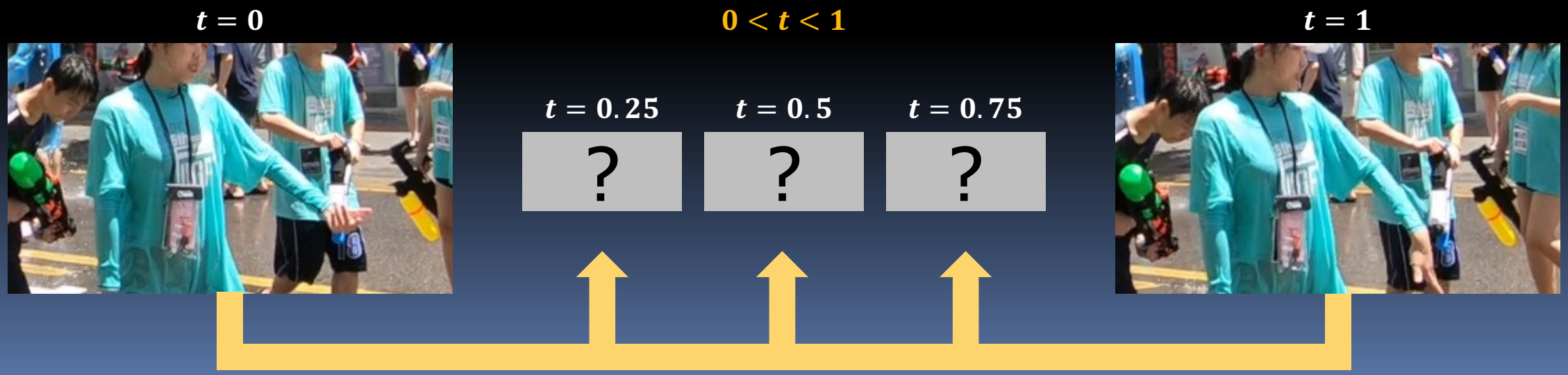


GoPro Hero4 240 fps

Video Frame Interpolation

- **Goal**
  - Enhance temporal smoothness by interpolating frames

- **Challenges**
  - Real-world motions are highly complex
  - Efficient algorithms are required to handle thousands of frames



REDS_VTSR Dataset

- Real-world dataset for AIM 2019 and 2020 VTSR Challenges
  - **Strong dynamics** with **nonlinear motion**
  - Diverse scenes and locations
  - Provides 15, 30, 60 fps videos for the same scene
  - High-quality, 1280×720 (HD) resolution


Zoo, Korea


Old Village, Korea


Street, Turkey


Harbor, Turkey


Uphill, Germany

VTSR Datasets: Quantitative Comparison

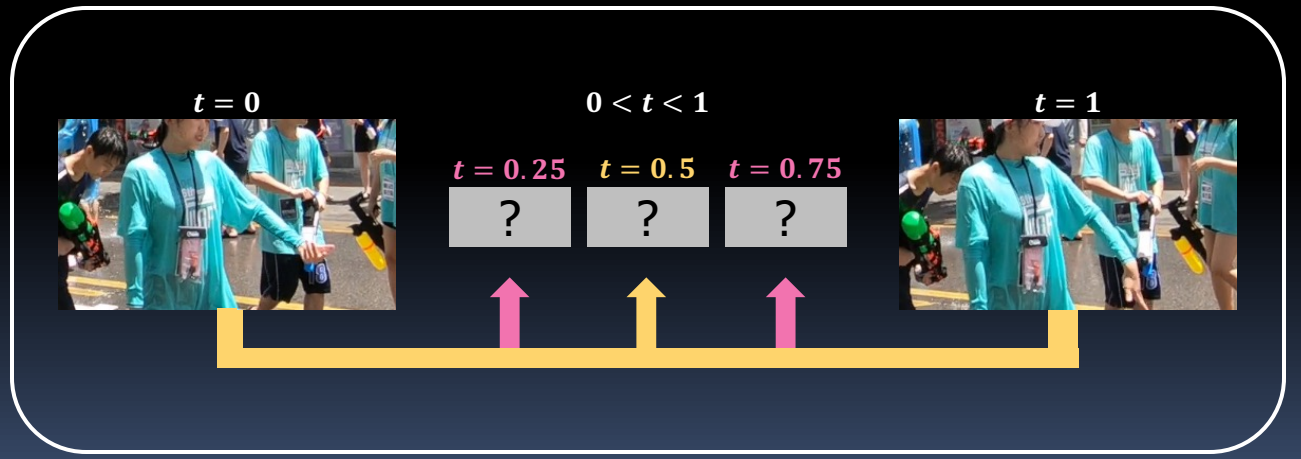
Dataset	Resolution	fps	#Sequences	#Frames	Note
For Training					
YouTube clips	Varying resolutions	Varying fps	-		Collected from YouTube
UCF101			13,320	-	
Vimeo-90k			73,171	219,513	Triplet dataset
KITTI raw			56	16,951	
GoPro (Nah *et al.**)	1280×720 (HD)		33	3,214	Captured by GoPro 4
DAVIS 2016+2017	720×480 (SD)		200	13,914	
REDS_VTSR	**1280×720 (HD)**	**Maximum 60 fps**	**240**	**43,200**	**Captured by GoPro 6**
For Evaluation (Validation + Test)					
Middlebury	Varying resolutions	Varying fps	8	58	
THUMOS 2015			5,613	-	
SlowFlow			46	-	Captured by TS5Q camera
Sintel			19	-	Synthetic
HD			11	-	
REDS_VTSR	**1280×720 (HD)**	**Maximum 60 fps**	**60**	**10,800**	**Captured by GoPro 6**

*Nah *et al.*, "Deep Multi-scale Convolutional Neural Network for Dynamic Scene Deblurring," in CVPR 2017.

AIM 2020 VTSR Challenge

- **Goal**
  - Encourage development of the **state-of-the-art** Video Temporal Super-Resolution (a.k.a. frame interpolation) algorithm
  - Compare various methods in a unified environment

- **Task**
  - 15 fps video → 30 fps video
    → 60 fps video



AIM 2020 VTSR Challenge

- **Evaluation**
  - **240** training + **30** validation sequences provided (with ground-truth)
  - Tested on disjoint **30** sequences (no ground-truth provided)

- **Metrics**
  - **PSNR** / SSIM
  - Reproducibility / Runtime (for reference)
  - LPIPS* (experimental)
    - Reference-based perceptual metric
    - Shows better correlation with human perception

*Zhang *et al.*, "The Unreasonable Effectiveness of Deep Features as a Perceptual Metric," in CVPR 2018.

AIM 2020 VTSR Challenge

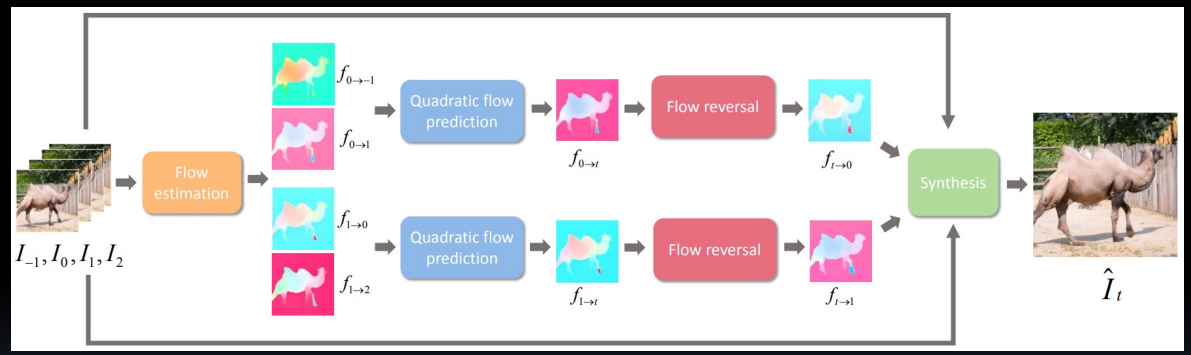
- **Challenge schedule**
  - May 1st, 2020: Development phase starts
  - July 10th, 2020: Testing phase starts
  - July 17th, 2020: Testing phase finishes
  - July 19th, 2020: Final rank announcement

- **Participants**
  - **68** participants in the CodaLab
  - **5** teams submitted their final solutions (one withdrawn)

Review: AIM 2019 VTSR Challenge

- The winner: **Quadratic Video Interpolation (QVI)**

from the **SenseSloMo** team



$$f_{0\to t} = \frac{f_{0\to1} + f_{0\to-1}}{2} t^2 + \frac{f_{0\to1} - f_{0\to-1}}{2} t$$

Li *et al.*, "Quadratic Video Interpolation for VTSR Challenge," in AIM 2019 Workshop in conjunction with ICCV 2019.
Xu *et al.*, "Quadratic Video Interpolation," in NeurIPS 2019.

AIM 2020 VTSR Challenge: Methods

- **XPixel (Challenge Winner)**
  - Enhanced Quadratic Video Interpolation
- **KAIST-VICLAB**
  - Quadratic Video Frame Interpolation with Multi-frame Synthesis Network
- **BOE-IOT-AIBD**
  - Multi Scale Quadratic Interpolation
- **TTI**
  - STARnet*

- **Top 3** methods are based on the **QVI** model!

*Haris *et al.*, "Space-Time-Aware Multi-Resolution Video Enhancement," in CVPR 2020.

AIM 2020 VTSR Challenge: Results

All values are reproduced by challenge organizers

15 fps → 30 fps	PSNR↑	SSIM↑	LPIPS↓*	Reproducibility	Runtime**
XPixel (Winner)	**24.78**	0.7118	0.268	O	12.4s / frame
KAIST-VICLAB	24.69	0.7142	**0.222**	O	1.5s / frame
BOE-IOT-AIBD	24.49	0.7034	0.249	O	1.0s / frame
TTI	23.59	0.6720	0.289	O	6.5s / frame
QVI (AIM 2019 Winner)	24.56	0.7065	-	O	-

15 fps → 60 fps	PSNR↑	SSIM↑	LPIPS↓*	Reproducibility	Runtime**
XPixel (Winner)	**25.69**	0.7425	0.214	O	-
KAIST-VICLAB	25.61	0.7462	**0.181**	O	-
BOE-IOT-AIBD	25.27	0.7269	0.230	O	-
TTI	24.36	0.6995	0.253	O	-
QVI (AIM 2019 Winner)	25.47	0.7383	-	O	-

*Experimental. Zhang *et al.*, "The Unreasonable Effectiveness of Deep Features as a Perceptual Metric," in CVPR 2018.
Measured on a single **Quadro RTX 8000 GPU for team TTI (due to the memory requirements), **RTX 2080 Ti** for the other teams.


banila co.
삼계탕
GT


CAFFÉ PASCUCCI
banila co.
PREMIUM STORE
삼계탕
XPixel


banila co.
KAIST-VICLAB

AIM 2020 VTSR Challenge: Results

- Frames with high performance variances
  - One shows superior performances than the others

Seq. 16 Frame 358
15 fps → 60 fps

GT
(PSNR/LPIPS)

XPixel
(**32.64**/**0.072**)

KAIST-VICLAB
(22.42/0.242)

BOE-IOT-AIBD
(23.86/0.254)

TTI
(30.64/0.168)

AIM 2020 VTSR Challenge: Results

- Frames with high performance variances
  - One shows superior performances than the others

Seq. 19 Frame 182
15 fps → 60 fps

GT
(PSNR/LPIPS)

XPixel
(29.33/0.171)

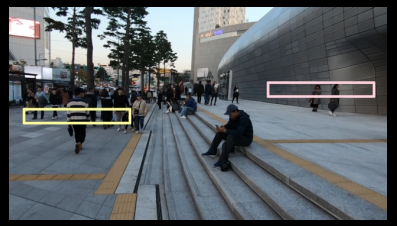
KAIST-VICLAB
(**30.19/0.143**)

BOE-IOT-AIBD
(27.89/0.197)

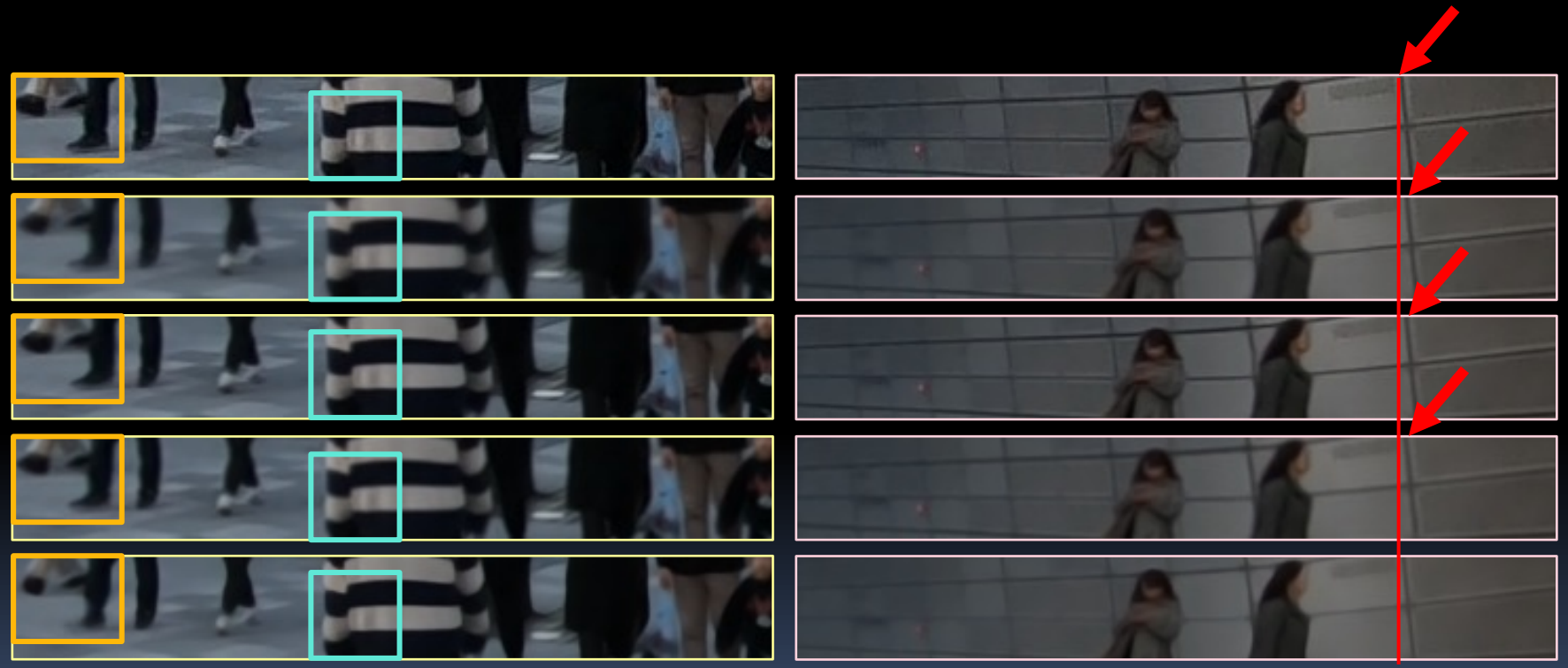
TTI
(20.30/0.226)

AIM 2020 VTSR Challenge: Results

- Frames with high performance variances
  - One shows superior performances than the others



Seq. 29 Frame 134
15 fps → 60 fps



GT
(PSNR/LPIPS)

XPixel
(23.33/0.167)

KAIST-VICLAB
(22.92/**0.100**)

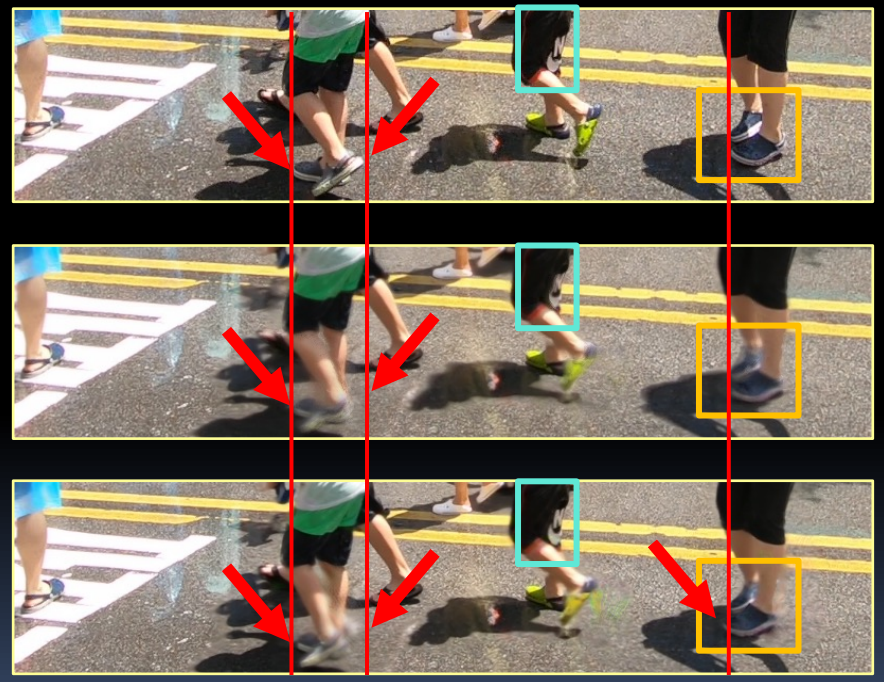
BOE-IOT-AIBD
(23.86/0.166)

TTI
(**29.11**/0.158)

AIM 2020 VTSR Challenge: Results

- PSNR vs. LPIPS (Experimental)

Seq. 7 Frame 358
15 fps → 60 fps



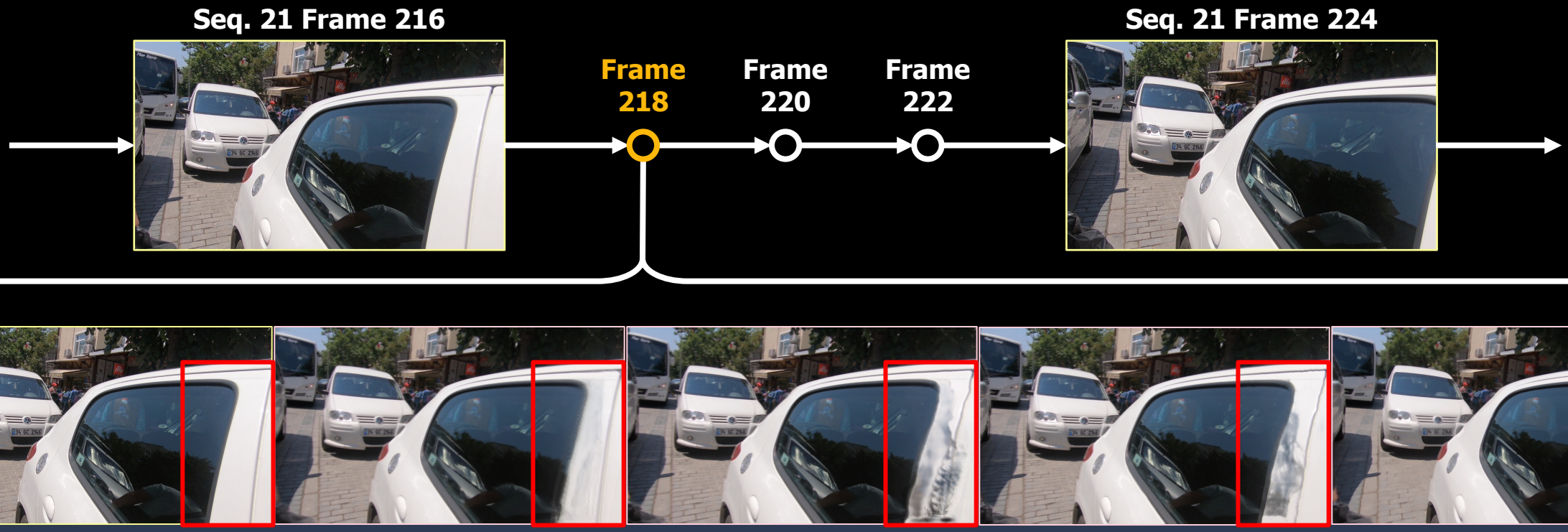
GT
(PSNR/LPIPS)

XPixel
(**27.04**/0.085)

KAIST-VICLAB
(25.67/**0.070**)

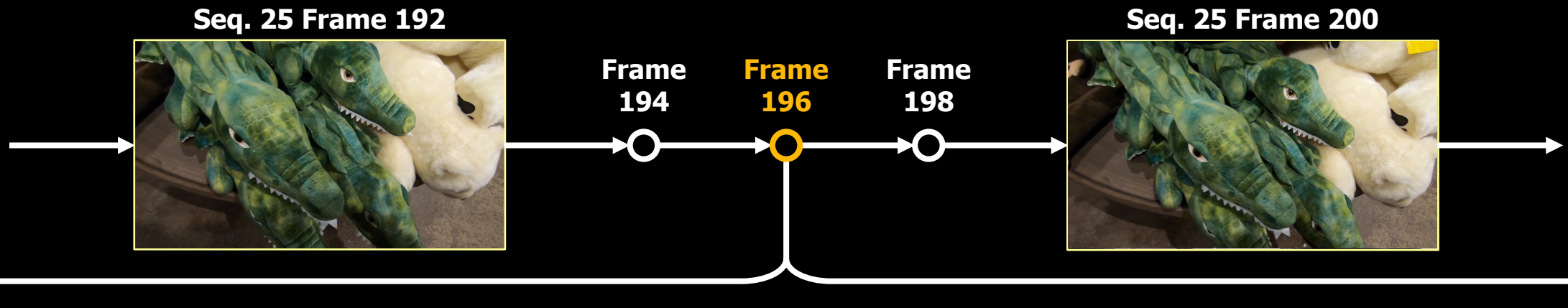
AIM 2020 VTSR Challenge: Results

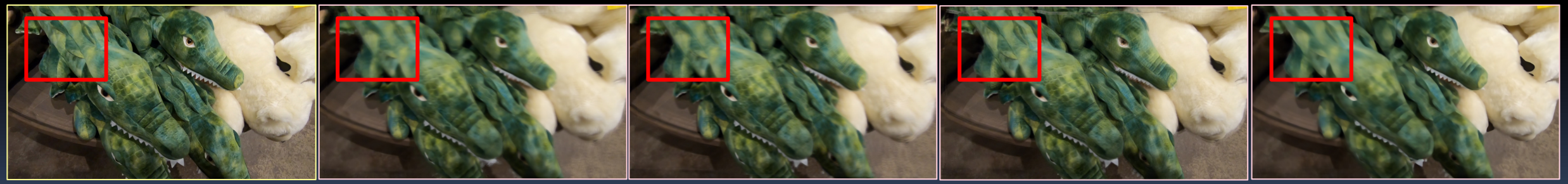
- Failure cases
  - Still have rooms for improvement!



AIM 2020 VTSR Challenge: Results

- Failure cases
  - Still have rooms for improvement!





GT
(PSNR/LPIPS)

XPixel
(24.25/0.416)

KAIST-VICLAB
(23.43/0.417)

BOE-IOT-AIBD
(22.70/**0.267**)

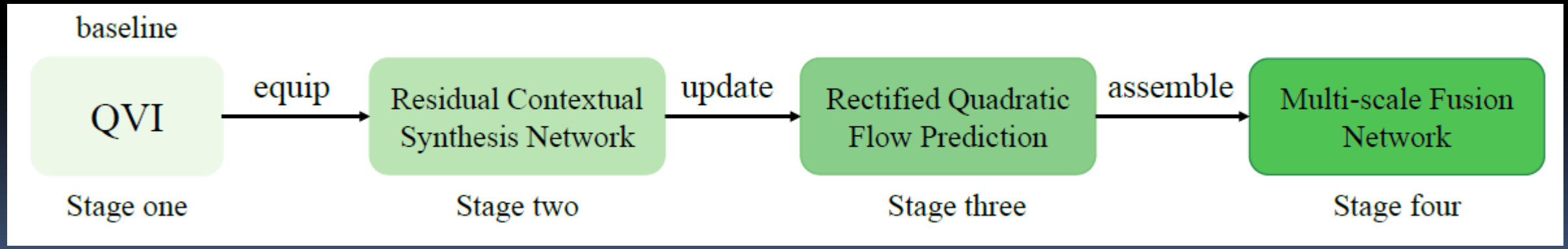
TTI
(**28.41**/0.354)

The Winning Method

Enhanced Quadratic Video Interpolation

from the **XPixel** team

Yihao Liu, Xie Liangbin, Li Siyao, Wenxiu Sun,
Yu Qiao, and Chao Dong



Sponsors


HUAWEI


Google


Qualcomm


NVIDIA


MEDIATEK


CVL
Computer
Vision
Lab

Thank you

https://data.vision.ee.ethz.ch/cvl/aim20/

Contact:
(AIM 2020 VTSR official) aim2020.vtsr@gmail.com
(Sanghyun Son) sonsang35@gmail.com