RoVISQ: Reduction of Video Service Quality via Adversarial Attacks on Deep Learning-based Video Compression

Jung-Woo Chang[1], Mojan Javaheripi[1], Seira Hidano[2], Farinaz Koushanfar[1]

[1]University of California, San Diego

[2]KDDI Research, Inc.

Introduction

- **Video traffic** has experienced an even higher growth with the advent of streaming services.
- Recent developments in deep learning (DL) have given rise to various video analytics such as health care diagnosis.

Remote vehicle control

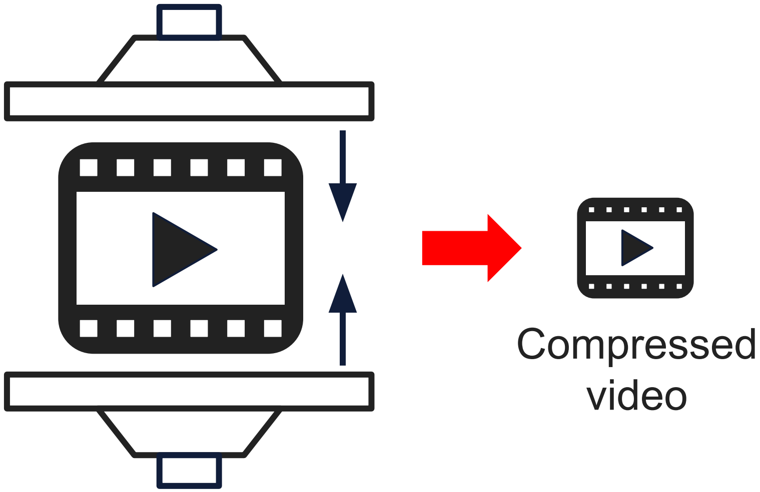
Video Streaming

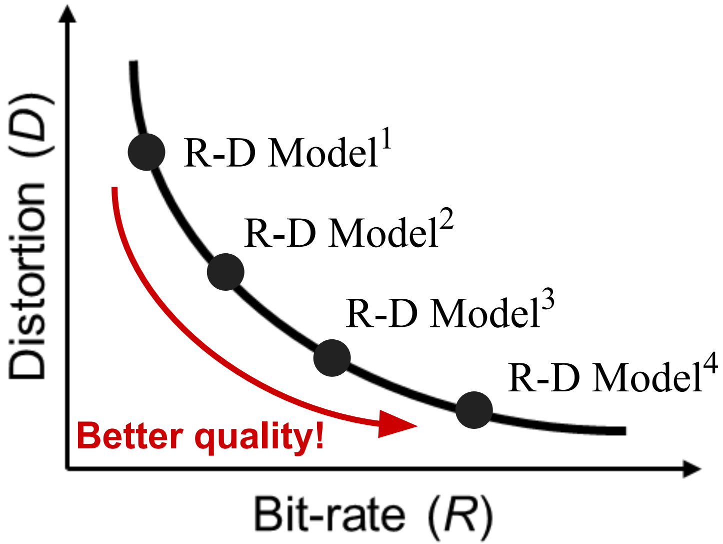
Metaverse

Health care diagnosis

Video Compression

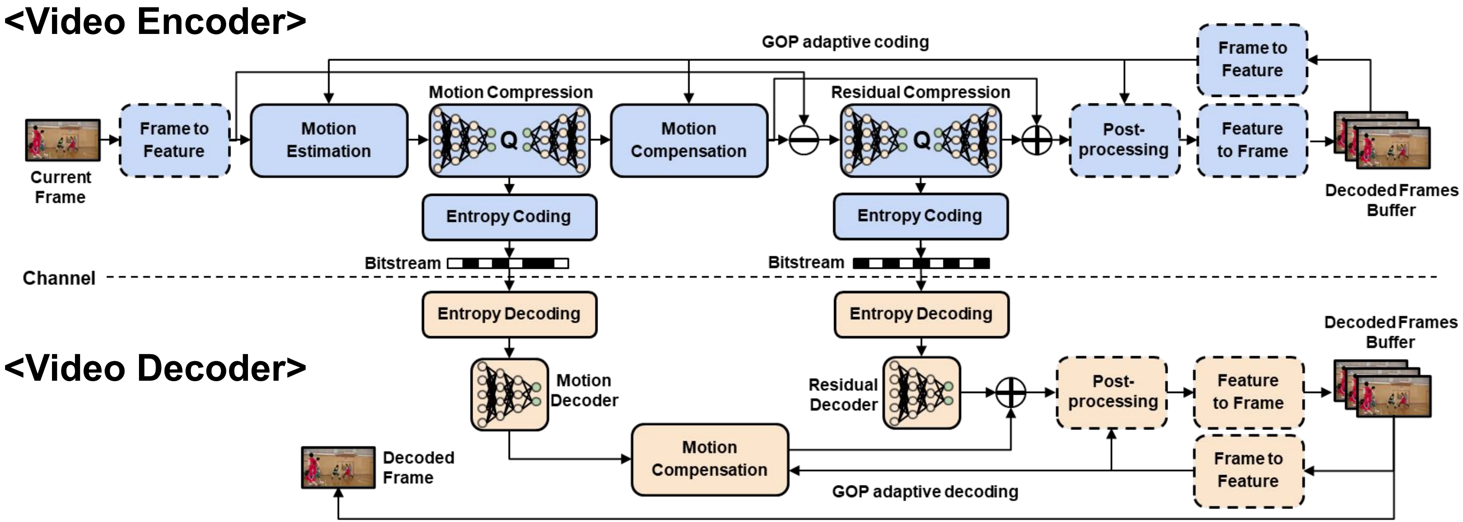
- In order to maximize the quality of experience (QoE), **video compression** is a key enabler for the aforesaid applications.
- Video compression employs rate-distortion (R-D) optimization to adapt to different **bandwidth constraints.**
    - Lower D requires higher R.





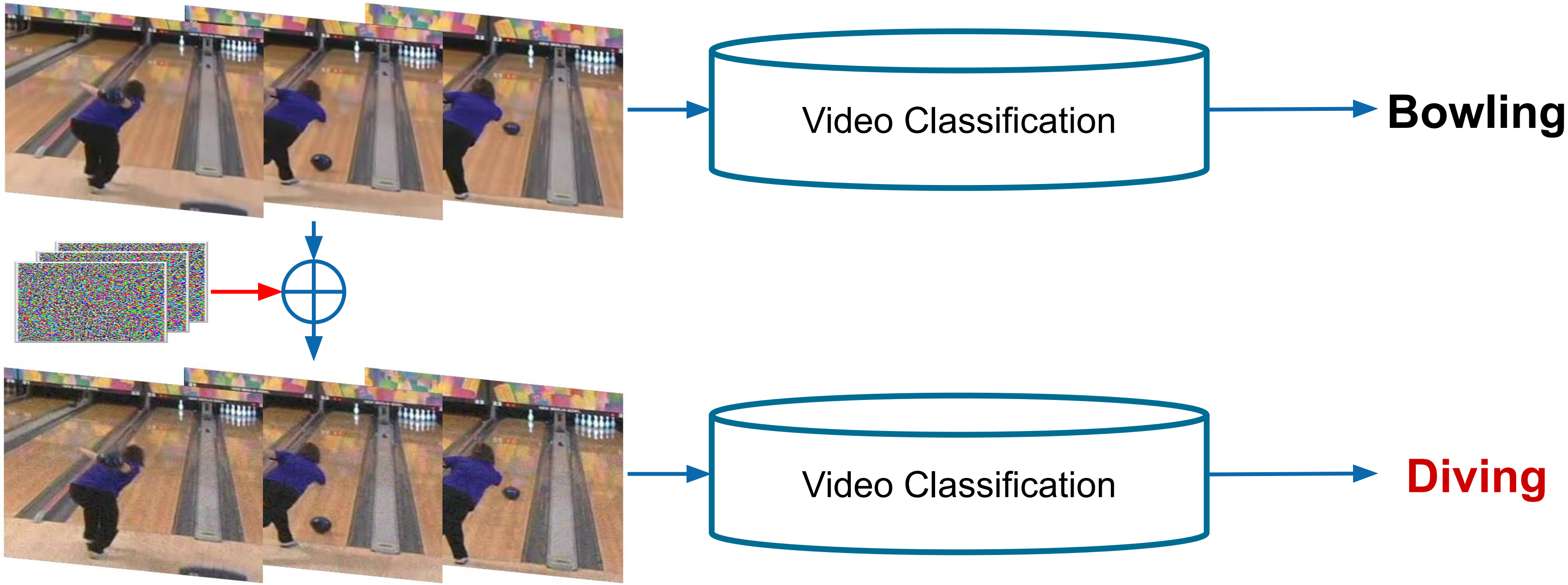
DL-based Video Compression

- Recently, **DL-based video compression** achieves impressive results by replacing all the components in the standard codecs with deep neural networks (DNNs).
  - It has been explored by the **Moving Picture Experts Group (MPEG)** for adoption in the next-generation video codecs.



Adversarial Attacks in DNNs

- Unfortunately, DNNs are known to be susceptible to **adversarial examples**.
  - Small perturbations added to the inputs of a DNN can cause it to misclassify the perturbed inputs.



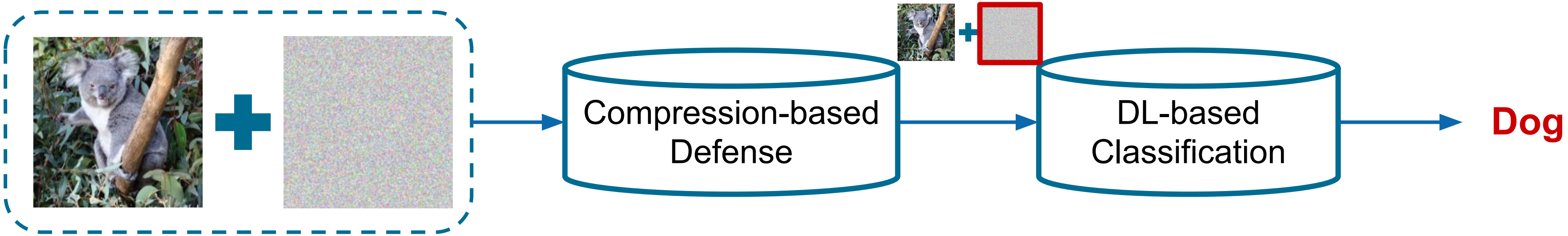
Motivation 1

- Compression techniques have been employed to remove the adversarial effect in several works[1-4].
- Video compression can **remove** the state-of-the-art video classification attacks.

[1] Jia, Xiaojun, et al. Comdefend: An efficient image compression model to defend adversarial examples. *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*. 2019.
[2] Zihao Liu, et al. Feature distillation: Dnn-oriented jpeg compression against adversarial examples. *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, 2019.
[3] Aaditya Prakash, et al. Protecting jpeg images against adversarial attacks. *Data Compression Conference*, 2018.
[4] Ayse Elvan Aydemir, Alptekin Temizel, and Tugba Taskaya Temizel. The effects of jpeg and jpeg2000 compression on attacks using adversarial examples. *CoRR, abs/1803.10418*, 2018

Motivation 1

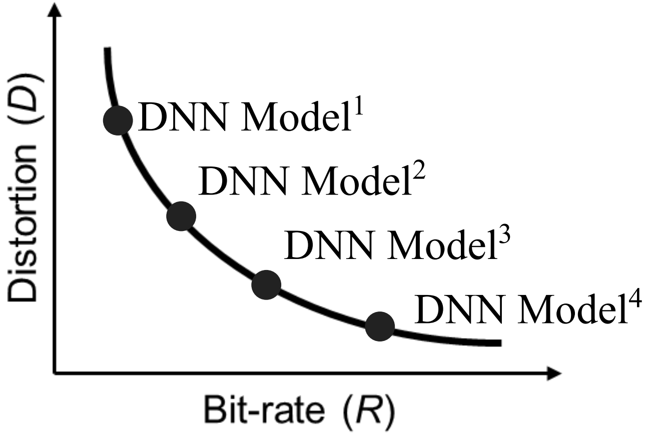
- Compression techniques have been employed to remove the adversarial effect in several works[1-4].
- Video compression can **remove** the state-of-the-art video classification attacks.
- ***Can a DL-based video compression be vulnerable to adversarial examples?***



[1] Jia, Xiaojun, et al. Comdefend: An efficient image compression model to defend adversarial examples. *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*. 2019.
[2] Zihao Liu, et al. Feature distillation: Dnn-oriented jpeg compression against adversarial examples. *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, 2019.
[3] Aaditya Prakash, et al. Protecting jpeg images against adversarial attacks. *Data Compression Conference*, 2018.
[4] Ayse Elvan Aydemir, Alptekin Temizel, and Tugba Taskaya Temizel. The effects of jpeg and jpeg2000 compression on attacks using adversarial examples. *CoRR, abs/1803.10418*, 2018

Motivation 2

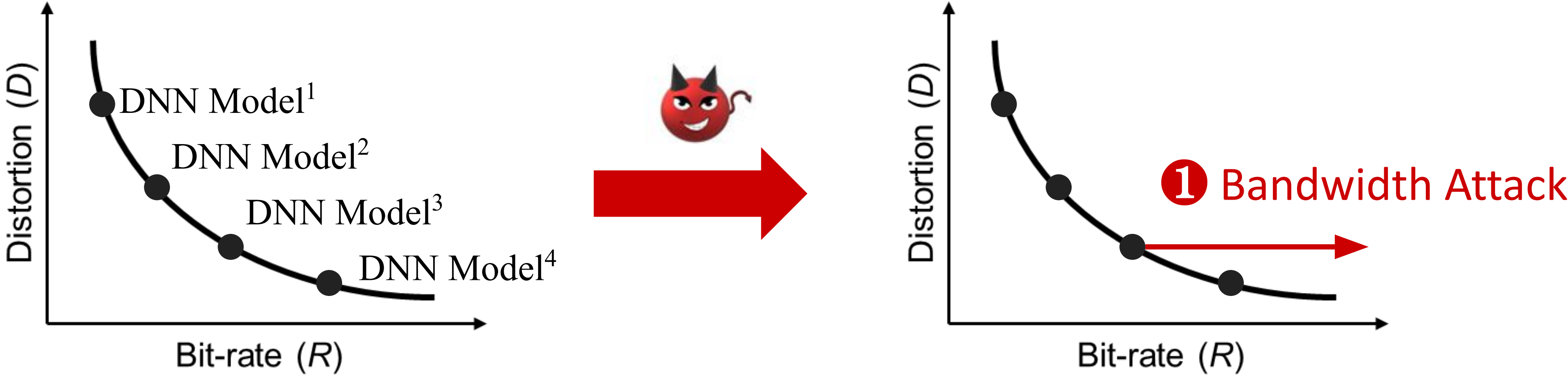
- DL-based video compression models[5-7] have **a fixed R-D relationship** through offline training.



[5] Guo Lu, et al. Dvc: An end-to-end deep video compression framework. *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, 2019.
[6] Ren Yang, et al. Learning for video compression with hierarchical quality and recurrent enhancement. *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, 2020.
[7] Zhihao Hu, et al. Fvc: A new framework towards deep video compression in feature space. *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, 2021.

Motivation 2

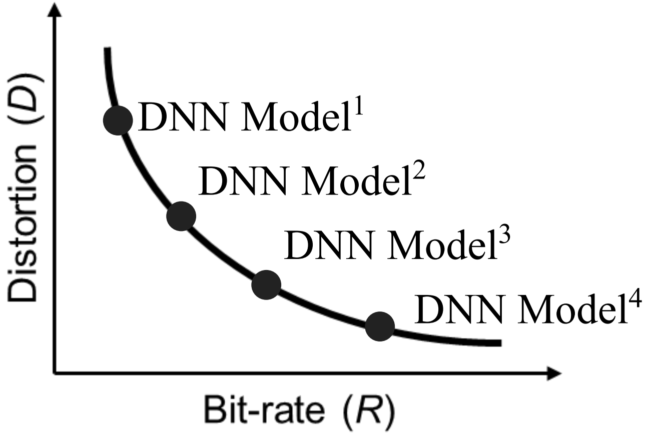
- DL-based video compression models[5-7] have **a fixed R-D relationship** through offline training.
- ***Can an adversary manipulate the R-D relationship arbitrarily?***

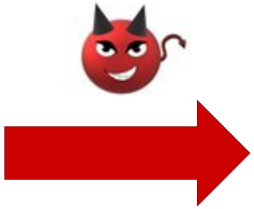


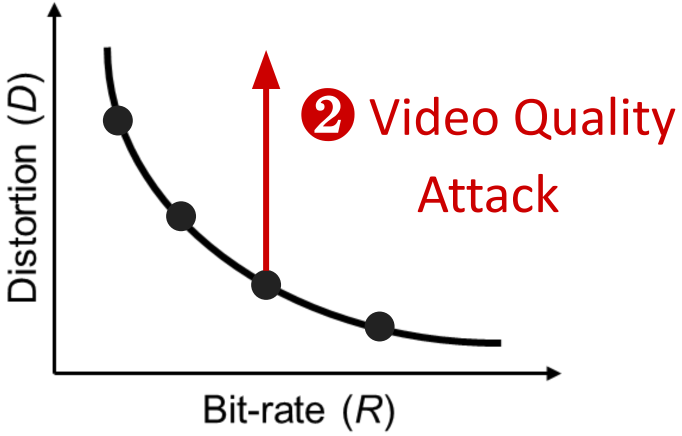
[5] Guo Lu, et al. Dvc: An end-to-end deep video compression framework. *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, 2019.
[6] Ren Yang, et al. Learning for video compression with hierarchical quality and recurrent enhancement. *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, 2020.
[7] Zhihao Hu, et al. Fvc: A new framework towards deep video compression in feature space. *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, 2021.

Motivation 2

- DL-based video compression models[5-7] have **a fixed R-D relationship** through offline training.
- ***Can an adversary manipulate the R-D relationship arbitrarily?***







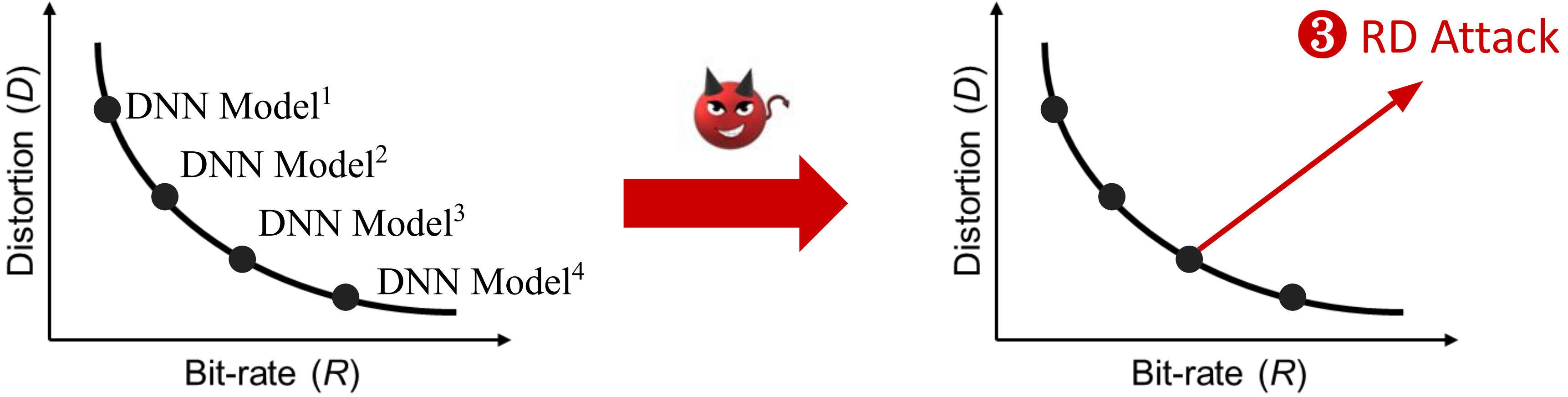
[5] Guo Lu, et al. Dvc: An end-to-end deep video compression framework. *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, 2019.

[6] Ren Yang, et al. Learning for video compression with hierarchical quality and recurrent enhancement. *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, 2020.

[7] Zhihao Hu, et al. Fvc: A new framework towards deep video compression in feature space. *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, 2021.

Motivation 2

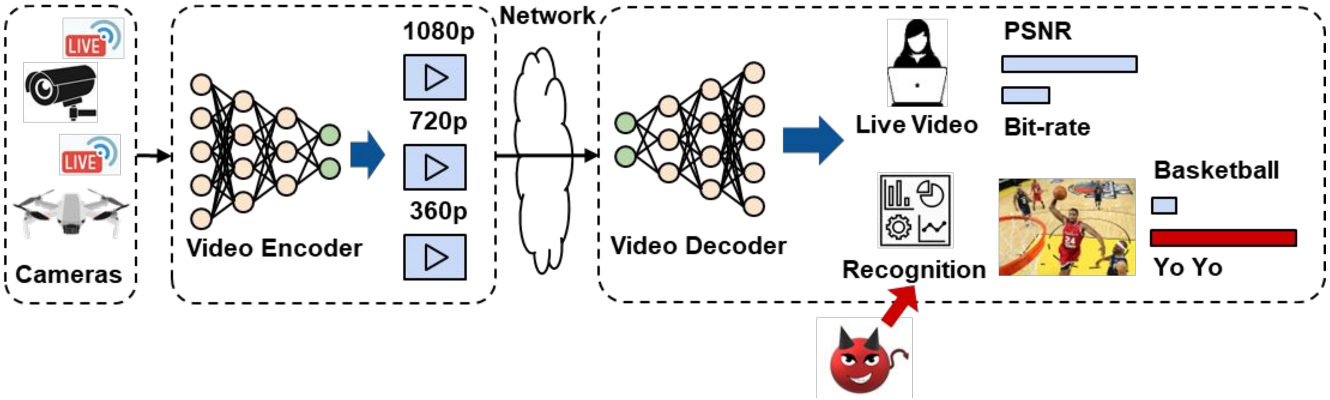
- DL-based video compression models[5-7] have **a fixed R-D relationship** through offline training.
- ***Can an adversary manipulate the R-D relationship arbitrarily?***



[5] Guo Lu, et al. Dvc: An end-to-end deep video compression framework. *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, 2019.
[6] Ren Yang, et al. Learning for video compression with hierarchical quality and recurrent enhancement. *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, 2020.
[7] Zhihao Hu, et al. Fvc: A new framework towards deep video compression in feature space. *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, 2021.

Motivation 3

- The state-of-the-art works on video classification attacks[8-9] didn’t consider video compression in their threat model.



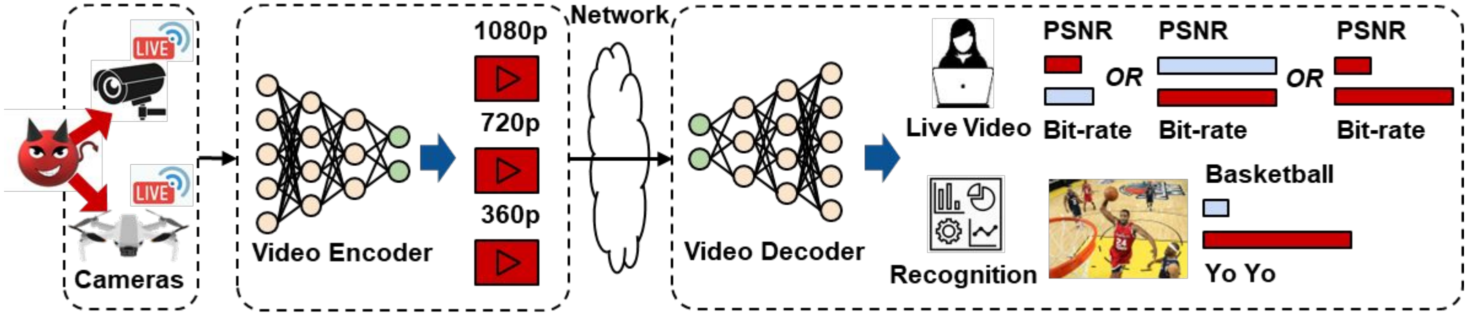
<Previous attack pipeline>

[8] Shasha Li, et al. Stealthy adversarial perturbations against real-time video classification systems. In Proceedings 2019 Network and Distributed System Security Symposium (NDSS), 2019.

[9] Shangyu Xie, et al. Universal 3-dimensional perturbations for black-box attacks on video recognition systems. In 2022 IEEE Symposium on Security and Privacy (SP), 2022.

Motivation 3

- The state-of-the-art works on video classification attacks[8-9] didn't consider video compression in their threat model.
- ***Can an adversary target towards front-end video sources and also affect a downstream video recognition system?***

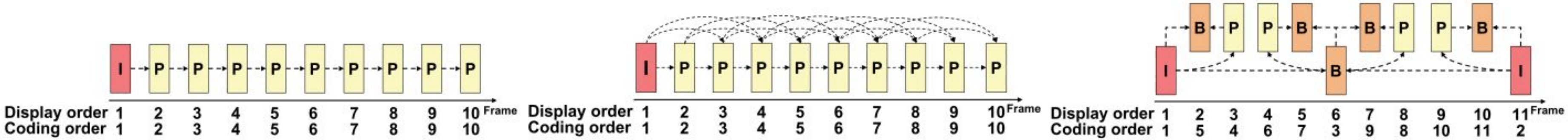


<Our proposed attack pipeline>

[8] Shasha Li, et al. Stealthy adversarial perturbations against real-time video classification systems. In Proceedings 2019 Network and Distributed System Security Symposium (NDSS), 2019.
[9] Shangyu Xie, et al. Universal 3-dimensional perturbations for black-box attacks on video recognition systems. In 2022 IEEE Symposium on Security and Privacy (SP), 2022.

Motivation 4

- Video compression group a series of frames into sequences called **Group of Pictures (GOP)**[5-7] to allow back-end users to access video streams at any time.
  - Three types of GOP structures are used in DNN-based video compression systems.

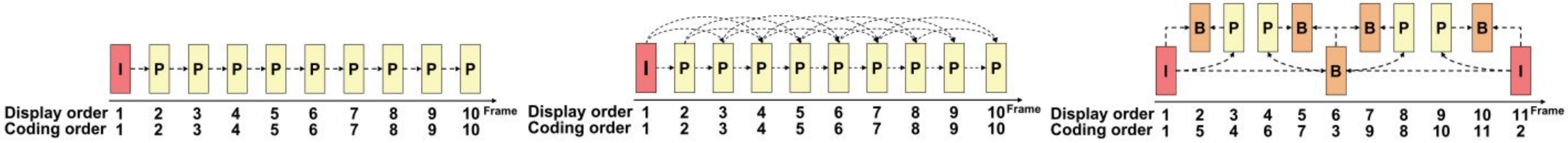


(a) Non-hierarchical [5] (b) Hierarchical-P [6] (c) Hierarchical-B [7]

[5] Guo Lu, et al. Dvc: An end-to-end deep video compression framework. *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, 2019.
[6] Ren Yang, et al. Learning for video compression with hierarchical quality and recurrent enhancement. *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, 2020.
[7] Zhihao Hu, et al. Fvc: A new framework towards deep video compression in feature space. *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, 2021.

Motivation 4

- Video compression group a series of frames into sequences called **Group of Pictures (GOP)**[5-7] to allow back-end users to access video streams at any time.
  - Three types of GOP structures are used in DNN-based video compression systems.
- ***Can well-crafted perturbations break down temporal coding structures?***

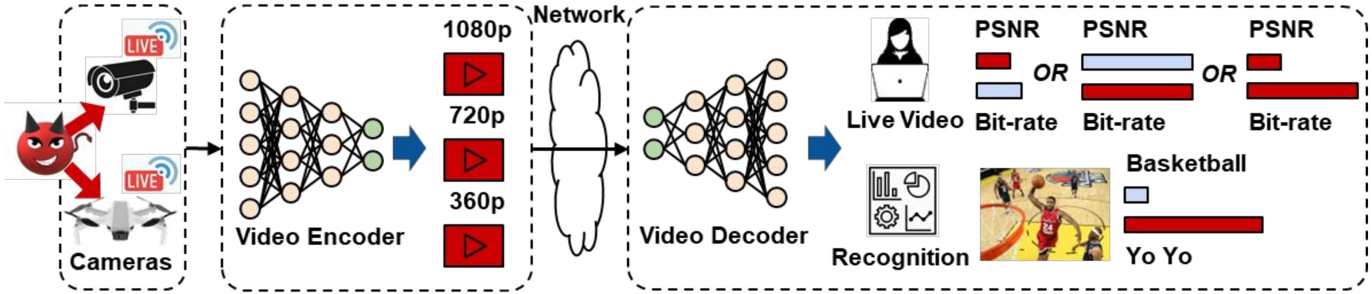


(a) Non-hierarchical [5] (b) Hierarchical-P [6] (c) Hierarchical-B [7]

[5] Guo Lu, et al. Dvc: An end-to-end deep video compression framework. *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, 2019.
[6] Ren Yang, et al. Learning for video compression with hierarchical quality and recurrent enhancement. *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, 2020.
[7] Zhihao Hu, et al. Fvc: A new framework towards deep video compression in feature space. *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, 2021.

Contributions

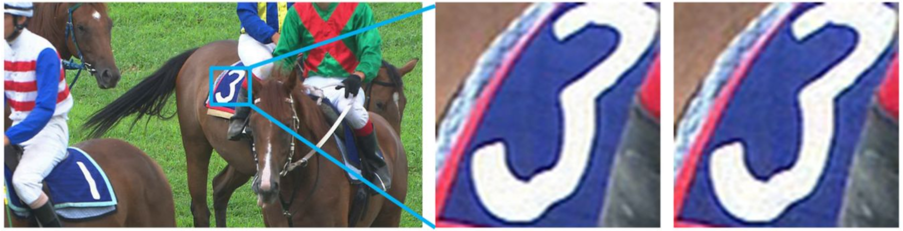
- Perform the **first** systematic study of adversarial attacks on DL-based video compression and downstream video recognition systems.
- Propose **four** new adversarial attacks, dubbed RoVISQ, that result in high-impact security and QoE consequences.
- Construct a well-designed **universal perturbation** that is invariant to the underlying DNN model, encoding parameters, and input videos.
- Show the **resiliency** of RoVISQ attacks against various defenses.



Threat Model

- Attack Scenarios
  - Adversary adds small perturbations to a stored video to subvert the video compression over **a long period of time**.

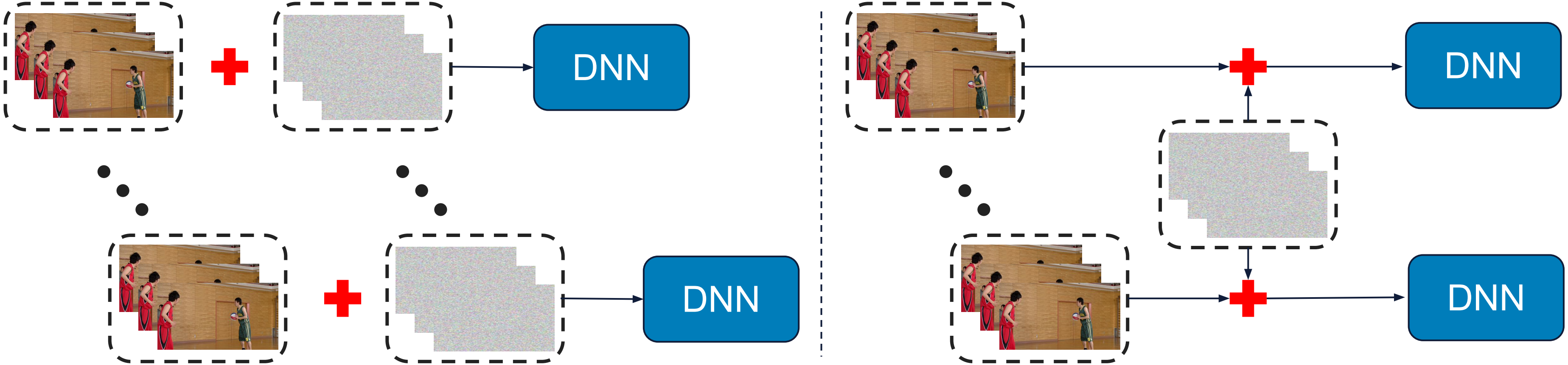
Raw Input Perturbed Input



Raw Input Perturbed Input

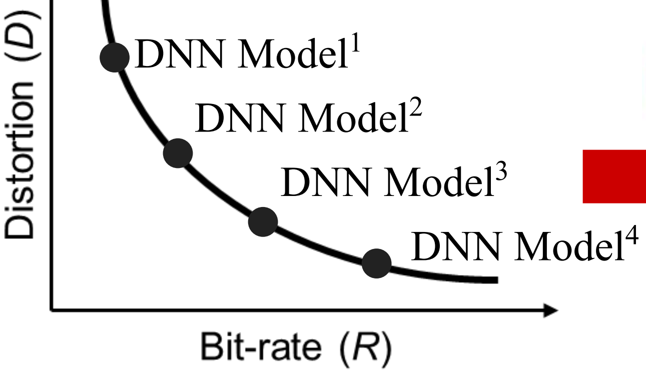
Threat Model

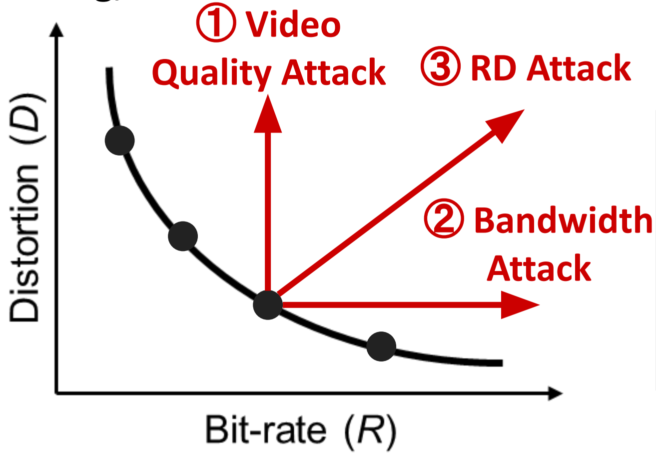
- Attack Scenarios
  - There are two attack scenarios.
    - **Offline Attack**: sample-wise perturbations that are independently added to each sample.
    - **Online Attack**: well- crafted universal perturbations that can be used to attack any given video sequence at any time step.



Threat Model

- Adversary’s Goal
  - Selectively degrade the bit-rate R and/or distortion level D compared to the R-D relationship from the pre-trained model.
    - **Video Quality Attack** -> **Low quality**
    - **Bandwidth Attack** -> **Buffering, Low-Resolution Video**
    - **RD Attack** -> **Low quality, Buffering, Low-Resolution Video**
    - **Video Classification Attack**





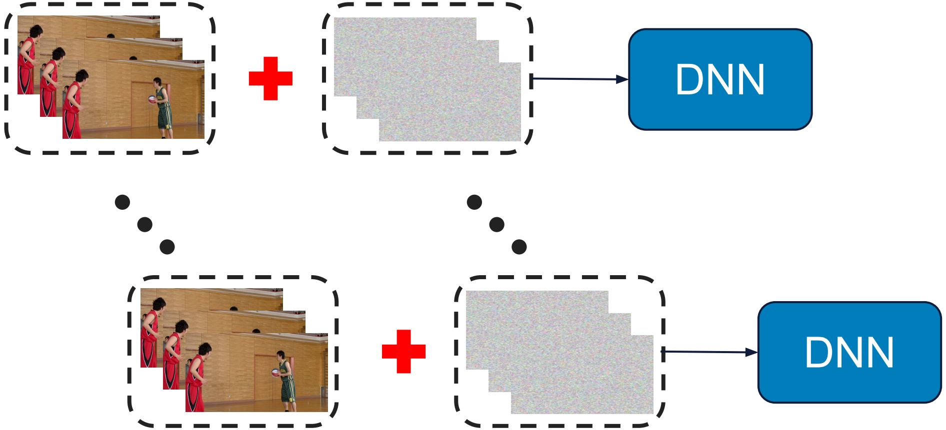
Video Analytics

Threat Model

- Adversary's Capability and knowledge
  - **Offline Scenario**

    *** Compression rate, GOP structure**

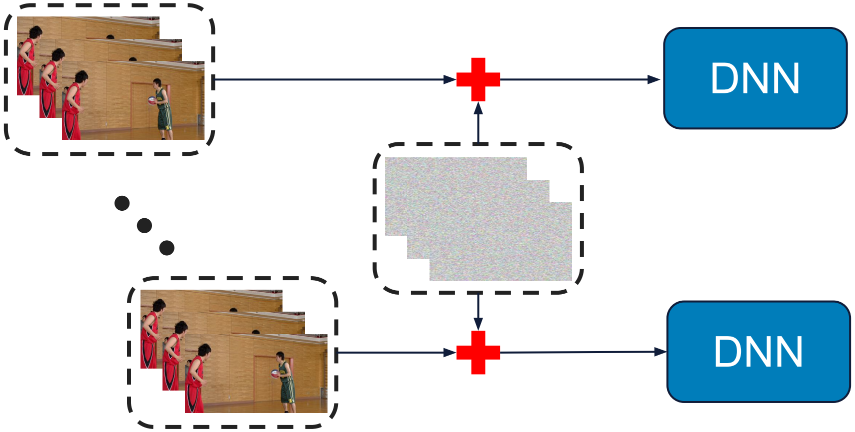
    - We assume that the adversary knows every **encoding parameters**.
    - We assume the attacker has **white-box** access to an open-source model.
    - Our perturbations are independently added to each sample because the attack latency is no constrained.



Threat Model

- Adversary's Capability and knowledge
  - **Online Scenario**

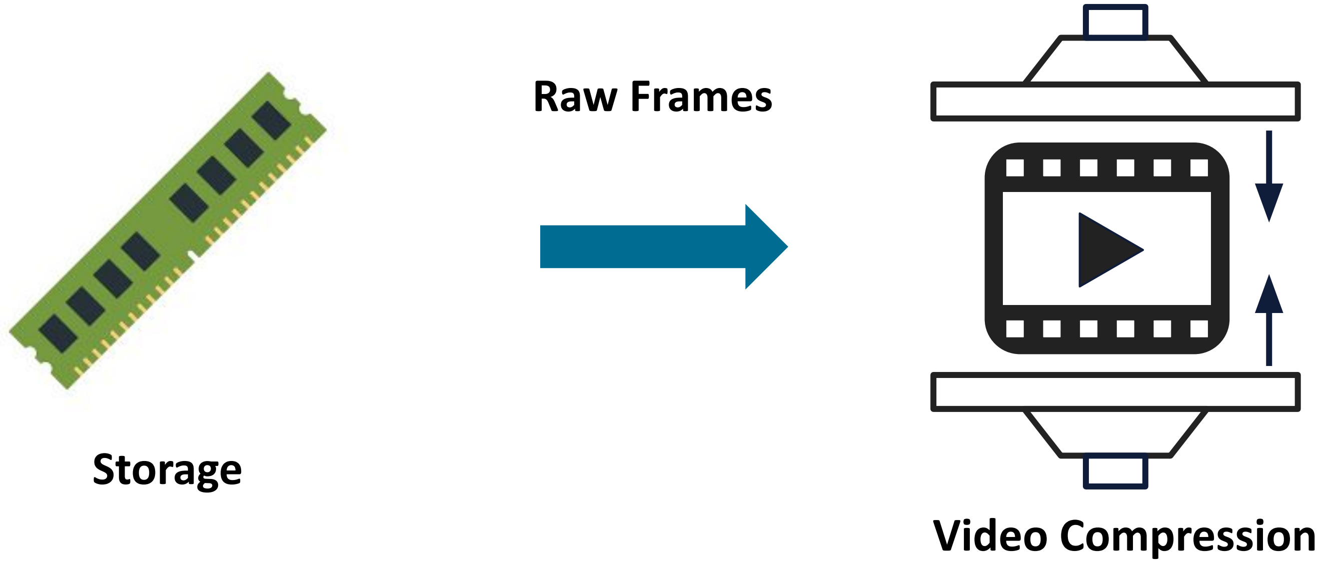
    * Compression rate, GOP structure
    - We assume that the adversary doesn't know any **encoding parameters**.
    - We study both **white-box** and **black-box** settings for DNN models.
    - Attacker is capable of injecting perturbations onto the real-time video stream.



Our Offline Attack Construction

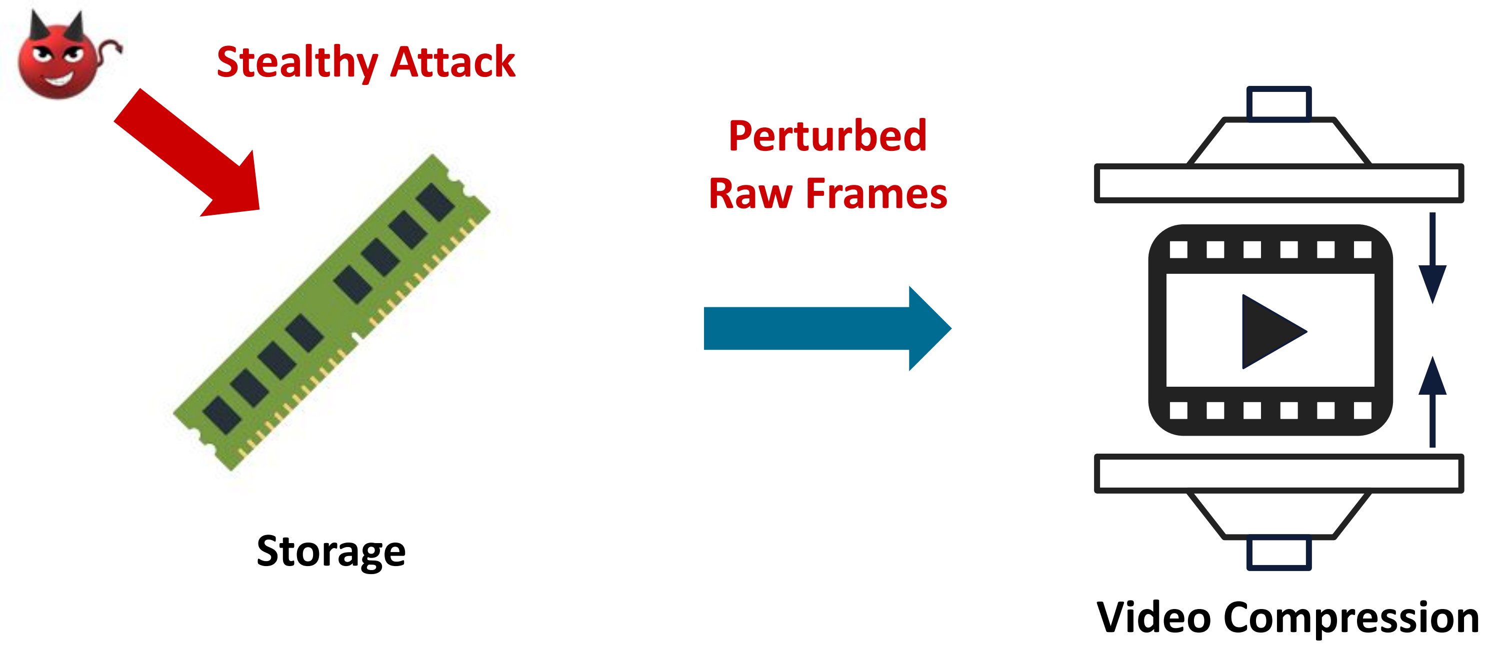
Offline Attack Construction

- In offline scenario, the raw frames are stored in the storage device.



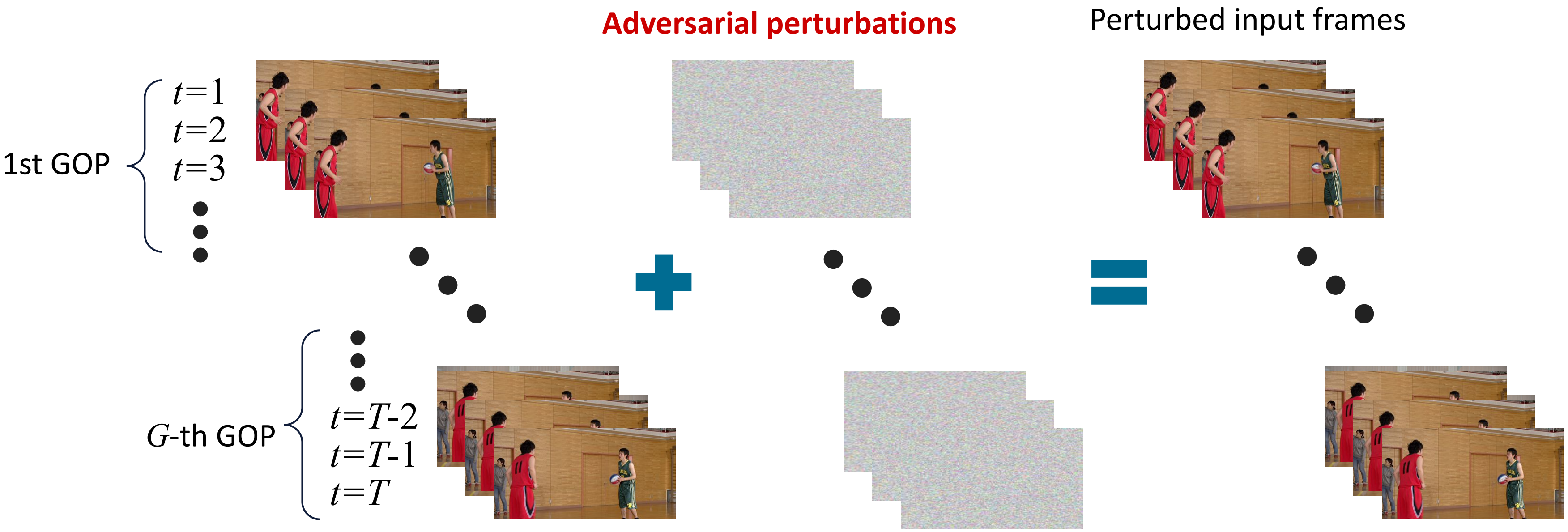
Offline Attack Construction

- Our adversary adds the small perturbations to the input frames stored in the storage.



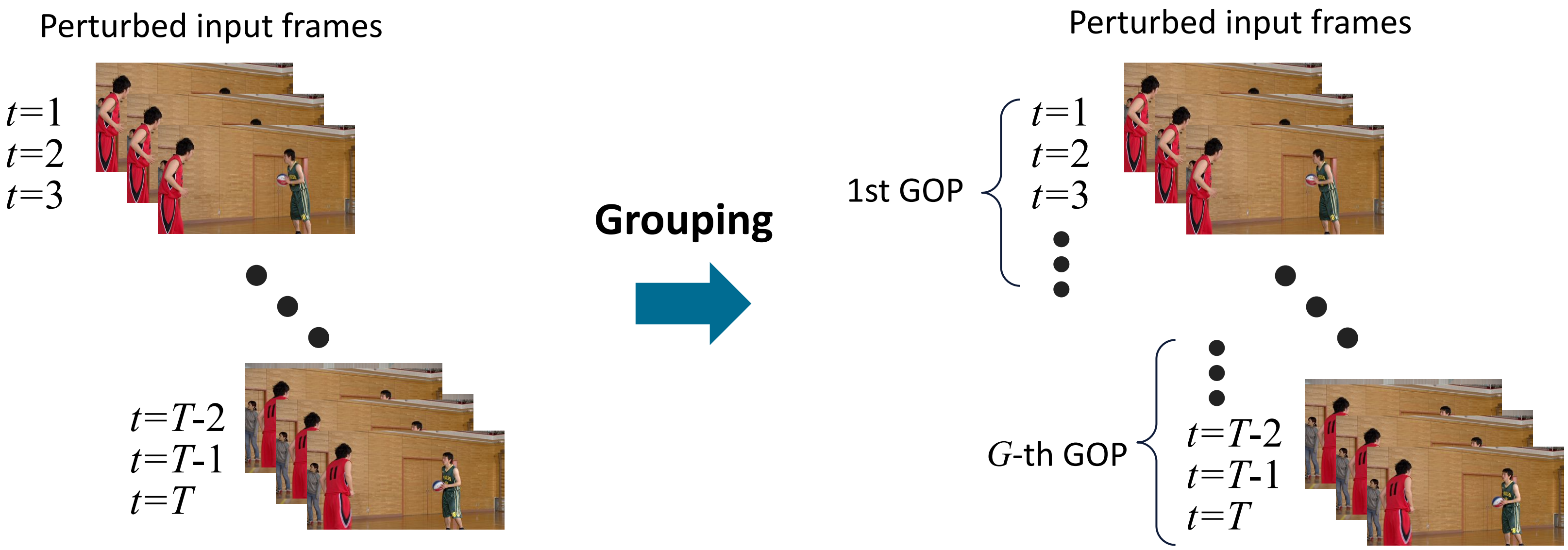
Offline Attack Construction

- For example,



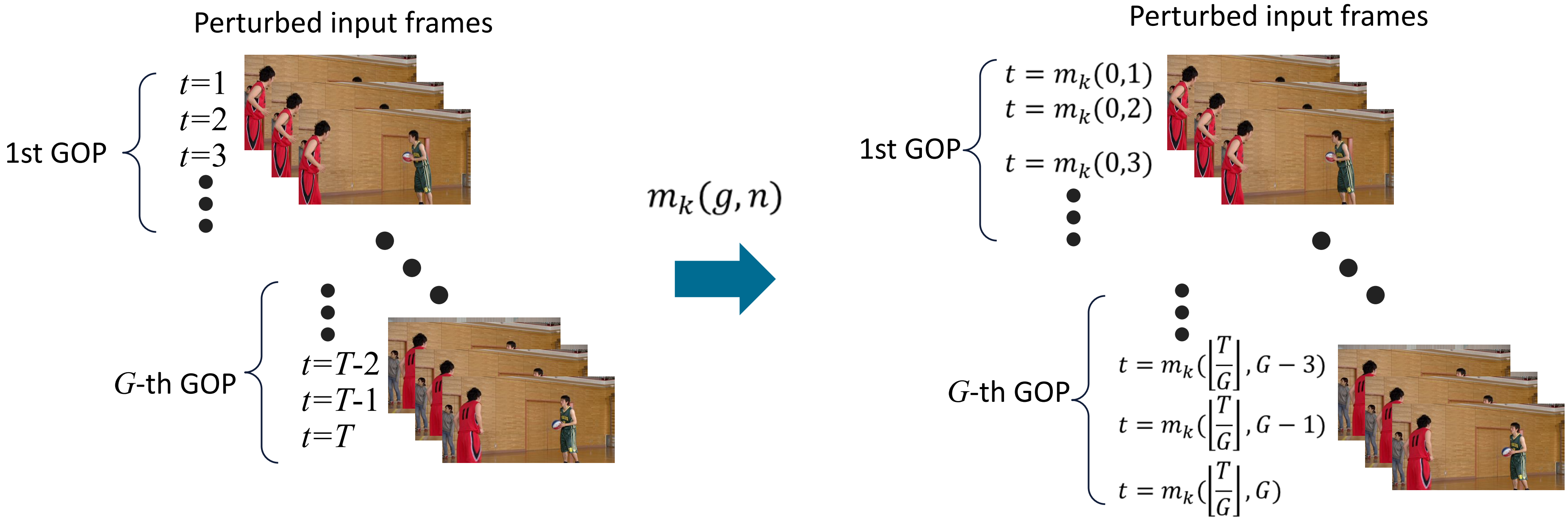
Offline Attack Construction

- Video Compression groups **a series of input frames** into **GOP**.



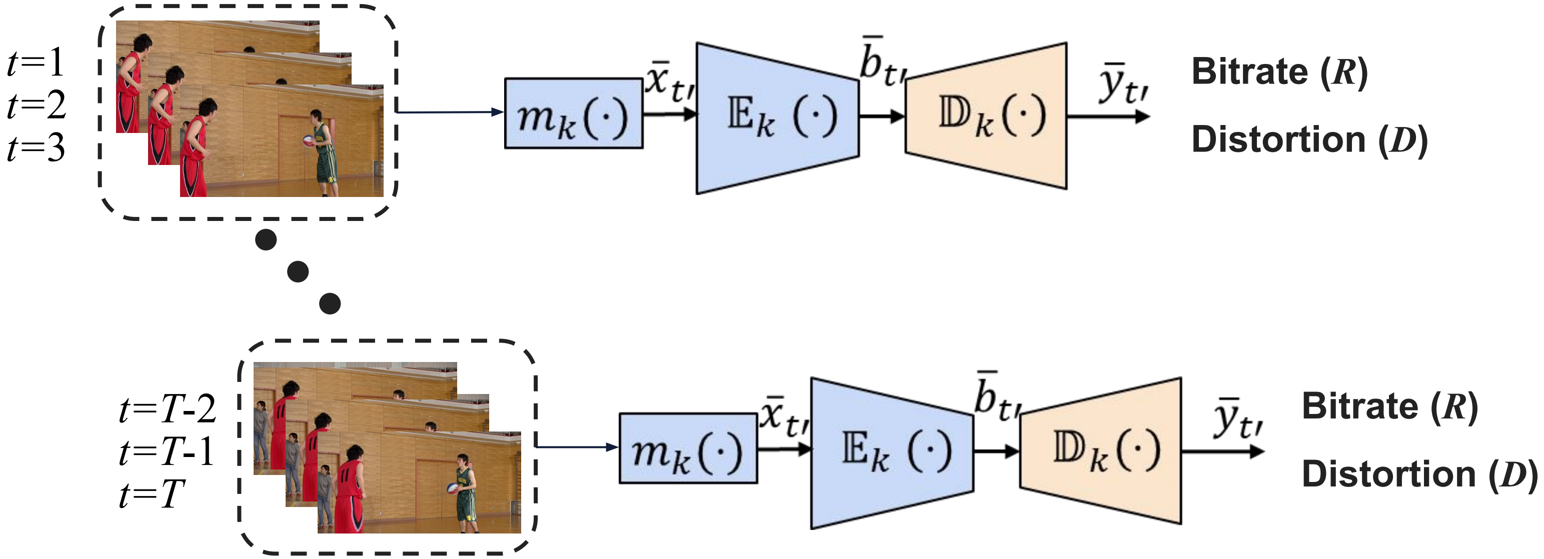
Offline Attack Construction

- For a given k, the n-th coding order in the g-th GOP is mapped to a new time step t using a deterministic function $m_k(g, n)$



Offline Attack Construction

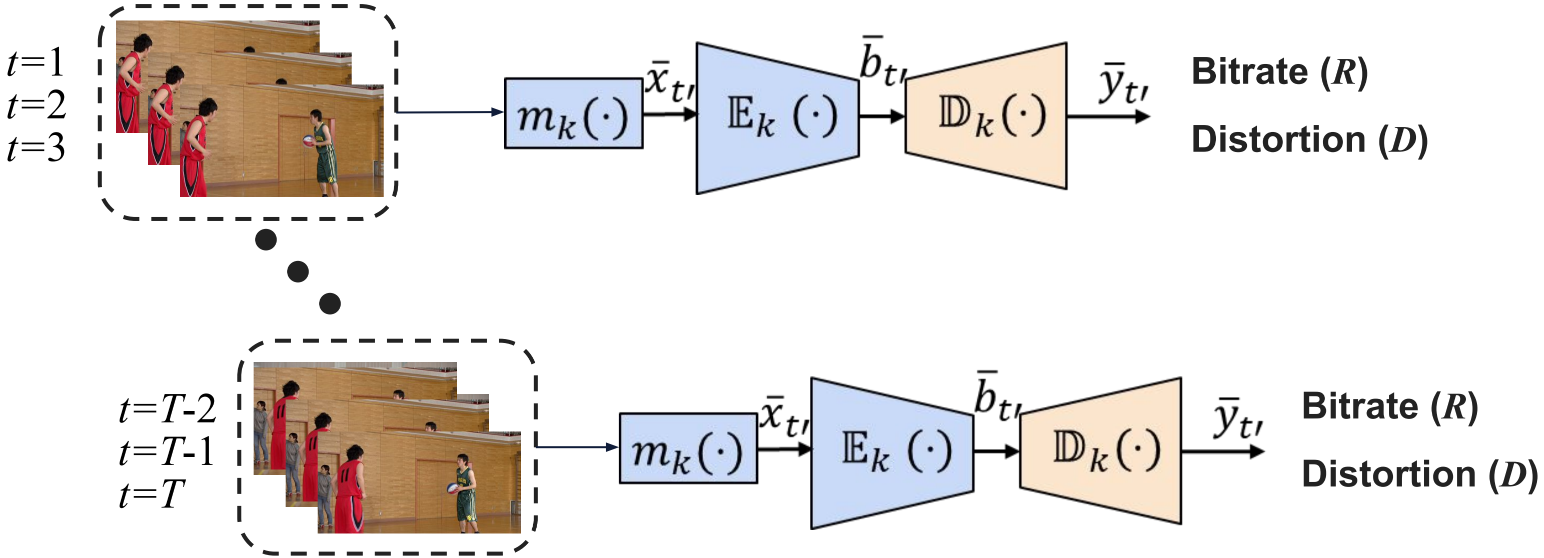
- We quantify the video compression performance based on two important measures.
  - **Bit-rate**
  - **Distortion (mean squared error)**



Offline Attack Construction

- We formulate the QoE factors for the g-th GOP from the bit-rate and the distortion:

$$Q_0(\bar{B}_g) = \frac{1}{G} \sum_{\bar{b}_t \in \bar{B}_g} R(\bar{b}_t) \qquad Q_1(X_g, \bar{Y}_g) = \frac{1}{G} \sum_{\bar{y}_t \in \bar{Y}_g} D(x_t, \bar{y}_t)$$

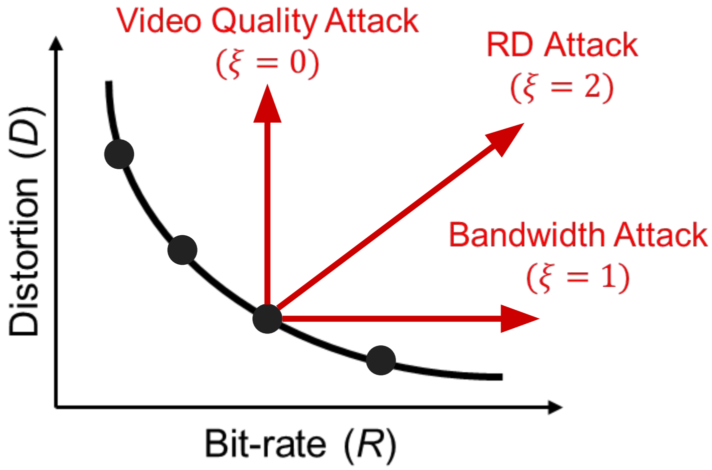


Offline Attack Construction

- To generate the perturbations, the adversary maximizes the following loss function.

$$\max_{\Delta_g} \quad \mathcal{L}_{comp}(g) \quad \text{s.t.} \quad \|\Delta_g\|_\infty \leq \epsilon_c$$

$$\mathcal{L}_{comp}(g) = \begin{cases} \mathbf{E}_0 + \lambda \cdot Q_1(X_g, \bar{Y}_g) & \text{if } \xi = 0 \\ Q_0(\bar{\mathcal{B}}_g) + \lambda \cdot \mathbf{E}_1 & \text{if } \xi = 1 \\ Q_0(\bar{\mathcal{B}}_g) + \lambda \cdot Q_1(X_g, \bar{Y}_g) & \text{if } \xi = 2 \end{cases}$$



ξ determines the attack type.

ϵ_c is the upper bound of the L-infinity norm of the perturbation.

λ determines the target video compression model by controlling R-D trade-off.

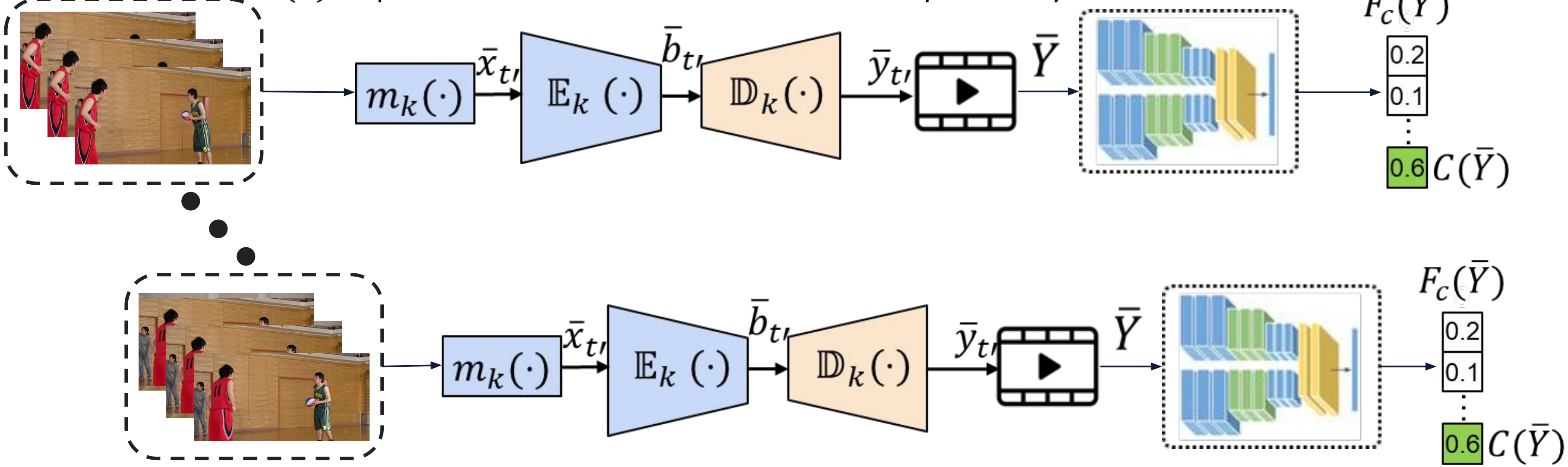
Offline Attack Construction

- Adversarial Loss for Downstream **Video Classification**

$$\mathcal{L}_{adv} = \begin{cases} F_{C(Y)}(\bar{Y}) - \max\limits_{c \neq C(Y)} F_c(\bar{Y}) & \text{(Untargeted)} \\ \max\limits_{c \neq c^*} F_c(\bar{Y}) - F_{c^*}(\bar{Y}) & \text{(Targeted)} \end{cases}$$

$F_c(\bar{Y})$ indicates the probability of the video belonging to a specific class C.

$C(\bar{Y})$ maps a video to the class with the maximum probability.

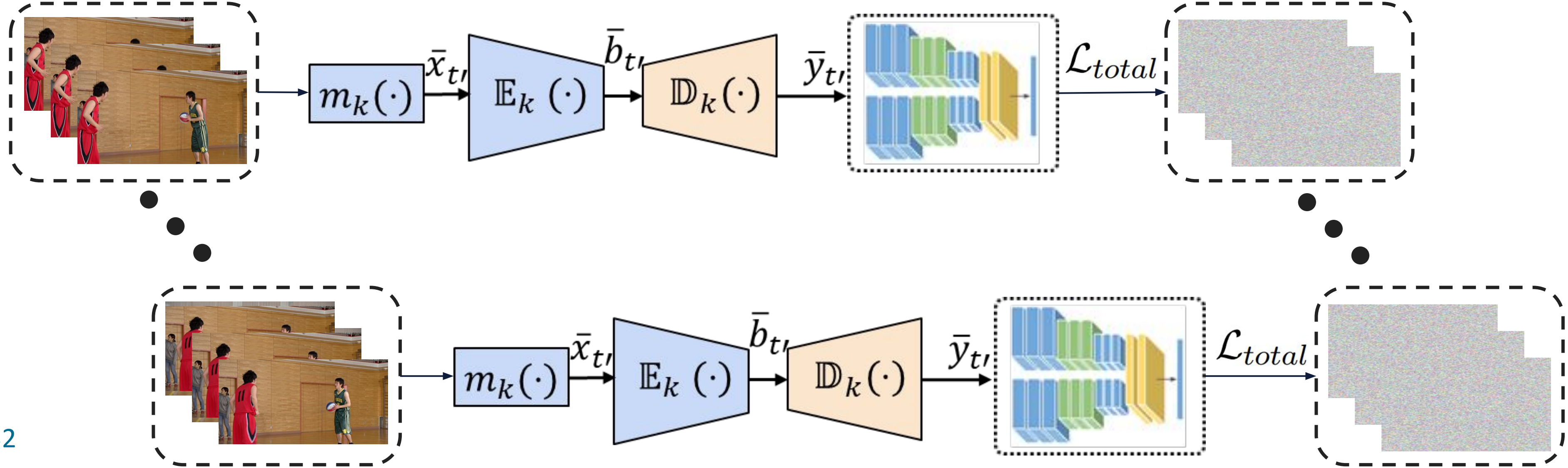


Offline Attack Construction

- Finally, we integrate all the loss functions to simultaneously derive perturbations on video compression and classification.

$$\max_{\Delta} \mathcal{L}_{total} = \frac{1}{\lfloor T/G \rfloor + 1} \sum_{g=0}^{\lfloor T/G \rfloor} \mathcal{L}_{comp}(g) - \beta \cdot \mathcal{L}_{adv}$$

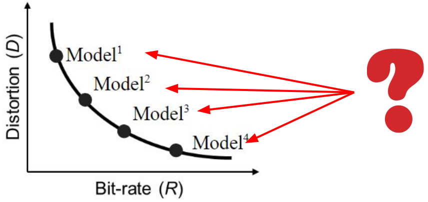
where β adjusts the scale of the two loss functions.



Our Online Attack Construction

Challenges of Online Attack

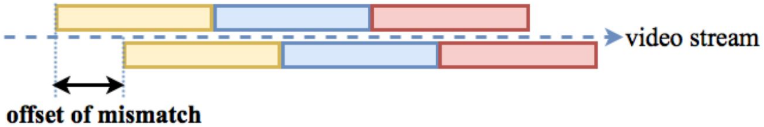
- Online adversarial attack is particularly challenging.
  - What is the compression rate of video compression?



  - Which mapping function $m_k(\cdot)$ does victim video compression use?

    Mapping function depends on the GOP structures.

  - How to align the perturbations with the target video sequence?

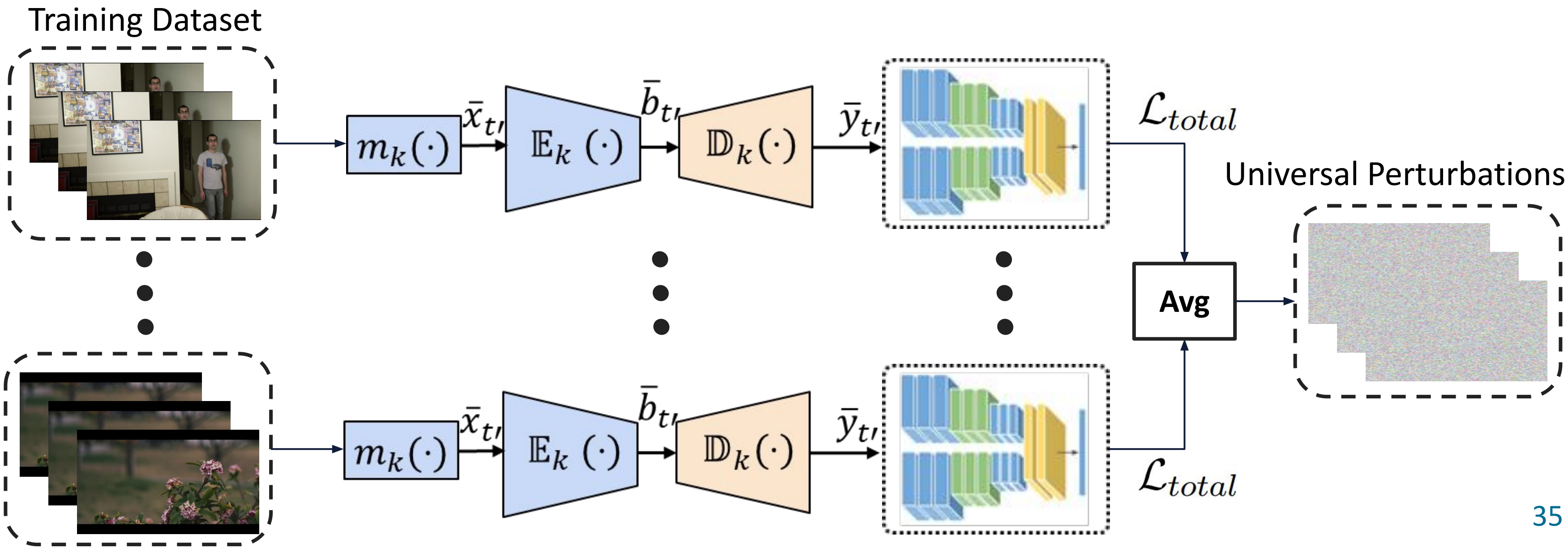


  - Contents of the video sequences are unknown.

    Each content has a different distribution of video data.

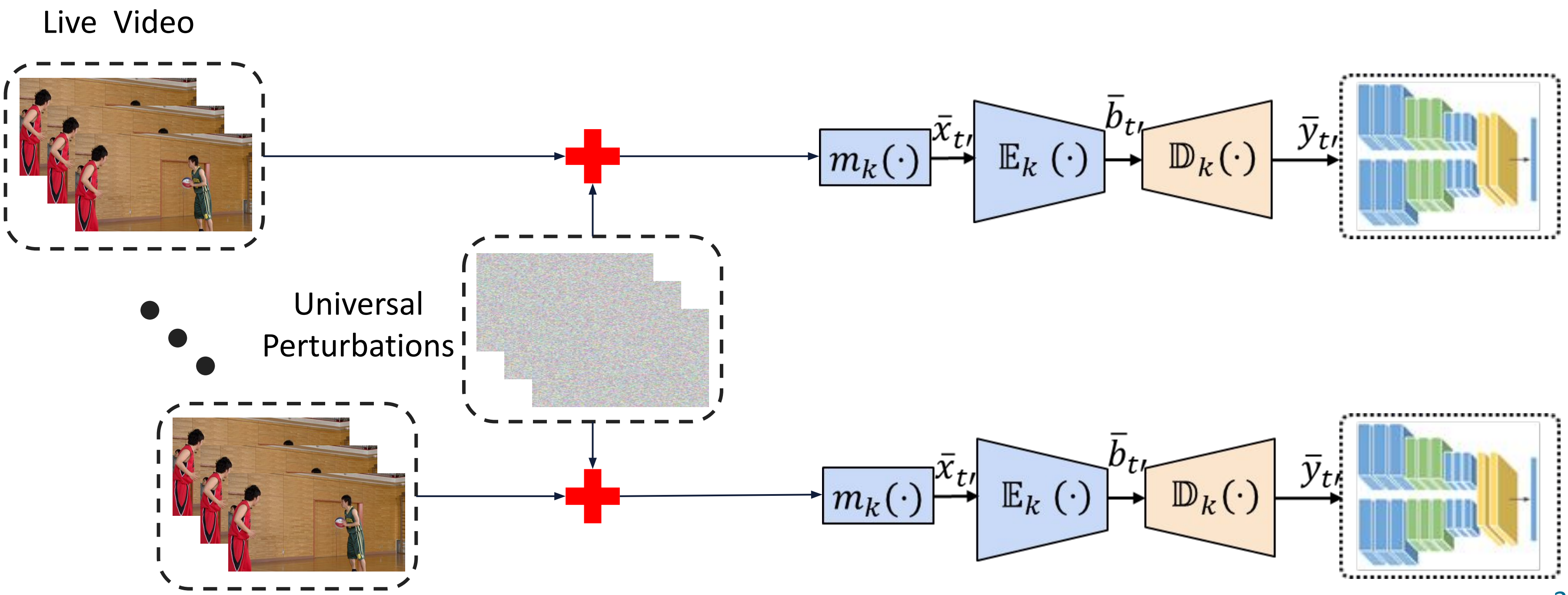
Online Attack Construction

- We train our universal perturbations that are **agnostic** to ❶compression ratio, ❷GOP structure, and ❸input, which is suitable for online attack.
  - We average the loss values across all training videos available to the attacker.



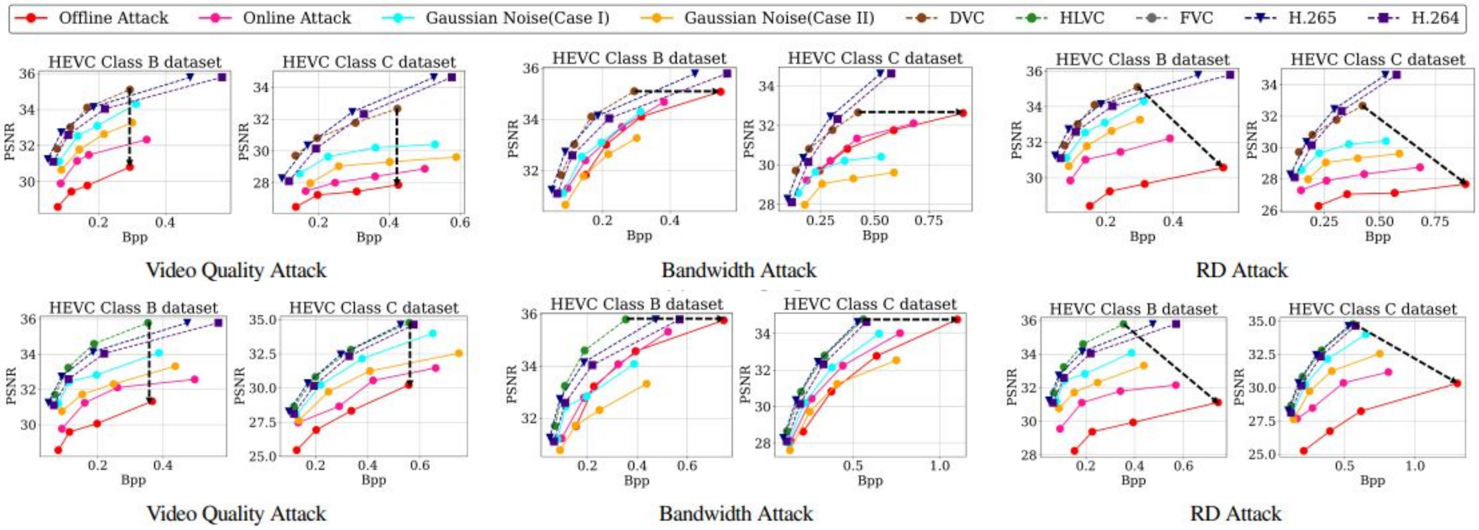
Online Attack Construction

- Real-time Adversarial Attacks on Entire Systems



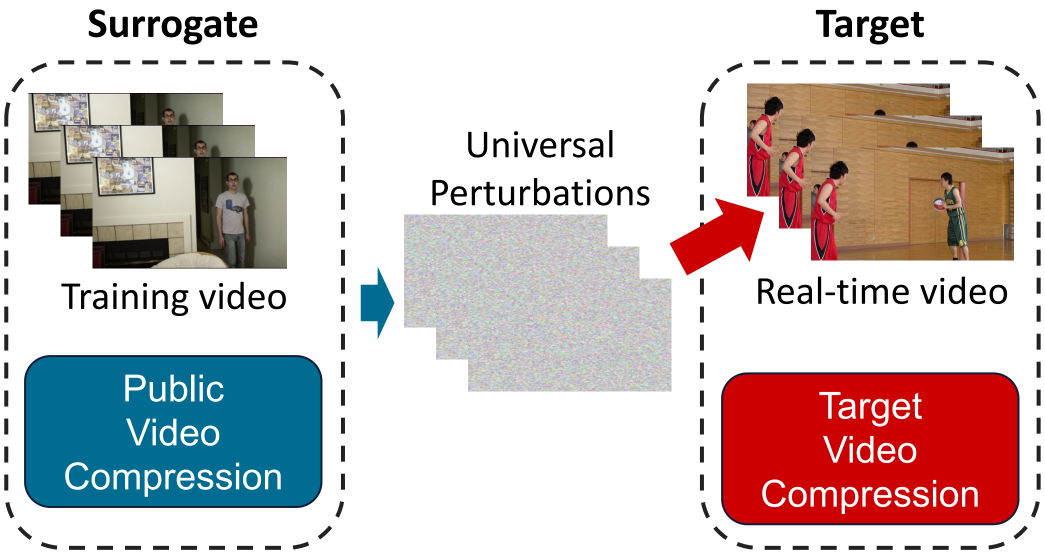
Experimental Results

- Evaluation Setup
  - Baselines
    - Gaussian (Case I) : $\sigma_I = \sigma_P = \sigma_B = \epsilon_c$, Gaussian (Case II) : $\sigma_I = 2 \cdot \epsilon_c, \sigma_P = \sigma_B = \epsilon_c$
- White-box Attack Performance



Experimental Results

- Black-box Attack Performance



<Attack performance against conventional codecs>

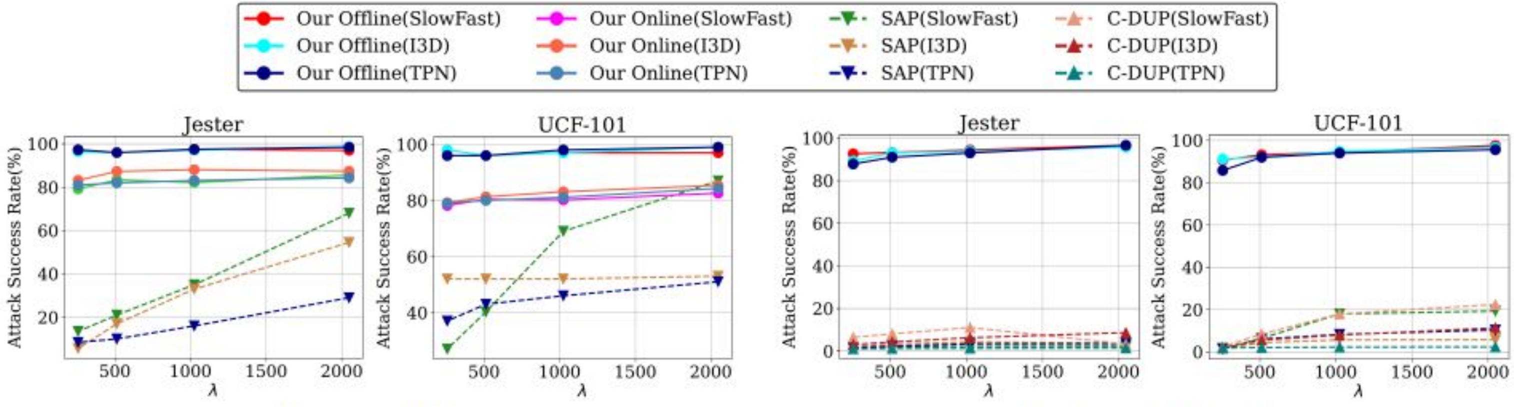
		Video Quality Attack	Bandwith Attack	RD Attack	Gaussian Noise
PSNR (dB)	H.265	-3.47	-1.55	-3.62	-1.71
	H.264	-3.19	-1.03	-3.48	-1.31
Bpp	H.265	+45.5%	+78.4%	+73.8%	+62.1%
	H.264	+34.7%	+65.2%	+61.8%	+45.9%

<Attack performance against unseen DNN models>

		Video Quality Attack	Bandwith Attack	RD Attack	Gaussian Noise
M1	PSNR (dB)	-2.37	-0.87	-2.46	-1.57
	Bpp	+18.4%	+32.5%	+29.7%	+17.3%
M2	PSNR (dB)	-2.31	-0.92	-2.48	-1.44
	Bpp	+19.1%	+30.4%	+27.7%	+17.8%
M3	PSNR (dB)	-2.44	-0.91	-2.55	-1.68
	Bpp	+19.5%	+31.7%	+31.1%	+14.8%
M4	PSNR (dB)	-2.47	-0.95	-2.51	-1.63
	Bpp	+18.6%	+29.4%	+30.2%	+15.2%
M5	PSNR (dB)	-2.49	-0.88	-2.53	-1.72
	Bpp	+17.6%	+32.8%	+30.6%	+17.4%
M6	PSNR (dB)	-2.38	-0.98	-2.36	-1.65
	Bpp	+18.3%	+31.4%	+32.1%	+17.8%

Experimental Results

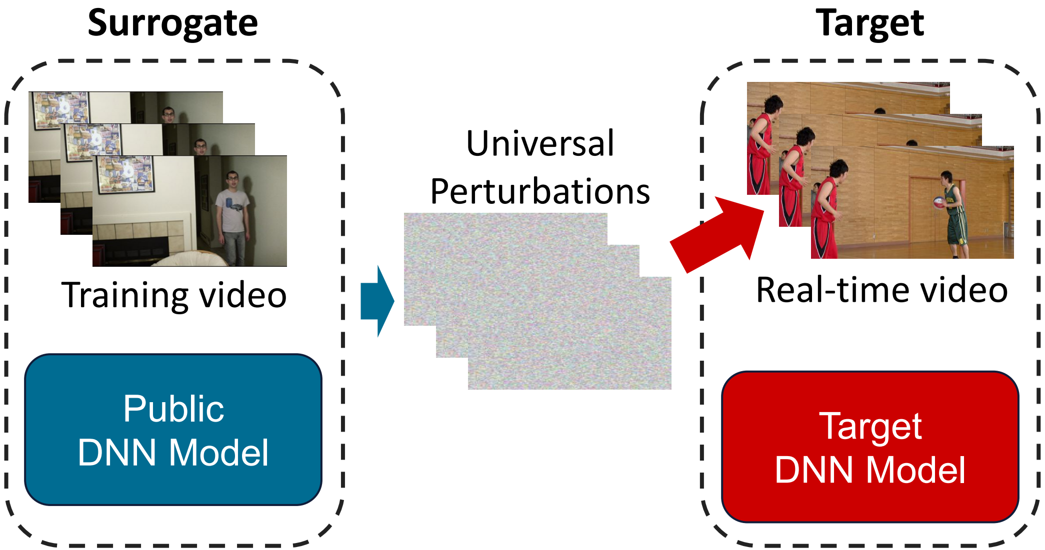
- White-box Attacks on Video Classification
  - We evaluate the success rate when directed towards a downstream video classifier and provide comparisons with state-of-the-art attacks on video classification.
  - As seen, our attack consistently achieves the highest success rate.
  - In particular, we obtain over 90% success rate on the UCF-101 and Jester datasets.



(a) Untargeted Attack (b) Targeted Attack

Experimental Results

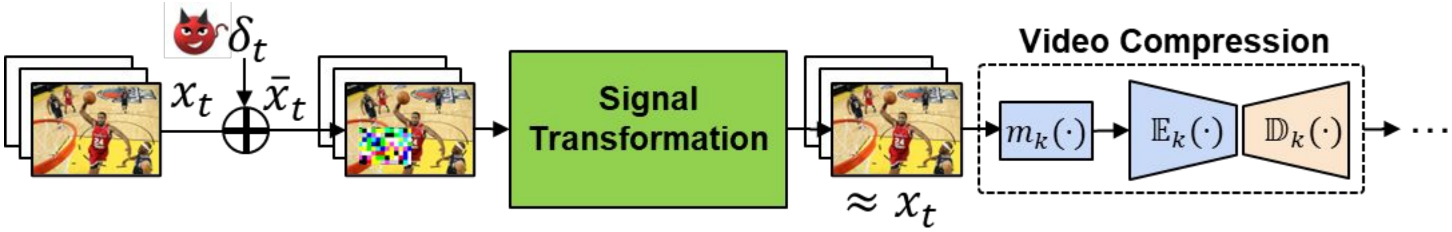
- Black-box Attacks on Video Classification
    - The proposed adversarial perturbations are transferable to unseen video classification models, outperforming previous attacks.



Victim Model	Attack	Attack Success Rate (%) $\lambda=256$	512	1024	2048
TPN [73]	GeoTrap [36]	6.4	16.8	18.5	32.4
	U3D [71]	7.4	17.5	19.4	36.1
	Bandwidth (I3D)	71.3	76.9	79.6	**82.4**
	Bandwidth (SlowFast)	**73.2**	**77.8**	**80.6**	81.5
SlowFast [21]	GeoTrap [36]	11.2	22.2	38.9	54.6
	U3D [71]	10.2	24.1	37.0	60.2
	Bandwidth (I3D)	73.2	**76.9**	78.7	81.5
	Bandwidth (TPN)	**74.1**	75.0	**80.6**	**82.4**
I3D [13]	GeoTrap [36]	8.3	24.1	41.7	42.6
	U3D [71]	6.5	16.7	39.8	48.1
	Bandwidth (SlowFast)	70.4	**76.9**	**81.5**	**83.3**
	Bandwidth (TPN)	**72.2**	74.1	76.9	80.6

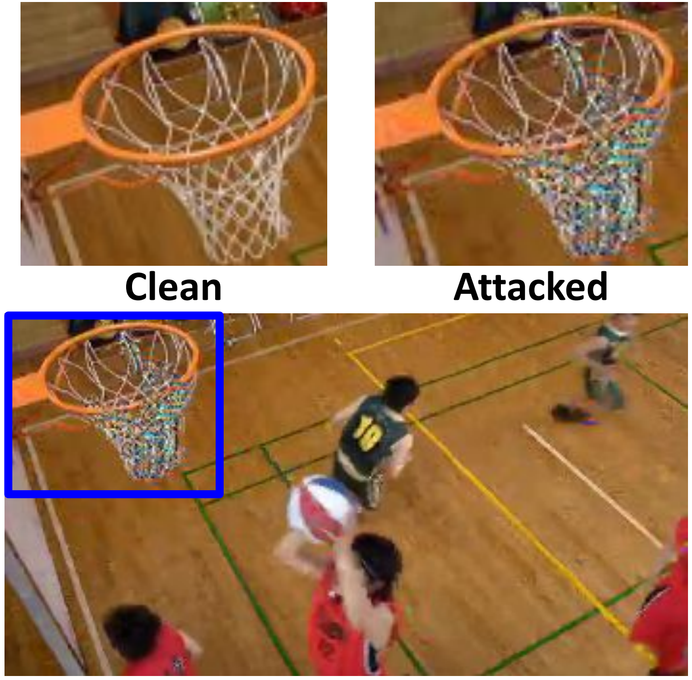
Evaluation of Existing Defenses

- Defense Construction
  - We comprehensively evaluate different defense mechanisms against our attacks. There are very few defenses available for adversarial video classification.
  - We implement new defense mechanisms that rely on signal transformations to remove adversarial perturbations
    - **Adversarial Training**
    - **Video Denoising**
    - **JPEG Image Compression**



Experimental Results

- Attack Visualization



Attacked Video

Conclusion

- We presents the first systematic study on adversarial attacks to deep learning-based video compression systems.
- Our comprehensive experiments show that our attacks outperform noise baselines and previously proposed attacks in both offline and online settings.
- Furthermore, our attacks still maintain high success rate in the presence of various defenses.
- Video demo is available at **https://sites.google.com/view/demo-of-rovisq/home**

Thank you!

Questions?

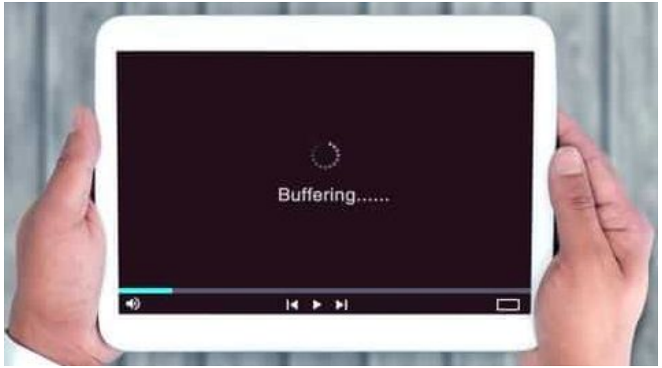
Supplementary Slides

Proposed Attacks

Original Video

Attacked Video

PSNR/Bpp 29.48 / 0.51576 | 29.47 / **0.9289**

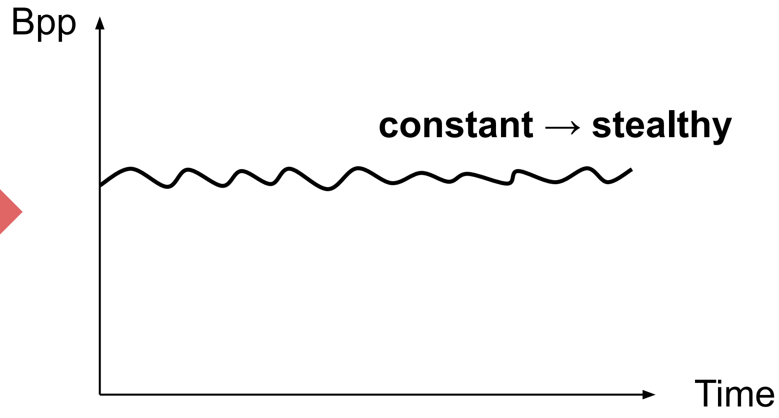


- Bandwidth Attack
  - This prevents legitimate users from successful communication with the streaming server and induces a high latency.
  - The end-users either experience **buffering** when downloading high-resolution videos due to increased bit-rate or a **reduced video resolution** at a fixed bit-rate.

Proposed Attacks

Original Video Attacked Video

PSNR/Bpp 29.48 / 0.51576 25.17 / 0.52034



- Video Quality Attack
  - This attack is particularly advantageous when the media server administrator is monitoring the network bandwidth in real time.
  - In this scenario, the service provider can detect anomalies in the bit-rate, but the proposed distortion attack remains stealthy.

Proposed Attacks

	Original Video	Attacked Video
		
PSNR/Bpp	29.48 / 0.51576	**24.22 / 0.8834**

- RD Attack
  - This attack combines the capabilities of the above two attacks by simultaneously targeting R and D to cause a high latency and video distortion.
  - The back-end users suffer from the **strongest** low-quality or denial-of-service.
  - If the media server lowers the video resolution to reduce network traffic, the RD attack is further exacerbated.

Experimental Results

- Defense against Adversarial Attacks on **Video Compression**
  - Our attacks still maintain high success rate in the presence of various defenses, such as adversarial training, video denoising, and JPEG coding.

Benchmark	w Defense PSNR (dB)	w Defense Bpp	w/o Defense PSNR (dB)	w/o Defense Bpp
DVC [44]	29.22	0.34	31.24	0.27
Video Quality (Offline)	-2.41	+0.6%	-3.52	+0.7%
Video Quality (Online)	-2.51	+16.4%	-3.05	+19.9%
Bandwidth (Offline)	-0.12	+84.2%	-0.01	+99.4%
Bandwidth (Online)	-0.75	+31.5%	-0.39	+35.7%
RD (Offline)	-2.88	+71.5%	-4.21	+85.3%
RD (Online)	-2.41	+25.6%	-3.10	+33.5%

Adversarial Training

Benchmark	w Defense PSNR (dB)	w Defense Bpp	w/o Defense PSNR (dB)	w/o Defense Bpp
DVC [44]	29.74	0.28	31.24	0.27
Video Quality (Offline)	-3.23	+0.5%	-3.52	+0.8%
Video Quality (Online)	-2.76	+14.3%	-3.05	+19.9%
Bandwidth (Offline)	-0.12	+64.8%	-0.01	+99.5%
Bandwidth (Online)	-0.43	+21.8%	-0.39	+35.7%
RD (Offline)	-3.81	+56.8%	-4.21	+85.3%
RD (Online)	-2.63	+18.4%	-3.10	+33.5%

Video Denoising

Benchmark	CF	w Defense PSNR (dB)	w Defense Bpp	w/o Defense PSNR (dB)	w/o Defense Bpp
DVC [44]	20	31.14	0.28	31.24	0.27
	40	29.26	0.21		
Video Quality (Offline)	20	-3.35	+0.7%	-3.52	+0.8%
	40	-3.14	+0.6%		
Video Quality (Online)	20	-2.86	+19.1%	-3.05	+19.9%
	40	-2.76	+18.4%		
Bandwidth (Offline)	20	-0.25	+95.4%	-0.01	+99.5%
	40	-0.45	+86.7%		
Bandwidth (Online)	20	-1.45	+34.2%	-0.39	+35.7%
	40	-1.76	+31.2%		
RD (Offline)	20	-4.09	+82.6%	-4.21	+85.3%
	40	-3.71	+70.5%		
RD (Online)	20	-2.95	+31.8%	-3.10	+33.5%
	40	-2.79	+28.6%		

JPEG Compression

Experimental Results

- Defense against Adversarial Attacks on **Video Classification**
  - Our attacks still maintain high success rate in the presence of various defenses, such as adversarial training, video denoising, and JPEG coding.

Video Classifier	Defense	ACC (%) w/o Defense	ACC Drop (%)	ASR (%) w Defense	ASR (%) w/o Defense
SlowFast [21]	AT [46]	85.4	-11.3	68.2	93.2
	JPEG [67]		-5.2	75.5	
	Denoising [16]		-7.5	76.9	
TPN [73]	AT [46]	74.3	-10.1	63.1	92.0
	JPEG [67]		-2.5	74.8	
	Denoising [16]		-4.0	75.3	
I3D [13]	AT [46]	71.7	-8.0	76.2	92.1
	JPEG [67]		-7.4	80.1	
	Denoising [16]		-5.8	81.8	

Targeted Attack

Video Classifier	Defense	ASR (%) w Defense		ASR (%) w/o Defense	
		Offline	Online	Offline	Online
SlowFast [21]	AT [46]	67.1	53.2	96.1	80.4
	JPEG [67]	72.3	64.6		
	Denoising [16]	73.3	64.1		
TPN [73]	AT [46]	64.2	58.2	95.8	81.3
	JPEG [67]	70.9	61.2		
	Denoising [16]	71.8	63.8		
I3D [13]	AT [46]	75.8	65.3	96.3	80.7
	JPEG [67]	80.8	72.2		
	Denoising [16]	82.7	68.5		

Untargeted Attack