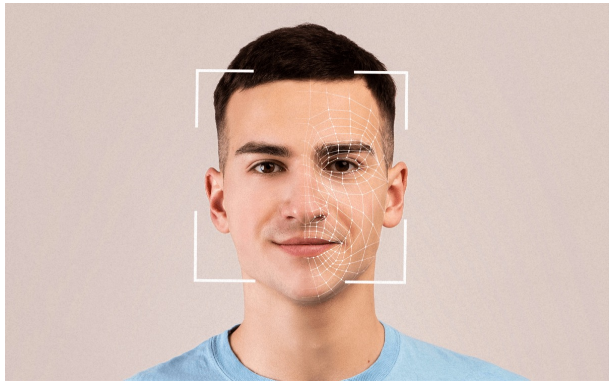
Privacy-Preserving Mechanisms on Data-Driven Deep Learning Applications

Department of Computer Science
Georgia State University

Dr. Zhipeng Cai
Email: zcai@gsu.edu

Data-Driven Deep Learning Applications



Face Recognition

Automatic Retailing

Automatic Driving

AI-aided Medical Diagnosis

AI-based Attacks

TECHNOLOGY EXECUTIVE COUNCIL

Artificial intelligence is playing a bigge
bad g

PUBLISHED TUE, SEP

Bob Violino

1. Twitter accused of covering up data breach that affects millions

On No
about
and EU
before
2022, a

3. Personal and medical data for 11 million people accessed in Optus data breach

Australi
2022 th

The info
home a

4. Hacker attempts to sell data of 500 million WhatsApp users on dark web

On No
be up

In the
numb
detail

9. SHEIN fined US$1.9mn over data breach affecting 39 million customers

In October, Zoetop Business Company, the firm that owns fast fashion brands SHEIN and ROMWE, was fined US$1.9mn by the state of New York after failing to disclose a data breach which affected 39 million customers.

Every Coin Has Two Sides

Convenience

Privacy Leakage

Tradeoff !!!

Utility v.s. Privacy

Data Privacy Leakage in Deep Learning Applications

Data Modality in Applications

Image/Video Data Privacy

Audio Data Privacy

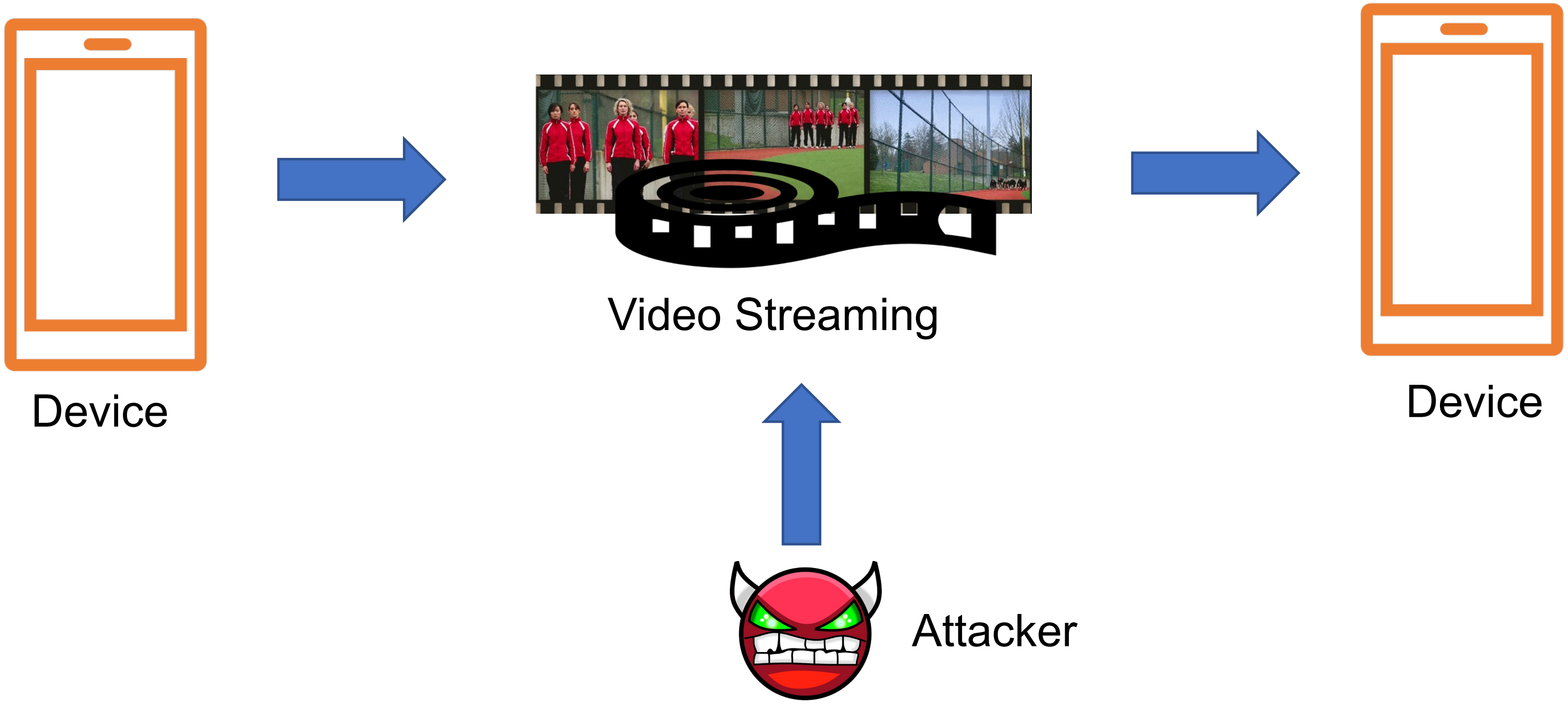
Text Data Privacy

Audio-Visual Autoencoding for Privacy-Preserving Video Streaming

1. Research Background
2. Existing Privacy-Preserving Mechanisms
3. New Challenges
4. Our Novel Design
5. Evaluation

1. Research Background
2. Existing Privacy-Preserving Mechanisms
3. New Challenges
4. Our Novel Design
5. Evaluation

Research Background

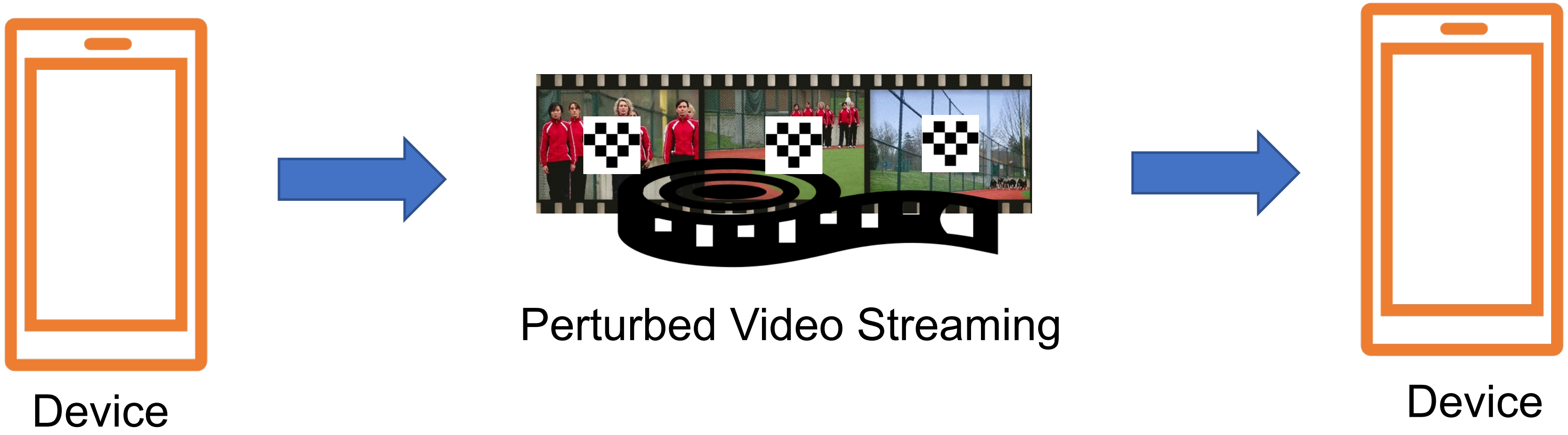


Audio-Visual Autoencoding for Privacy-Preserving Video Streaming

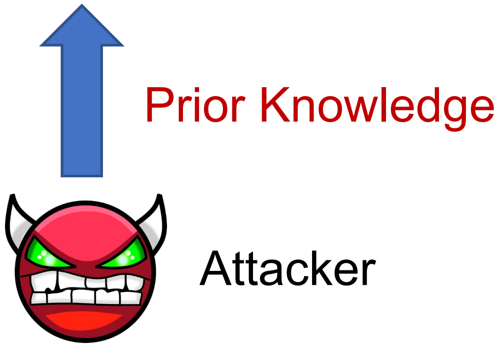
1. Research Background
2. Existing Privacy-Preserving Mechanisms
3. New Challenges
4. Our Novel Design
5. Evaluation

Existing Privacy-Preserving Mechanisms

1. Noise-based Privacy-Preserving Models



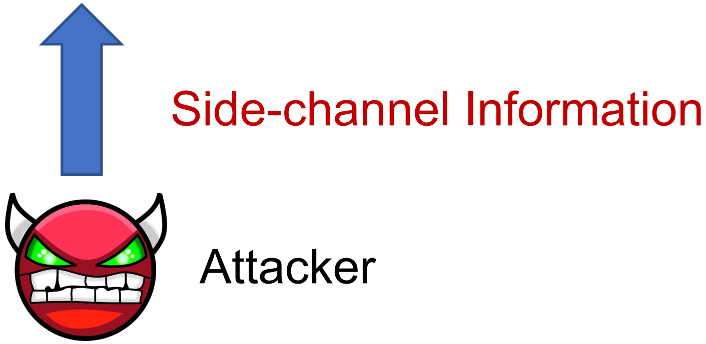
Attention: Random noise follows some patterned distributions (e.g., normal distribution).



Existing Privacy-Preserving Mechanisms

2. Encryption-based Privacy-Preserving Models

Attention: The traffic data can also be used to infer some private information.



Audio-Visual Autoencoding for Privacy-Preserving Video Streaming

1. Research Background
2. Existing Privacy-Preserving Mechanisms
3. New Challenges
4. Our Novel Design
5. Evaluation

New Challenges

In the design of a privacy-preserving mechanism:

1) To avoid the utilization of random noise with patterned distribution

2) To hide the side-channel information during transmission

Also, for this specific scenario (video streaming transmission)

3) Is it possible to consider the temporal information in terms of privacy preservation design?

Audio-Visual Autoencoding for Privacy-Preserving Video Streaming

1. Research Background
2. Existing Privacy-Preserving Mechanisms
3. New Challenges
4. Our Novel Design
5. Evaluation

Our Novel Design

We perturb the video streaming with its extracted audio.

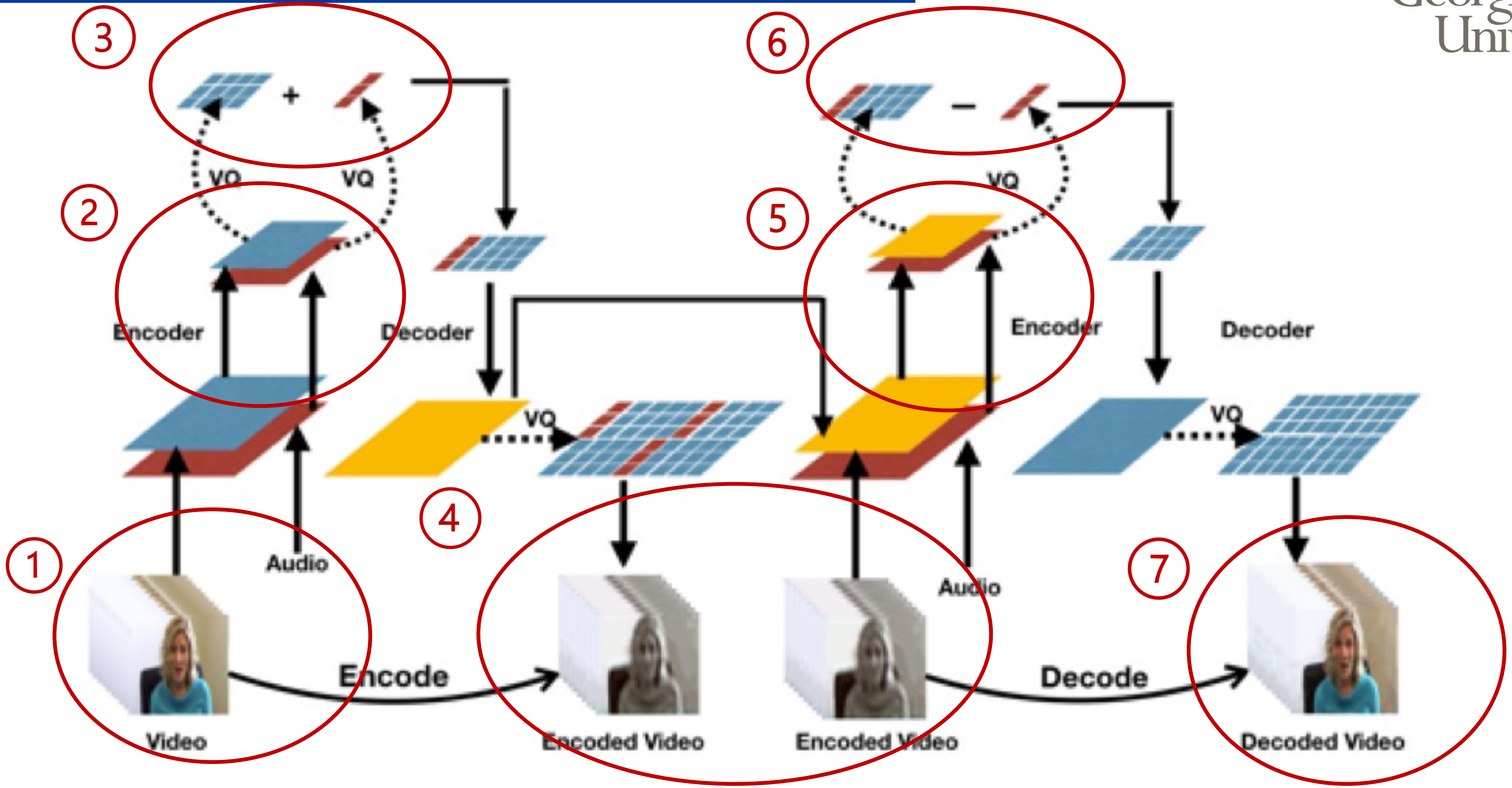
1) The extracted audio is unique and has no patterned distribution. No prior knowledge

We use the audio to encode the video streaming.

2) The traffic data flow is smoothed due to the video compression during transmission. No side-channel information

Audio-Visual Autoencoding Scheme

Cycle-VQ-VAE



3
+
VQ
VQ
2
Encoder
Decoder
VQ
1
Audio
Video
Encode
4
Encoded Video
6
−
VQ
5
Encoder
Decoder
VQ
Encoded Video
Audio
Decode
7
Decoded Video

Two Versions

Audio-Visual Autoencoding Scheme

Cycle-VQ-VAE

Is it possible to consider the temporal information in terms of privacy preservation design?

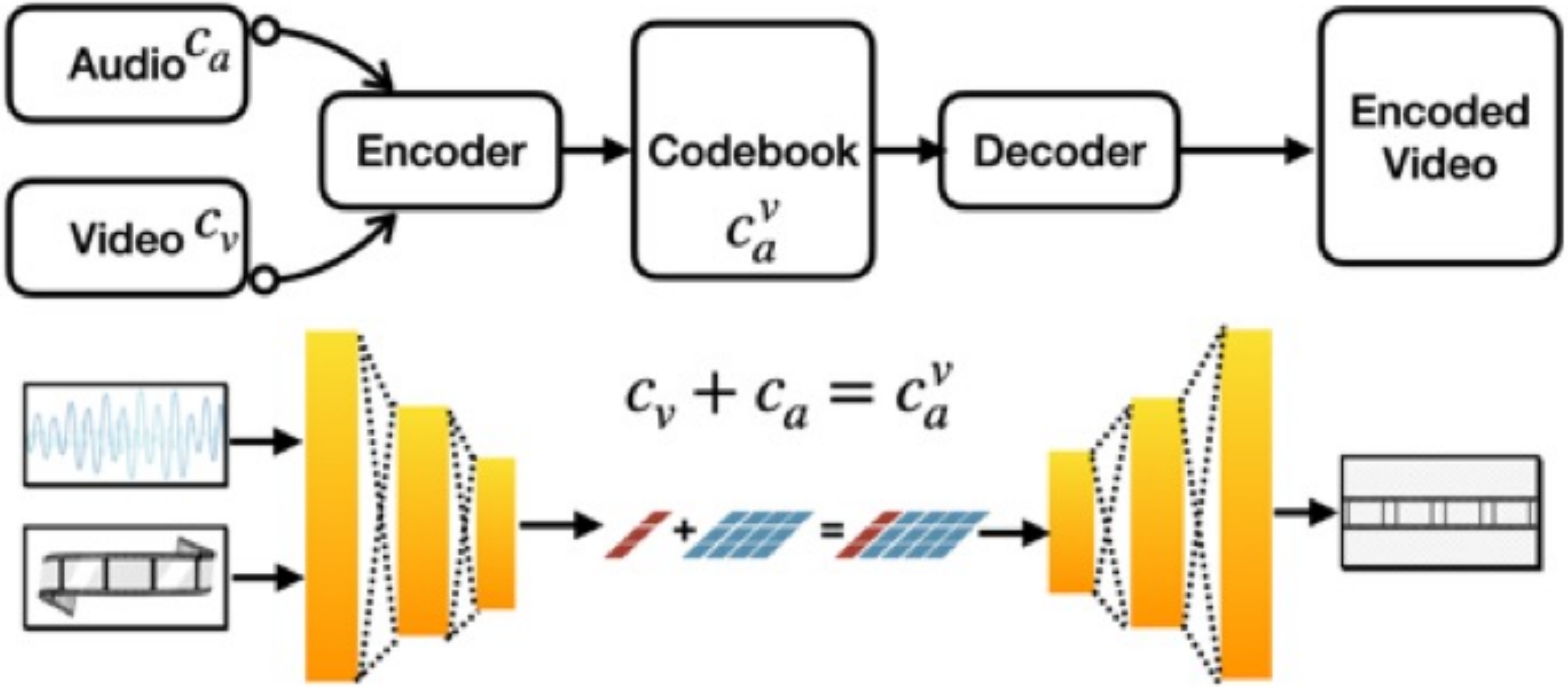
Without considering temporal information

Frame-to-Frame (F2F) Model

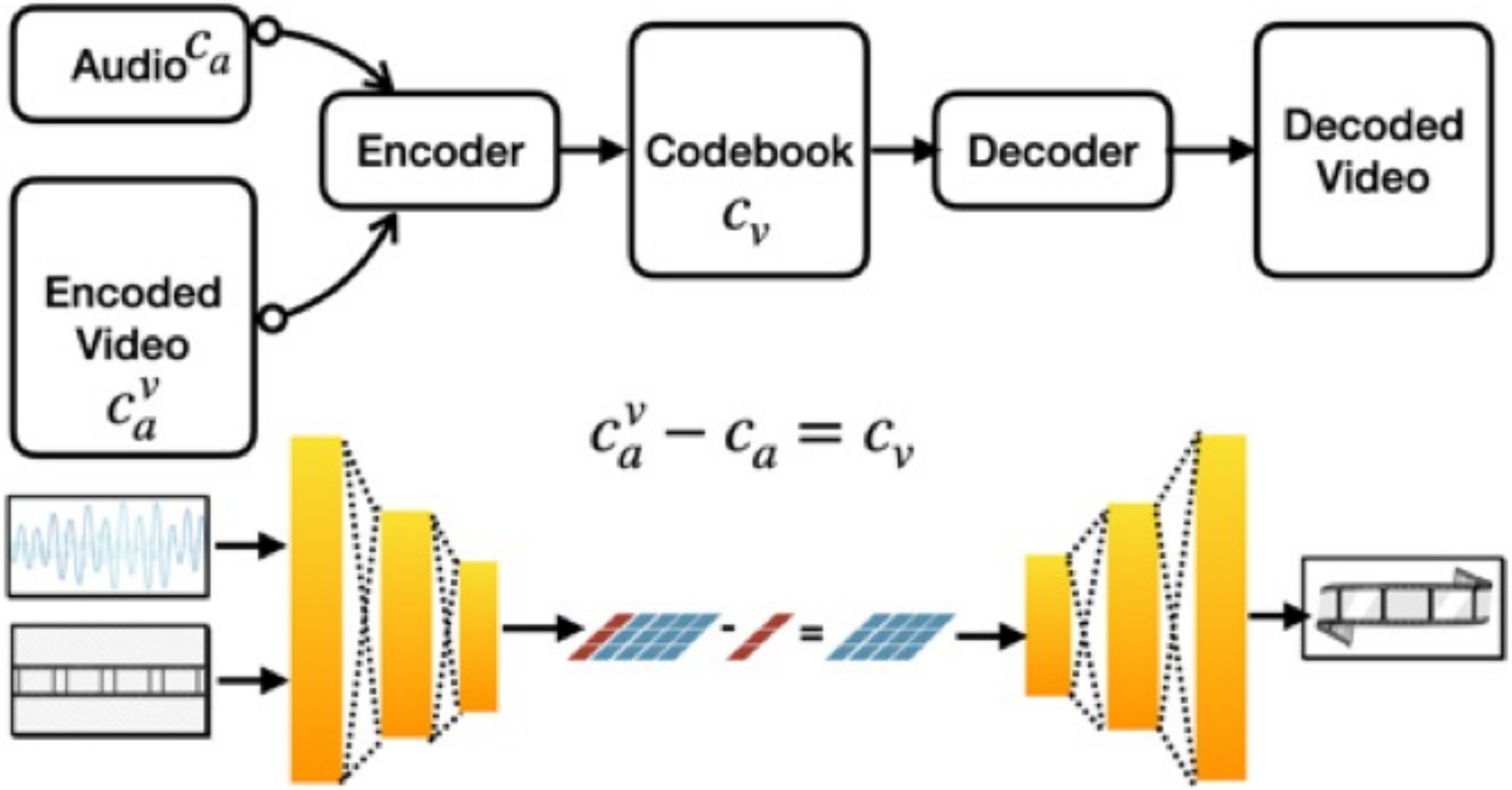
Considering temporal information

Video-to-Video (V2V) Model

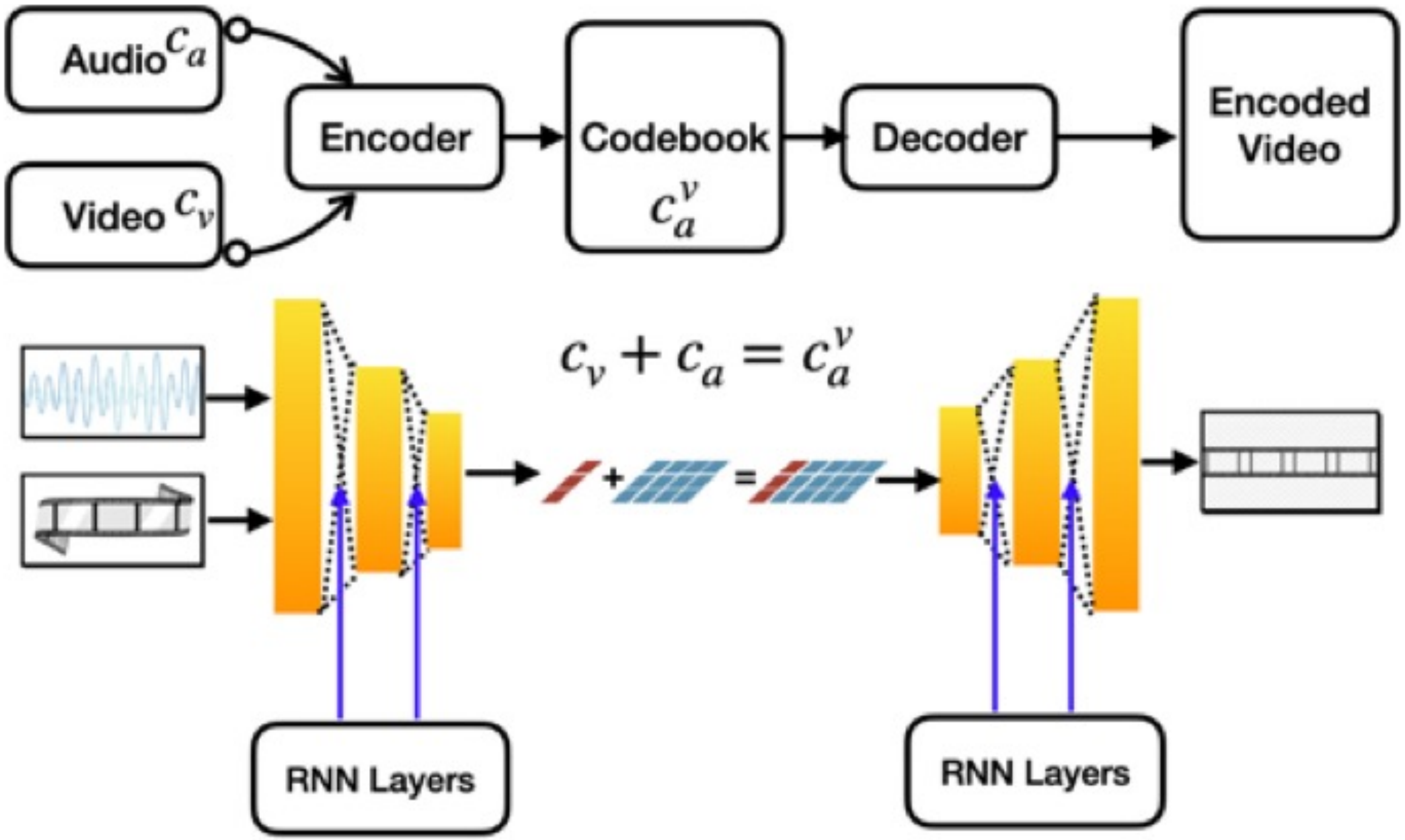
F2F Model --- Encoding



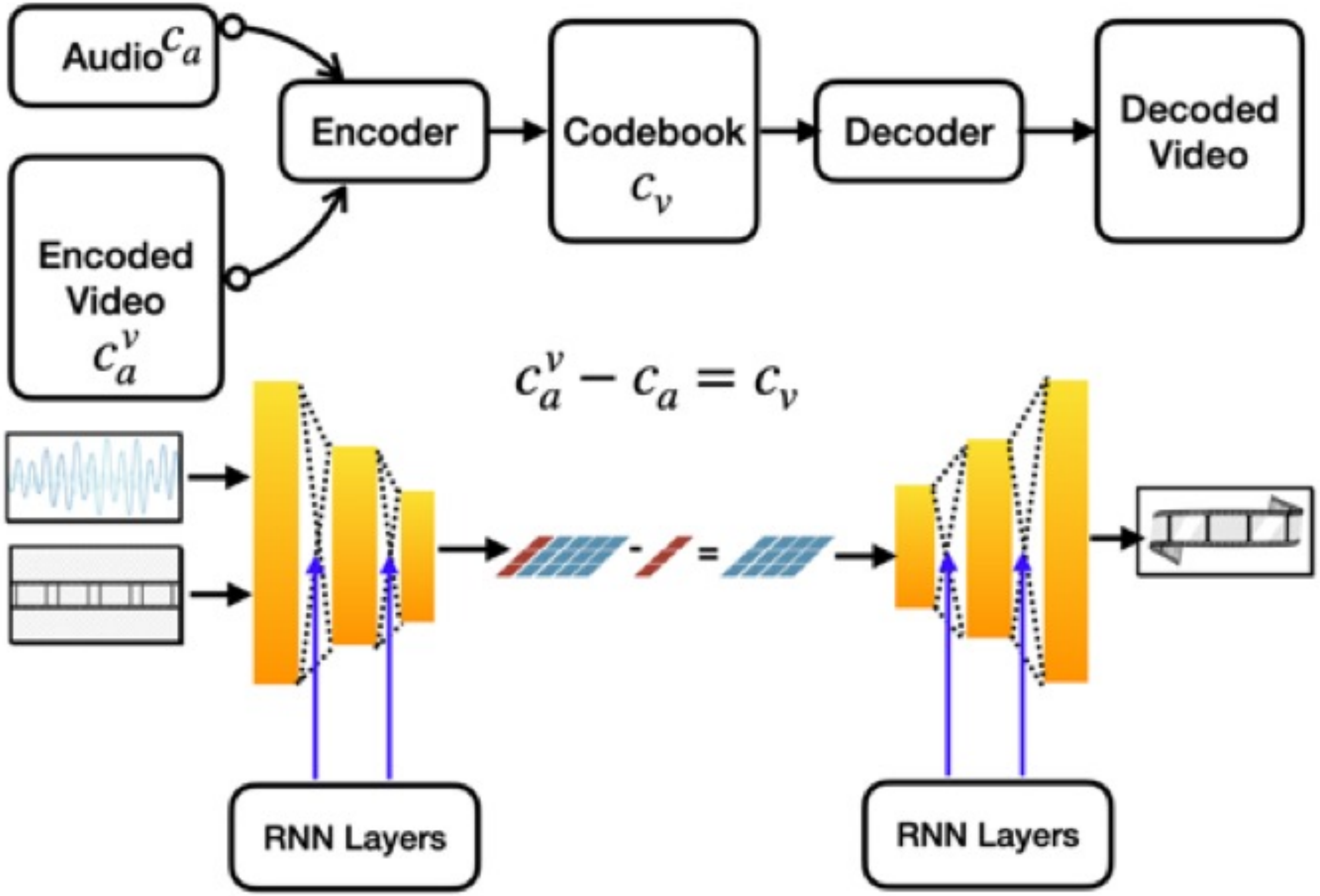
F2F Model --- Decoding



V2V Model --- Encoding



V2V Model --- Decoding



Audio-Visual Autoencoding for Privacy-Preserving Video Streaming

1. Research Background
2. Existing Privacy-Preserving Mechanisms
3. New Challenges
4. Our Novel Design
5. Evaluation

Experiment Settings

1. Dataset

Extract the video frames and the audio from 200 videos in the AVE dataset

2. AI Detection Model

1) Face Detection Model; 2) Semantic Segmentation Model

3. Side-channel Inference Attack Model → Effectiveness in terms of privacy protection

Analyze traffic data to achieve the activity recognition

4. Two Baselines → Superiority of Our proposed Models

1) AE-based Model; 2) Style Translator-based Model

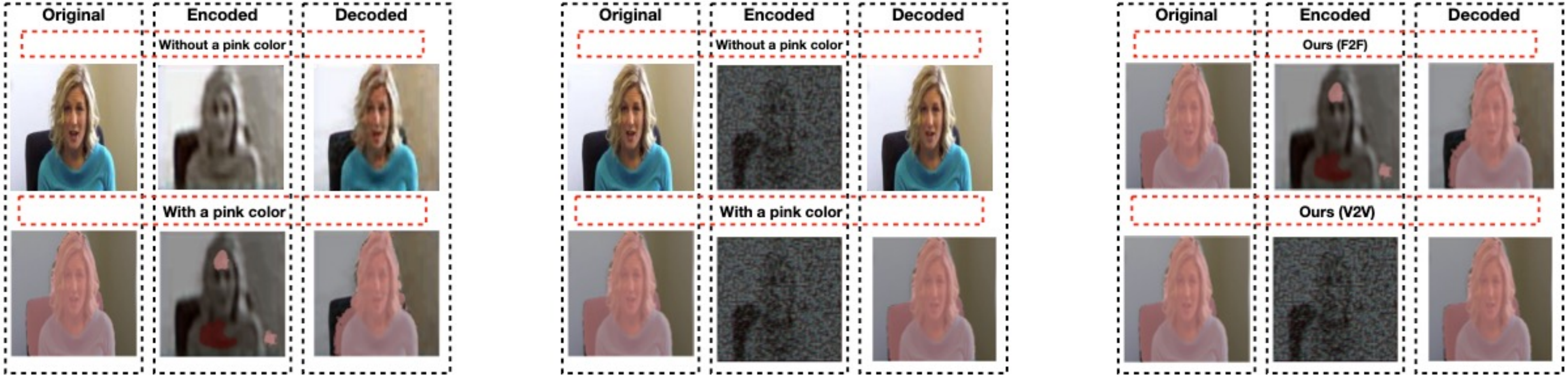
Face Detection Performance (Ours)

(a) Face Detection on F2F Video Frames

(b) Face Detection on V2V Video Frames

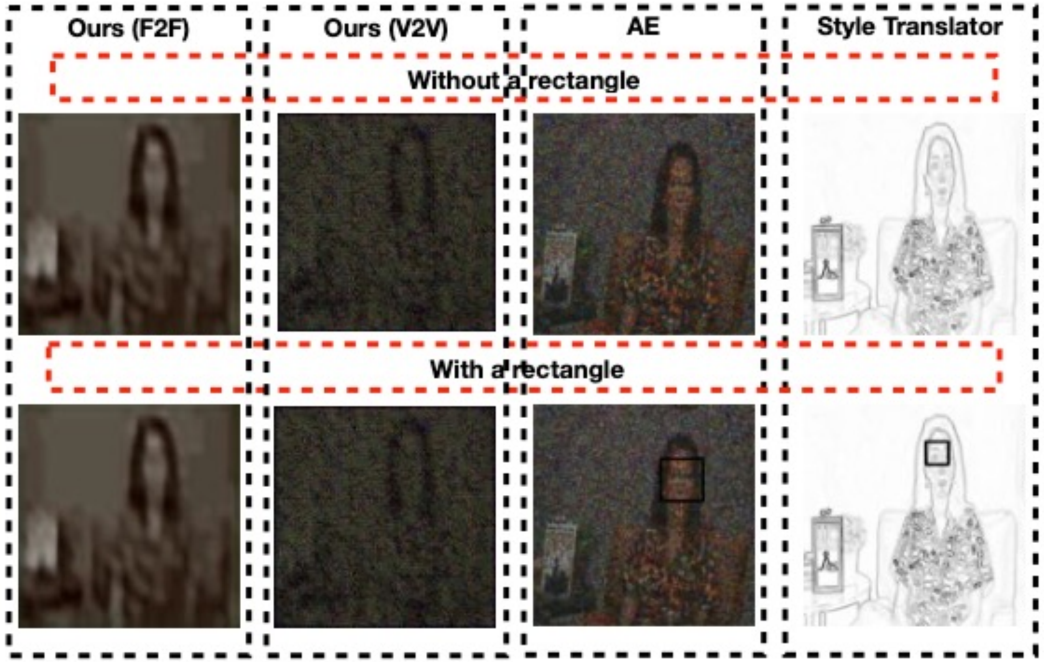
(c) Face Detection Comparison

Semantic Segmentation Performance (Ours)

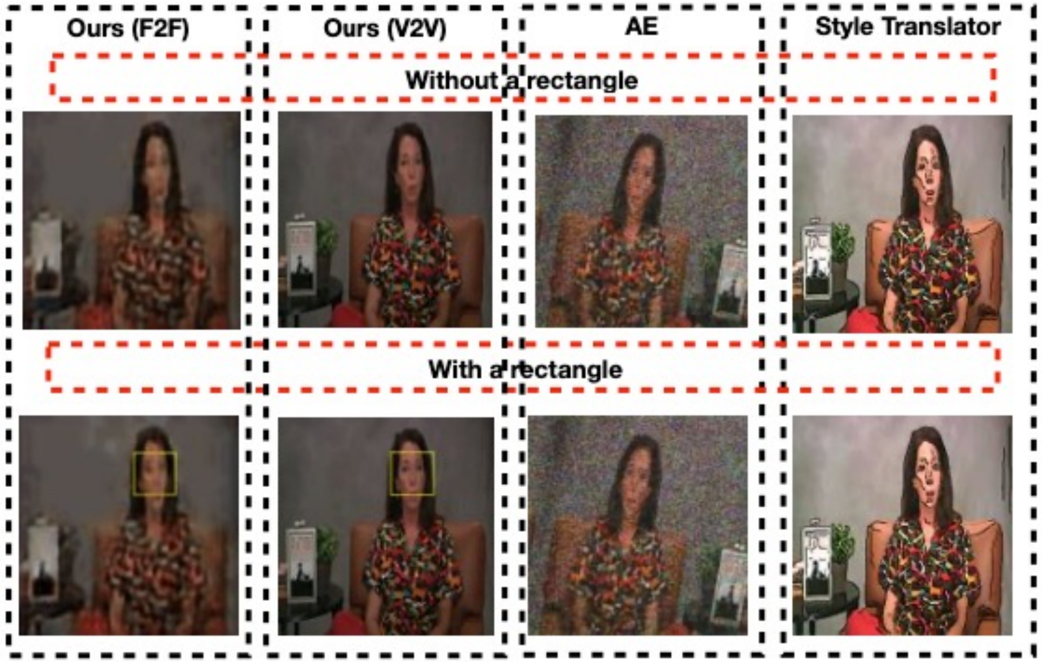


(d) Semantic Segmentation on F2F Video Frames (e) Semantic Segmentation on V2V Video Frames (f) Semantic Segmentation Comparison

Face Detection (Ours v.s. Baselines)

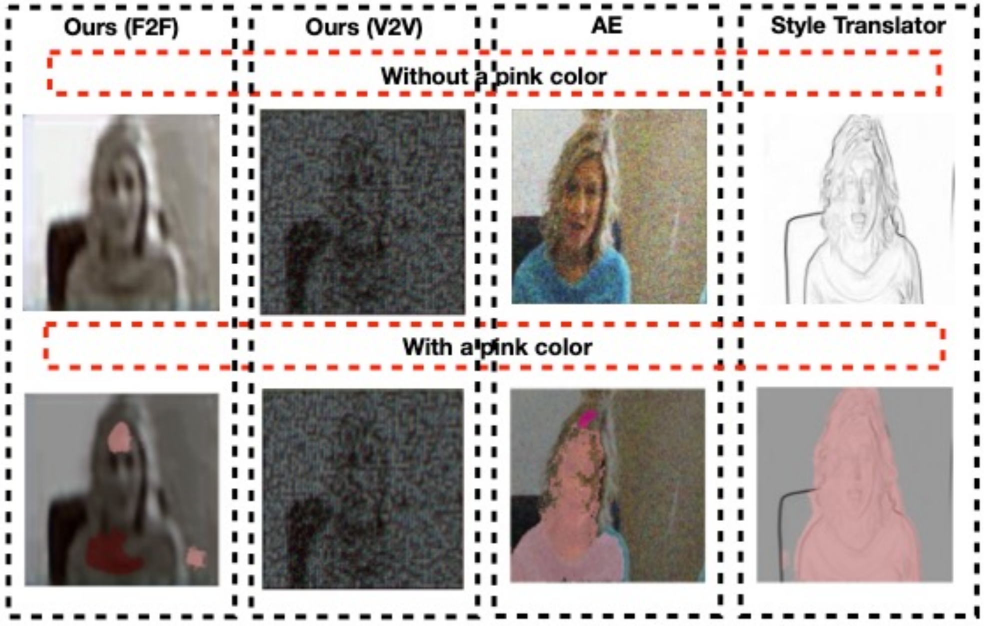


(a) Face Detection on Encoded Video Frames: Ours *v.s.* Others



(b) Face Detection on Decoded Video Frames: Ours *v.s.* Others

Semantic Segmentation (Ours v.s. Baselines)



(c) Semantic Segmentation on Encoded Video Frames: Ours *v.s.* Others

(d) Semantic Segmentation on Dencoded Video Frames: Ours *v.s.* Others

Accuracy of Face Detection

TABLE II
ACCURACY OF FACE DETECTION

	Ours(F2F)	Ours(V2V)	AE	Style Translator
Original	96.67%	96.67%	96.67%	96.67%
Encoded	6.00%	0.00%	26.67%	36.67%
Decoded	80.00%	96.67%	46.67%	63.33%

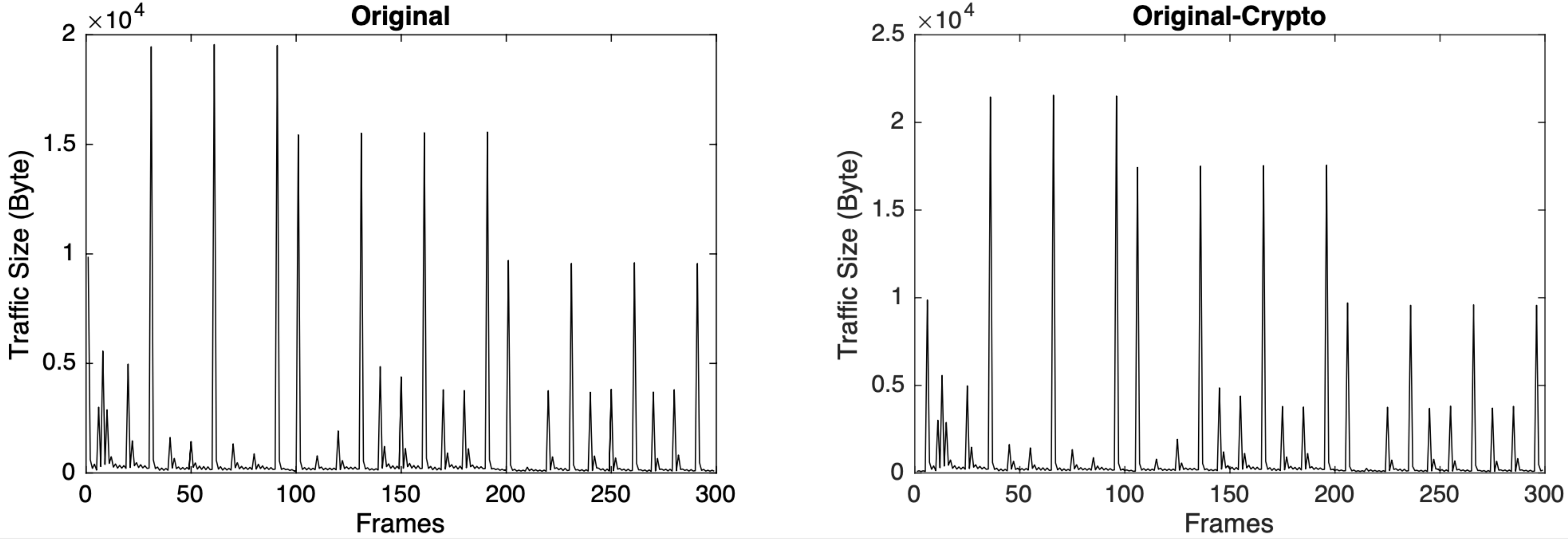
Accuracy of Semantic Segmentation

TABLE III

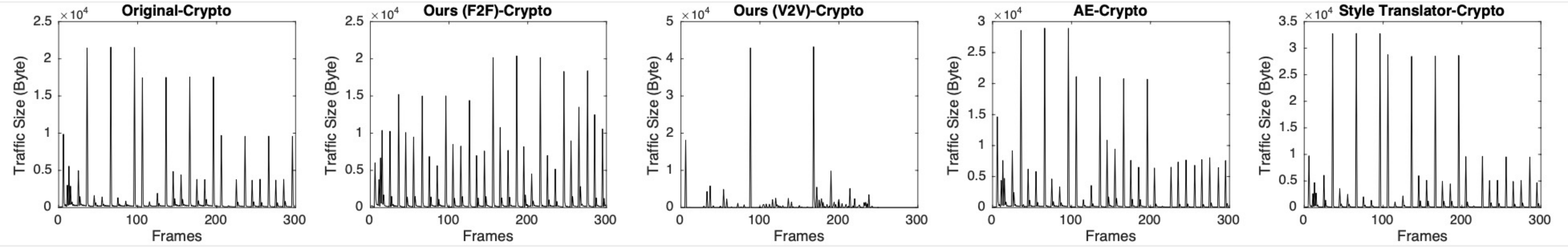
ACCURACY OF SEMANTIC SEGMENTATION

	Ours(F2F)	Ours(V2V)	AE	Style Translator
Original	93.30%	93.30%	93.30%	93.30%
Encoded	6.70%	0.00%	20.00%	36.67%
Decoded	73.33%	93.30%	43.30%	60.00%

Side-Channel Privacy Leakage



Traffic Size after Privacy Protection



Activity Inference

TABLE IV
RESULTS OF ACTIVITY INFERENCE

	Accuracy		Accuracy
Original	95.80%	Original-Crypto	94.90%
Ours (F2F)	42.86%	Ours (F2F)-Crypto	41.98%
Ours (V2V)	0.00%	Ours (V2V)-Crypto	0.00%
AE	95.60%	AE-Crypto	94.80%
Style Translator	94.50%	Style Translator-Crypto	93.70%

Encoding Process helps Video Compression

Transmission Efficiency

TABLE V
TRANSMISSION TIME AT DIFFERENT BANDWIDTHS (OURS (F2F) V.S. OTHERS)

	Original	Ours (F2F)	Ours (V2V)	AE	Style Translator
0.5MB/s	3.84s	3.24s(↓ 15.6%)	1.75s(↓ 54.4%)	5.6s(↑ 45.8%)	4.2s(↑ 9.3%)
1MB/s	1.87s	1.57s(↓ 16.1%)	0.87s(↓ 53.1%)	2.68s(↑ 43.3%)	2.05s(↑ 9.6%)
2MB/s	0.94s	0.78s(↓ 17.1%)	0.44s(↓ 52.7%)	1.34s(↑ 42.5%)	1.02s(↑ 8.5%)
Average		↓ 16.2%	↓ 53.4%	↑ 43.8%	↑ 9.1%

Recap

Data Modality in Applications

Image/Video Data Privacy | Audio Data Privacy | Text Data Privacy

Privacy-Preserving Mechanisms

Non-theoretical | Theoretical

Conclusion

Audio-Visual Autoencoding for Privacy-Preserving Video Streaming

Image/Video Data Privacy

Non-theoretical

More Works

Privacy-Preserving Mechanisms for Multi-label Image Recognition

Image Data Privacy

Theoretical

Privacy-Preserving Multimodal Sentiment Analysis

Multimodal Data Privacy

Theoretical

Future Work

From the viewpoint of users to the viewpoint of service providers

On-going work: Defense for Side-Channel Attack on Deep Learning Architecture

Thank You !!!