IEICE English Webinar
Distinguished Lecturer Program Series
October 7, 2022

Real or Fake?
From Biometric Data Protection to Fake Media Detection

Isao Echizen

Director / Professor, Information and Society Research Division,

National Institute of Informatics

Director, Global Research Center for Synthetic Media,

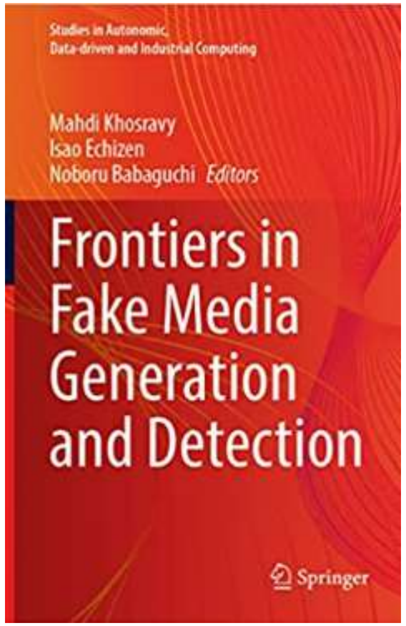
National Institute of Informatics

Joint work with Prof. Junichi Yamagishi,
Dr. Trung-Nghia Le, and Dr. Huy H. Nguyen

SynMedia Center

CREST FakeMedia

Short bio: Isao Echizen



M. Khosravy, I. Echizen, and N. Babaguchi, eds. Springer, June 2022

<u>Titles</u>

1995	BSc., Tokyo Institute of Technology
1997	MSc., Tokyo Institute of Technology
2003	Dr.Eng.,Tokyo Institute of Technology

<u>Career</u>

1997-2007	Systems Development Laboratory, Hitachi, Ltd.
2007-2014	Associate Professor, National Institute of Informatics (NII)
2014-Current	Professor, NII
2018-2020	Deputy Director General, NII
2019-Current	Professor, Graduate School of Information Science and Technology, The University of Tokyo
2021-Current	Director, Information and Society Research Division, NII
2021-Current	Director, Global Research Center for Synthetic Media, NII

SynMedia Center

<u>Other important positions</u>

2010	Visiting Professor, University of Freiburg, Germany
2011	Visiting Professor, University of Halle-Wittenberg, Germany
2020-Current	Japanese Representative, IFIP TC11 (Security and Privacy Protection)
2020-2026	Research Director, JST CREST FakeMedia (Research Area: Trusted quality AI systems)

CREST FakeMedia

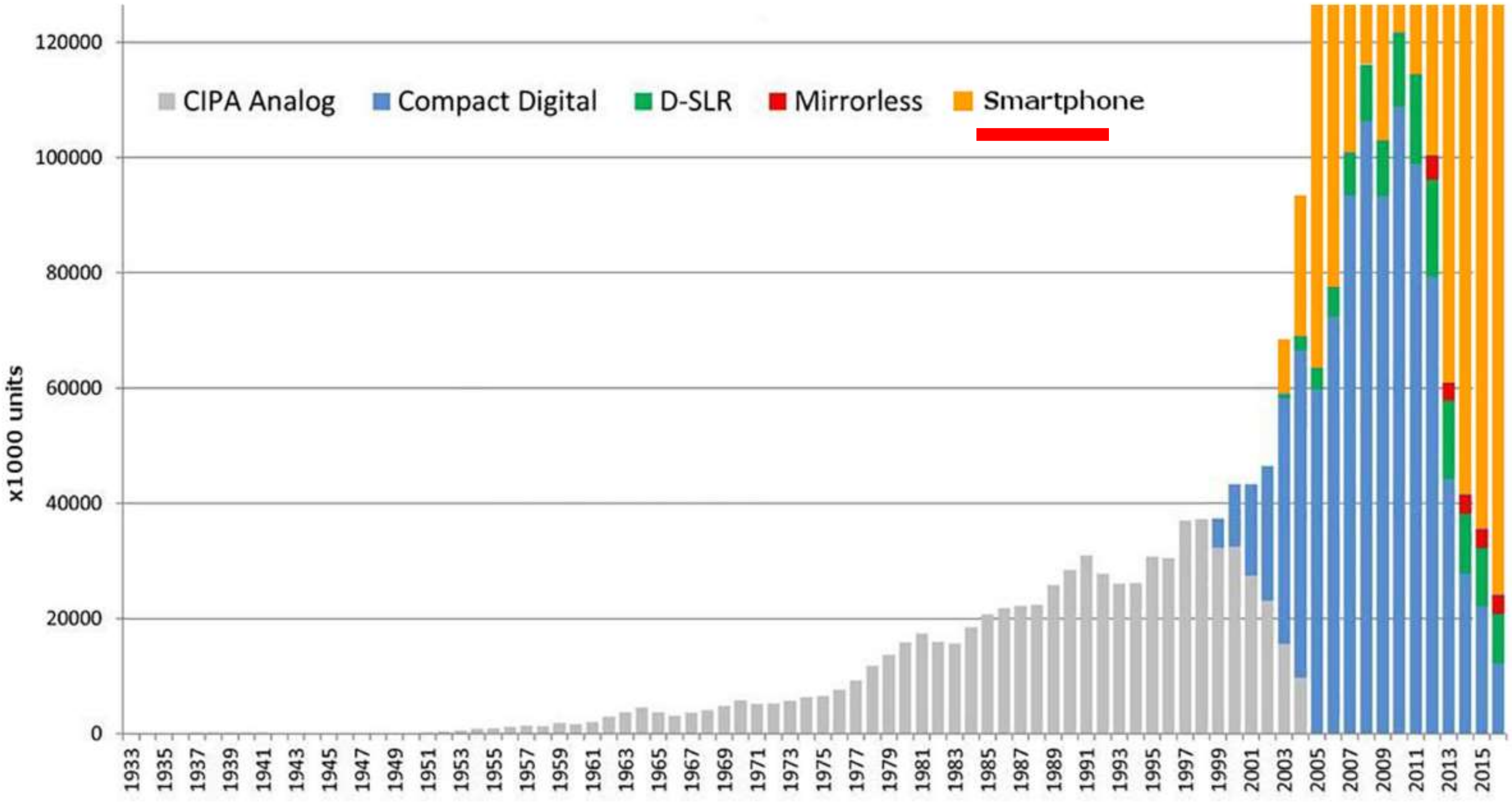
<u>Awards</u>

Information Security Cultural Award(2016), DOCOMO Mobile Science Award(2014), Best Paper Award(WIFS17), etc.

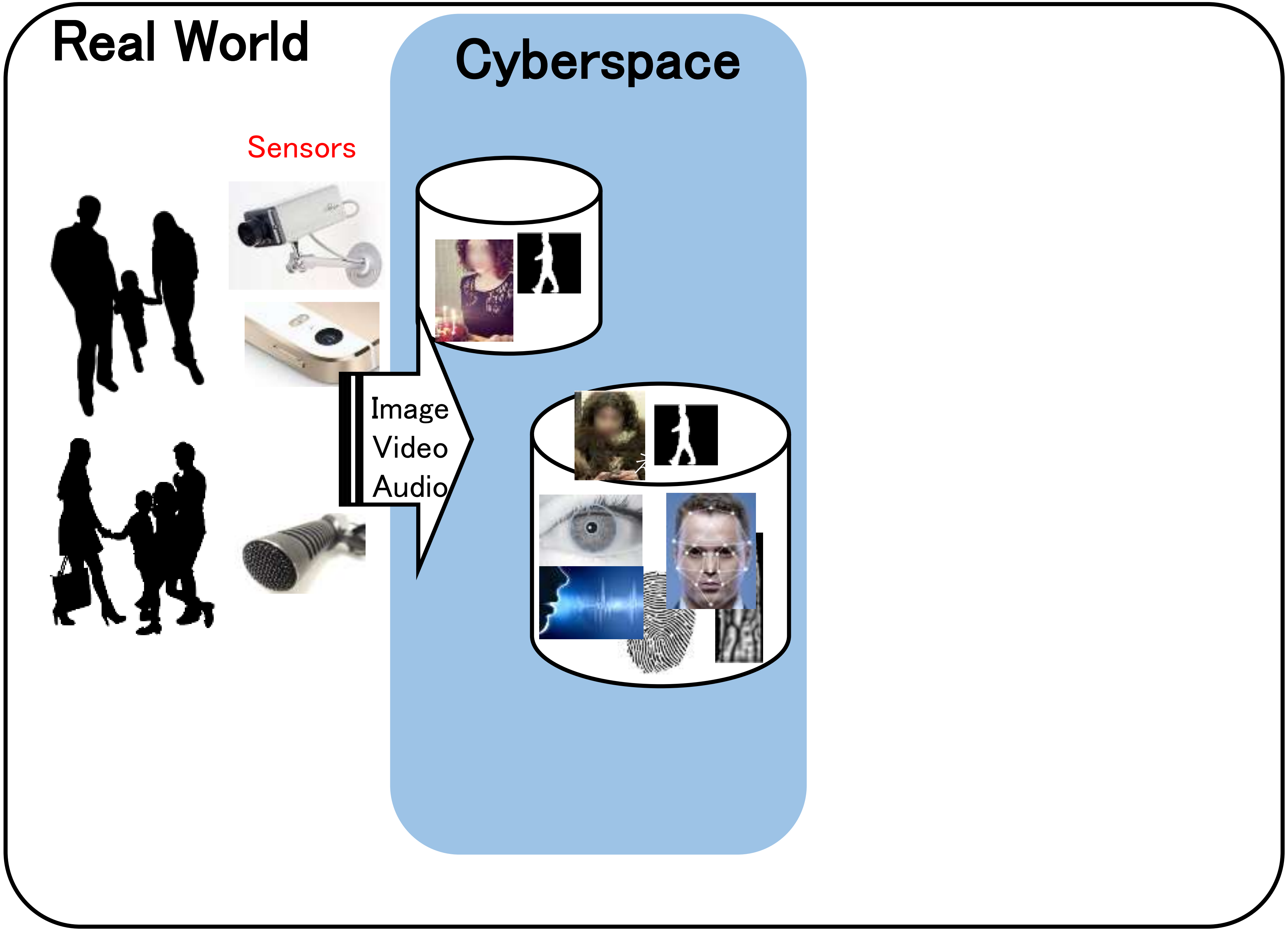
of cameras produced: explosive growth of smartphones

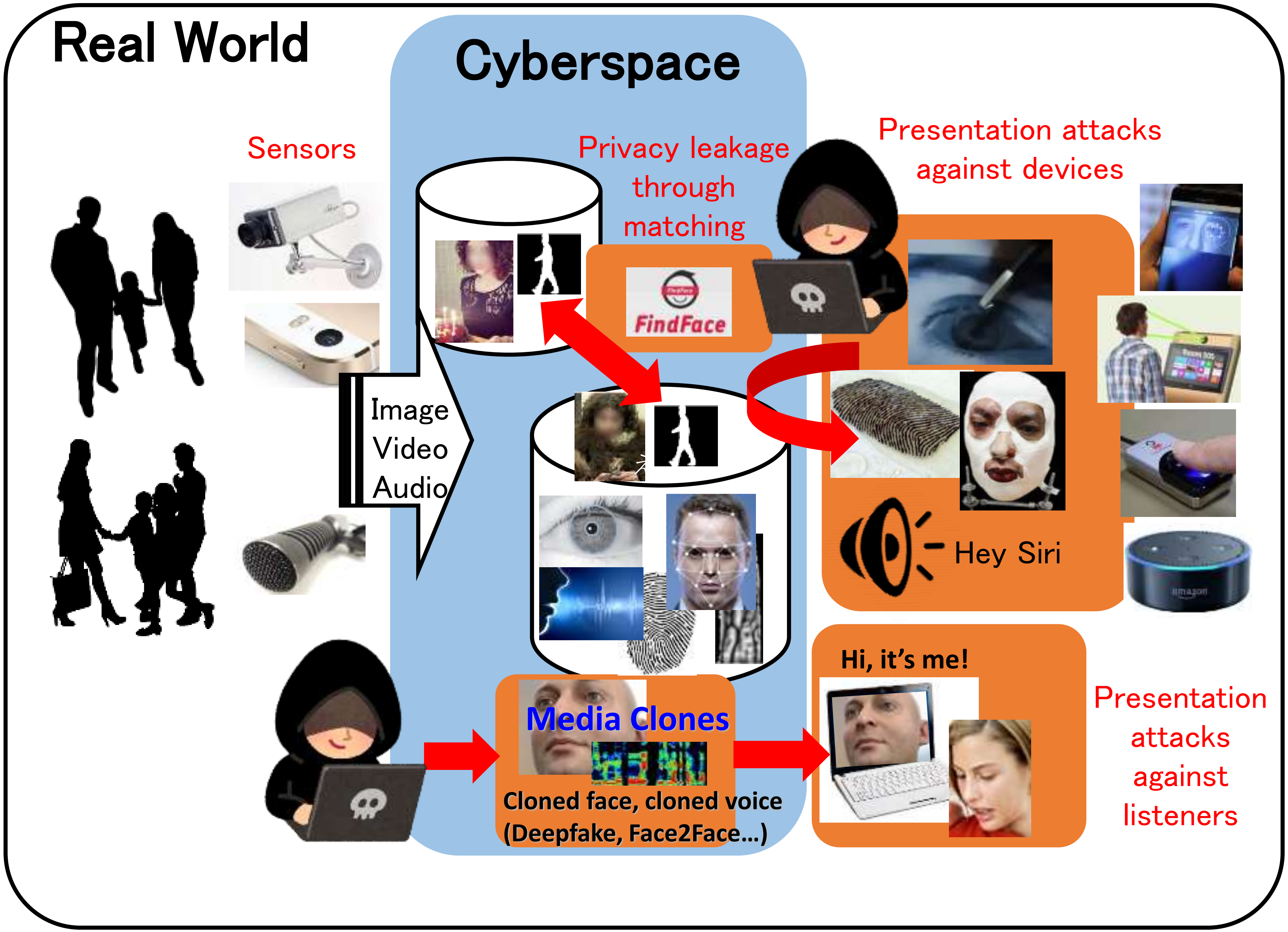
Annual production volume of cameras : 40 million (2001) → 1.5 billion (2016)

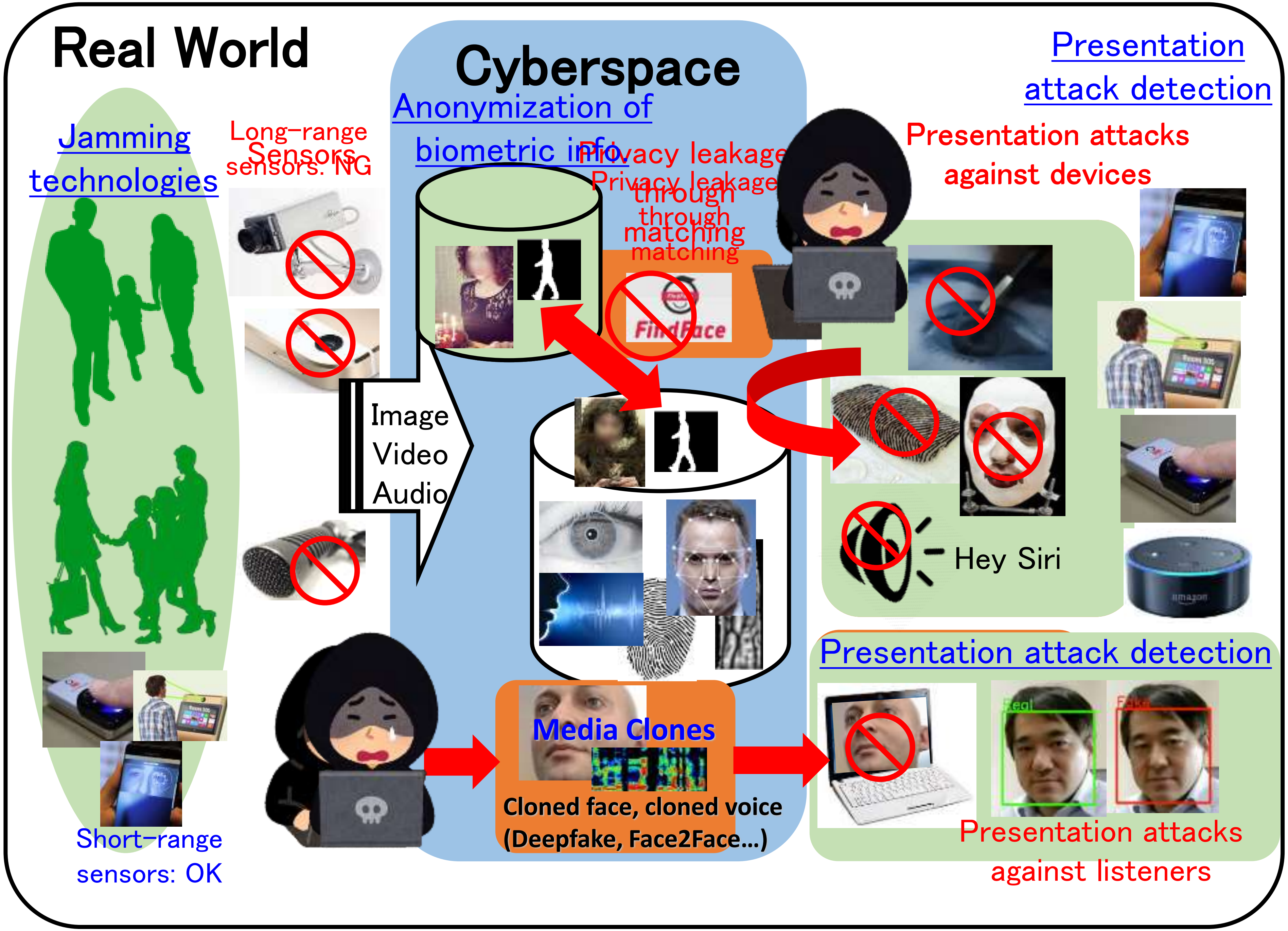
Security and privacy issues in sharing biometric information in cyberspace



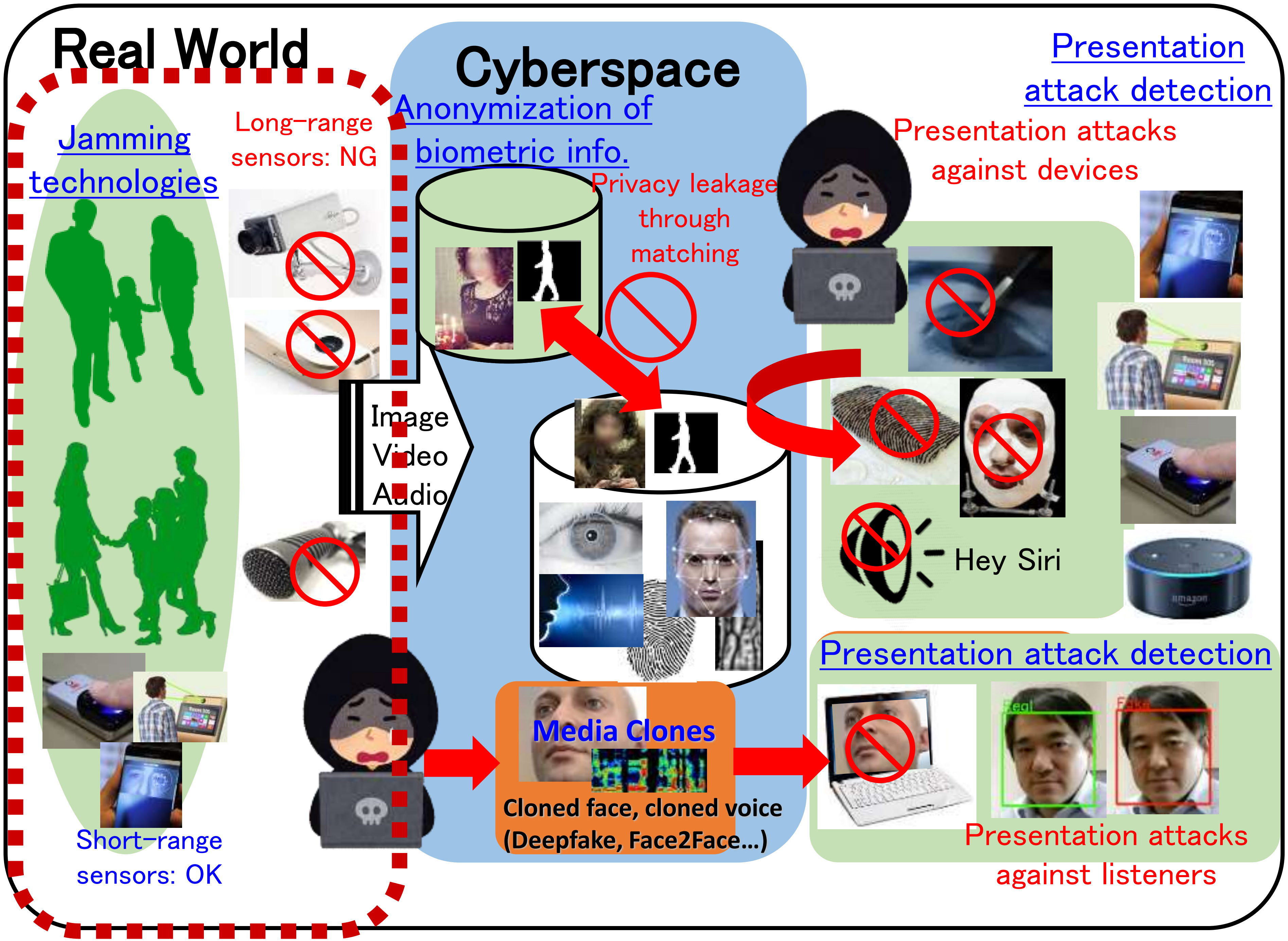
CIPA camera production 1933-2016

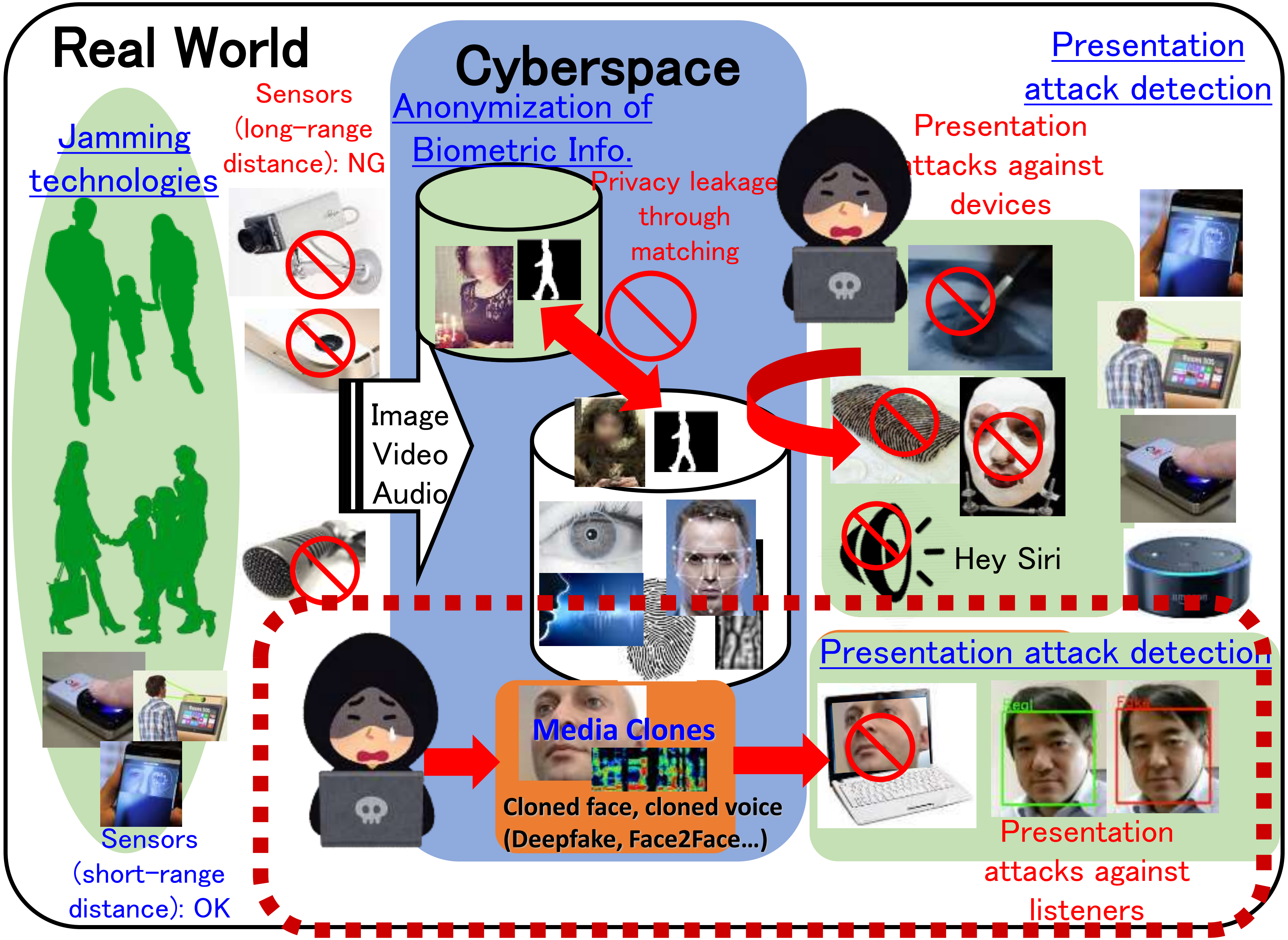

Real World
Cyberspace
Sensors
Image
Video
Audio


Real World
Cyberspace
Sensors
Privacy leakage
through
matching
FindFace
Presentation attacks
against devices
Image
Video
Audio
Hey Siri
amazon
Hi, it's me!
Media Clones
Cloned face, cloned voice
(Deepfake, Face2Face…)
Presentation
attacks
against
listeners



Real World
Cyberspace
Jamming technologies
Long-range sensors: NG
Sensors
Anonymization of biometric info.
Privacy leakage through matching
Privacy leakage through matching
FindFace
Presentation attack detection
Presentation attacks against devices
Image
Video
Audio
Hey Siri
amazon
Short-range sensors: OK
Media Clones
Cloned face, cloned voice
(Deepfake, Face2Face...)
Presentation attack detection
Real
Fake
Presentation attacks against listeners


Real World
Cyberspace
Jamming technologies
Long-range sensors: NG
Anonymization of biometric info.
Privacy leakage through matching
Presentation attack detection
Presentation attacks against devices
Image
Video
Audio
Hey Siri
Short-range sensors: OK
Media Clones
Cloned face, cloned voice (Deepfake, Face2Face...)
Presentation attack detection
Presentation attacks against listeners


Real World
Cyberspace
Jamming technologies
Sensors (long-range distance): NG
Anonymization of Biometric Info.
Privacy leakage through matching
Presentation attack detection
Presentation attacks against devices
Image
Video
Audio
Hey Siri
Sensors (short-range distance): OK
Media Clones
Cloned face, cloned voice (Deepfake, Face2Face...)
Presentation attack detection
Presentation attacks against listeners

Detection of computer generated fake media
(2018-current)

1. D. Afchar, V. Nozick, J. Yamagishi, and I. Echizen, " MesoNet: a Compact Facial Video Forgery Detection Network, " Proc. of the IEEE International Workshop on Information Forensics and Security (WIFS 2018), pp.1-7, December 2018
2. Huy H. Nguyen, Junichi Yamagishi, and Isao Echizen, Capsule-forensics: using capsule networks to detect forged images and videos, Proc. IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP), 5 pages, (May 2019)
3. Huy H. Nguyen, Fuming Fang, Junichi Yamagishi, Isao Echizen,"Multi-task Learning For Detecting and Segmenting Manipulated Facial Images and Videos," Proc. of the BTAS 2019,8 pages,(September 2019)

Outline

- Introduction, Generating Fake Media Using Human-Related Information
- Methods for Generating Fake Media Based on Faces
- Methods for Detecting Fake Media Based on Faces
- Advanced Fake Media Generation and Detection Methods
- Toward Countering Infodemics (JST CREST FakeMedia, NII SynMedia Center)

Fake or Real?

Fake

Real

StyleGAN / StyleGAN 2 (Karras et al. 2019/2020).
Using progressive training strategy and a style-based image generation approach.

Fake or Real?

Real

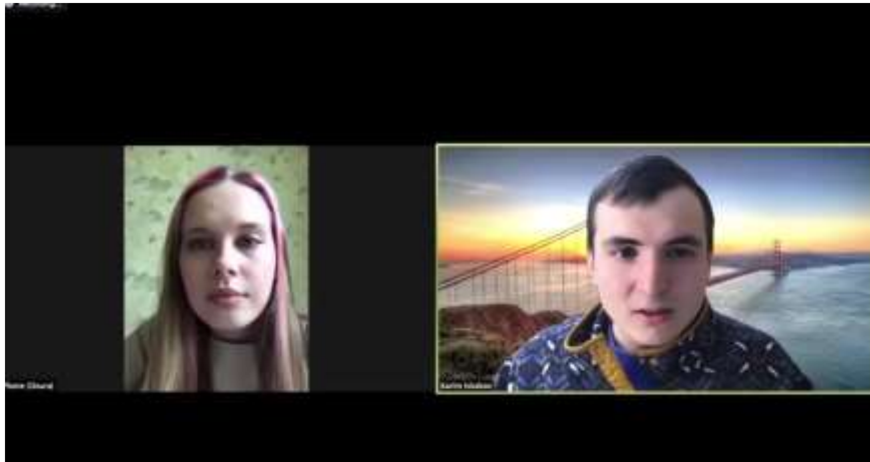
Fake

StyleGAN / StyleGAN 2 (Karras et al. 2019/2020).
Using progressive training strategy and a style-based image generation approach.

A. Rossler, D. Cozzolino, L. Verdoliva, C. Riess, J. Thies, and M. Niessner. Faceforensics++: Learning to detect manipulated facial images. In International Conference on Computer Vision, pages 1–11, Oct 2019.

Fake media generation using human-related information

- AI learns from human-related information such as faces, voices, bodies, and natural language to generate fake media
  - Deepfake (fake facial video, 2018-), GROVER (fake news, 2019-)
  - Impersonate CEO with fake voice and exploit cash (2019)
  - Impersonate a fictitious person to manipulate stock prices (2019)
  - Participate in the Zoom conference by pretending to be Elon Musk with a fake face (2020)



Elon Musk joined our Zoom call | Avatarify
https://www.youtube.com/watch?v=lONuXGNqLO0

THE WALL STREET JOURNAL.

PRO CYBER NEWS

Fraudsters Used AI to Mimic CEO's Voice in Unusual Cybercrime Case

Scams using artificial intelligence are a new challenge for companies

WSJ, August 30, 2019

The CEO of a British energy company received a fake voice call pretending to be the CEO of the parent company and wired EUR 220,000 to the company.

https://www.wsj.com/articles/fraudsters-use-ai-to-mimic-ceos-voice-in-unusual-cybercrime-case-11567157402

FAST COMPANY

04-30-19

How to spot the realistic fake people creeping into your timelines

A remarkable advance in artificial portrait generation adds a new potential layer of deception to online fraudsters, astroturfers, and propagandists.

FastCompany, April 30, 2019

Using the AI-generated profile image, he created a fake Twitter account named Maisy Kinsley (a Bloomberg journalist), contacted Tesla shareholders to obtain their personal information, and then planned to manipulate Tesla's stock price.

https://www.fastcompany.com/90332538/how-to-spot-the-creepy-fake-faces-who-may-be-lurking-in-your-timelines-deepfaces

Outline

- Introduction, Generating Fake Media Using Human-Related Information
- Methods for Generating Fake Media Based on Faces
- Methods for Detecting Fake Media Based on Faces
- Advanced Fake Media Generation and Detection Methods
- Toward Countering Infodemics (JST CREST FakeMedia, NII SynMedia Center)

Generating Fake Media for Faces: Five Types

1. Entire face synthesis
   - Generate (non-real-world) facial images from noise (latent variables) (StyleGAN, VQ-VAE, etc.)
2. Attribute manipulation: hair, skin color, expression
   - Generate a facial image of the target with a different hair color, skin color, expression, etc. (StarGAN, ELEGANT, etc.)
3. Facial reenactment
   - Generate facial images of the target that are synchronized with the attacker's facial expressions (Face2Face, ICFace, etc.)
4. Speaking manipulation
   - Generate facial images of the target speaking the voice / text by synthesizing the voice / text with the source facial images of the target (e.g., Synthesizing Obama)
5. Face swap
   - Replace the face part of the source video with the target face (e.g. Faceswap)

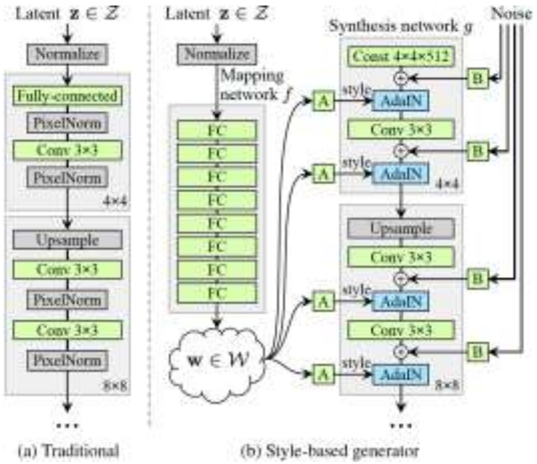
Generating Fake Media for Faces: Five Types

1. Entire face synthesis

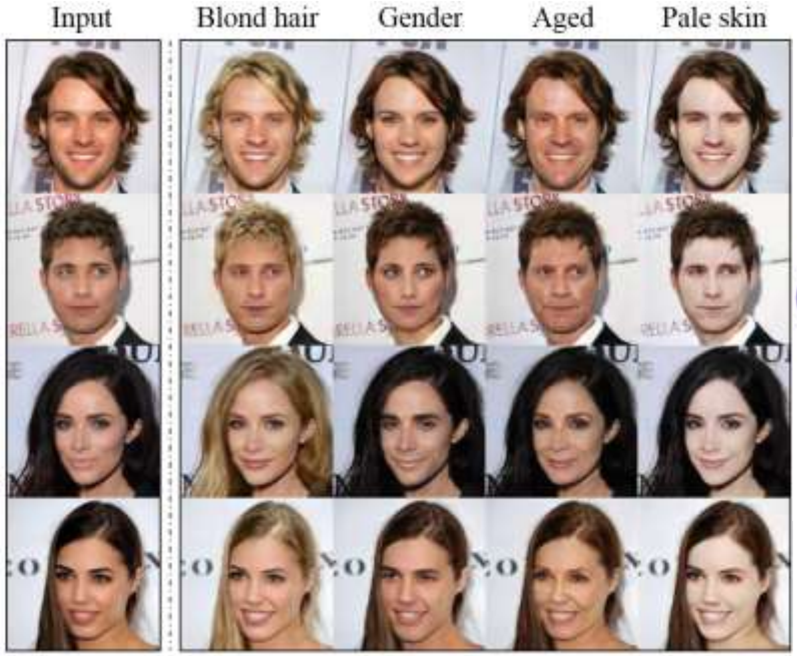
- Generate (non-real-world) facial images from noise (latent variables) (StyleGAN, VQ-VAE, etc.)

2. Attribute manipulation: hair, skin color, expression

- Generate a facial image of the target with a different hair color, skin color, expression, etc. (StarGAN, ELEGANT, etc.)



StyleGAN / StyleGAN 2[1] (Karras et al. 2019/2020). Using progressive training strategy and a style-based image generation approach.



StarGAN (Choi et al. 2018). Image-to-image translation for multiple domains.

Generating Fake Media for Faces: Five Types

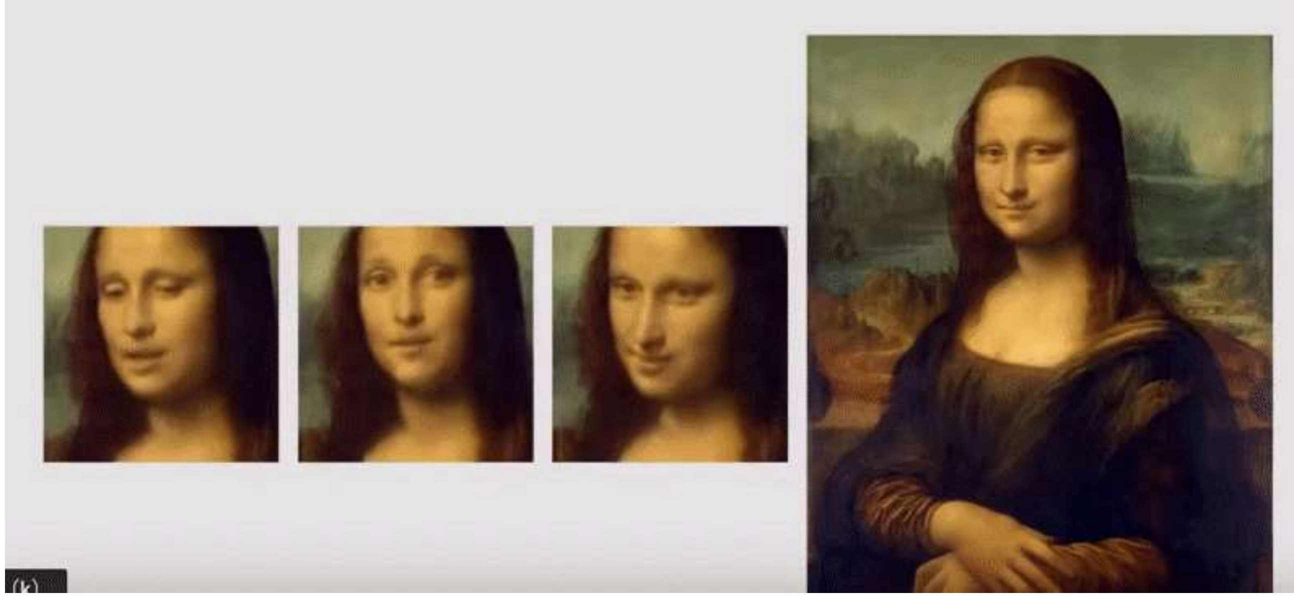
3. Facial reenactment

- Generate facial images of the target that are synchronized with the attacker's facial expressions (Face2Face, ICFace, etc.)

Video (attacker) + video (victim)→ forged video

Face2Face (Thies et al. 2016).
Transferring facial movements of one person to the other one.

Video (attacker) + image (victim)→ forged video



Neural Talking Head Models
(Zakharov et al. 2019)

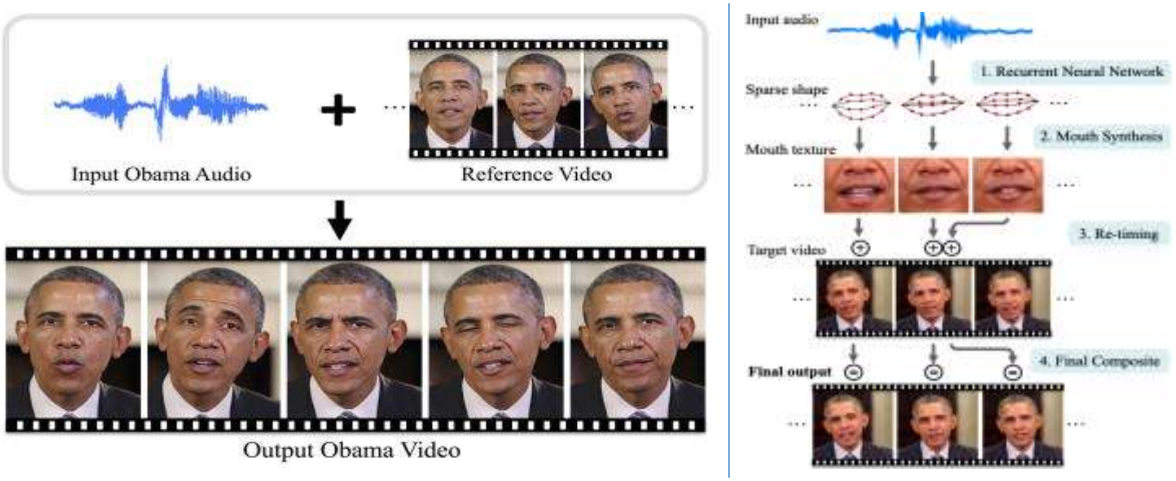
Generating Fake Media for Faces: Five Types

4. Speaking manipulation

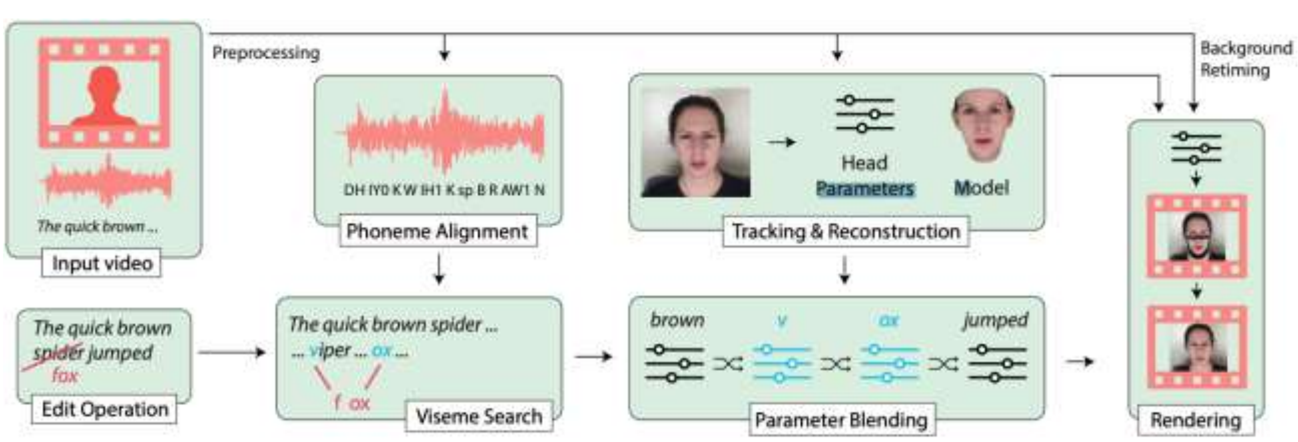
- Generate facial images of the target speaking the voice / text by synthesizing the voice / text with the source facial images of the target (e.g., Synthesizing Obama)

Synthesized speech (attacker) + image/video (victim)
→ forged video

Modified text (attacker) + video (victim)
→ forged video



Synthesizing Obama
(Suwajanakorn et al. 2017)



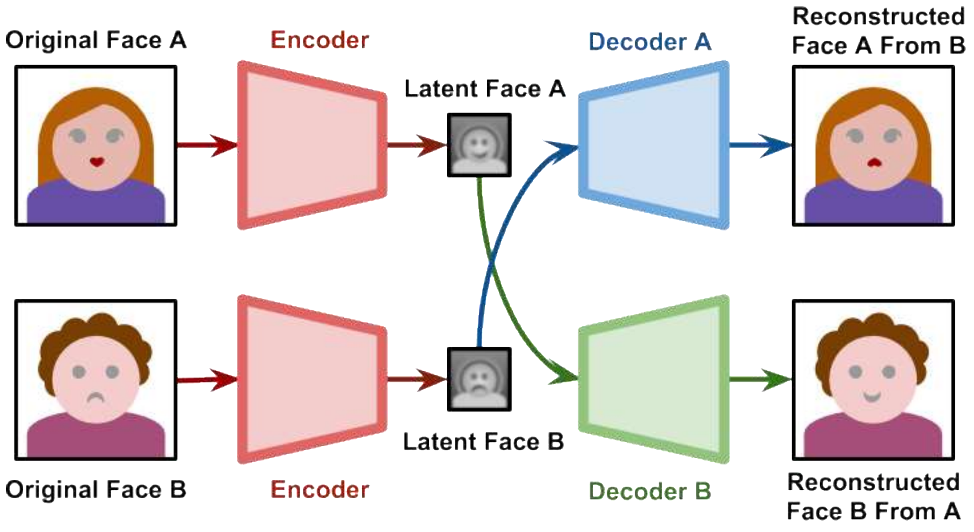
Text-based Editing of Talking-head Video
(Fried et al. 2019)

Generating Fake Media for Faces: Five Types

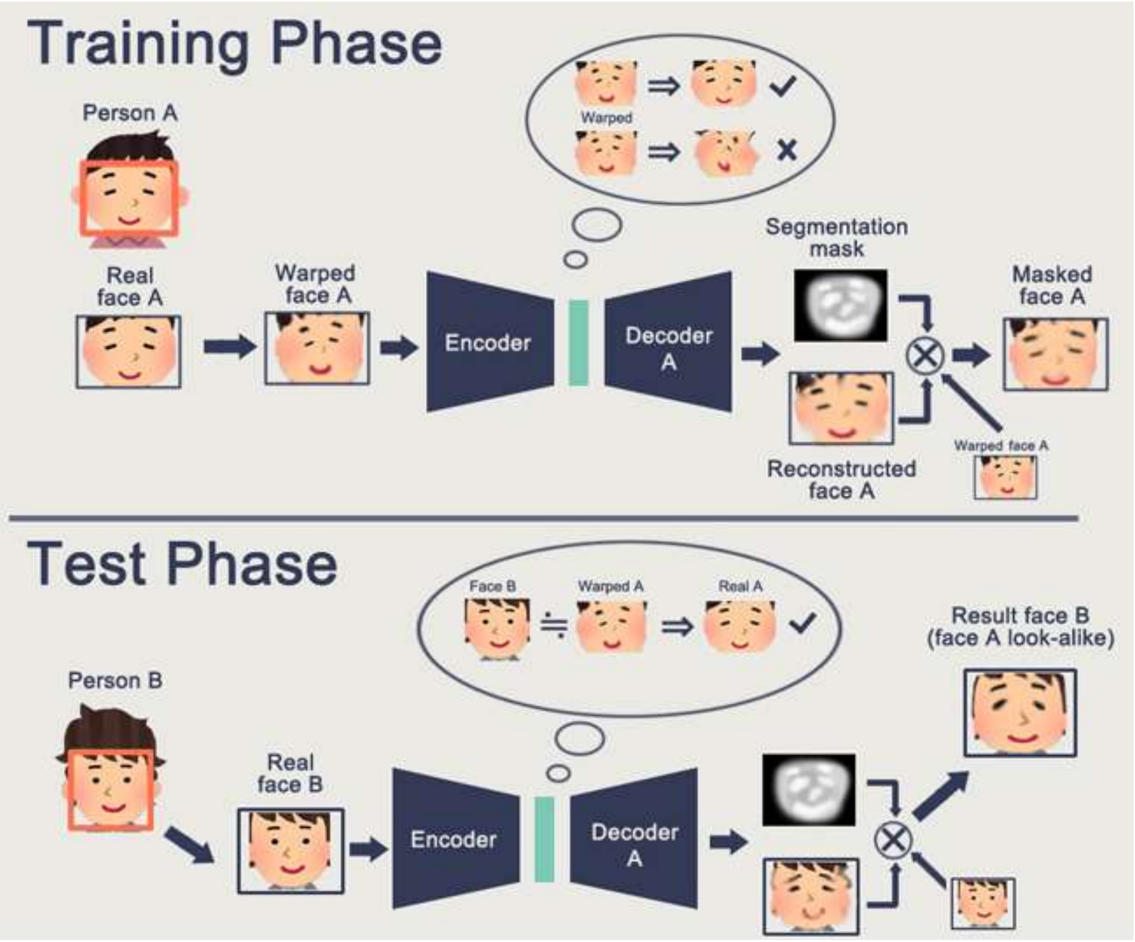
5. Face swap

- Replace the face part of the source video with the target face (e.g. Faceswap)

Deep learning based face swap



Original Deepfake (Faceswap)[1]
Image: Alan Zucconi

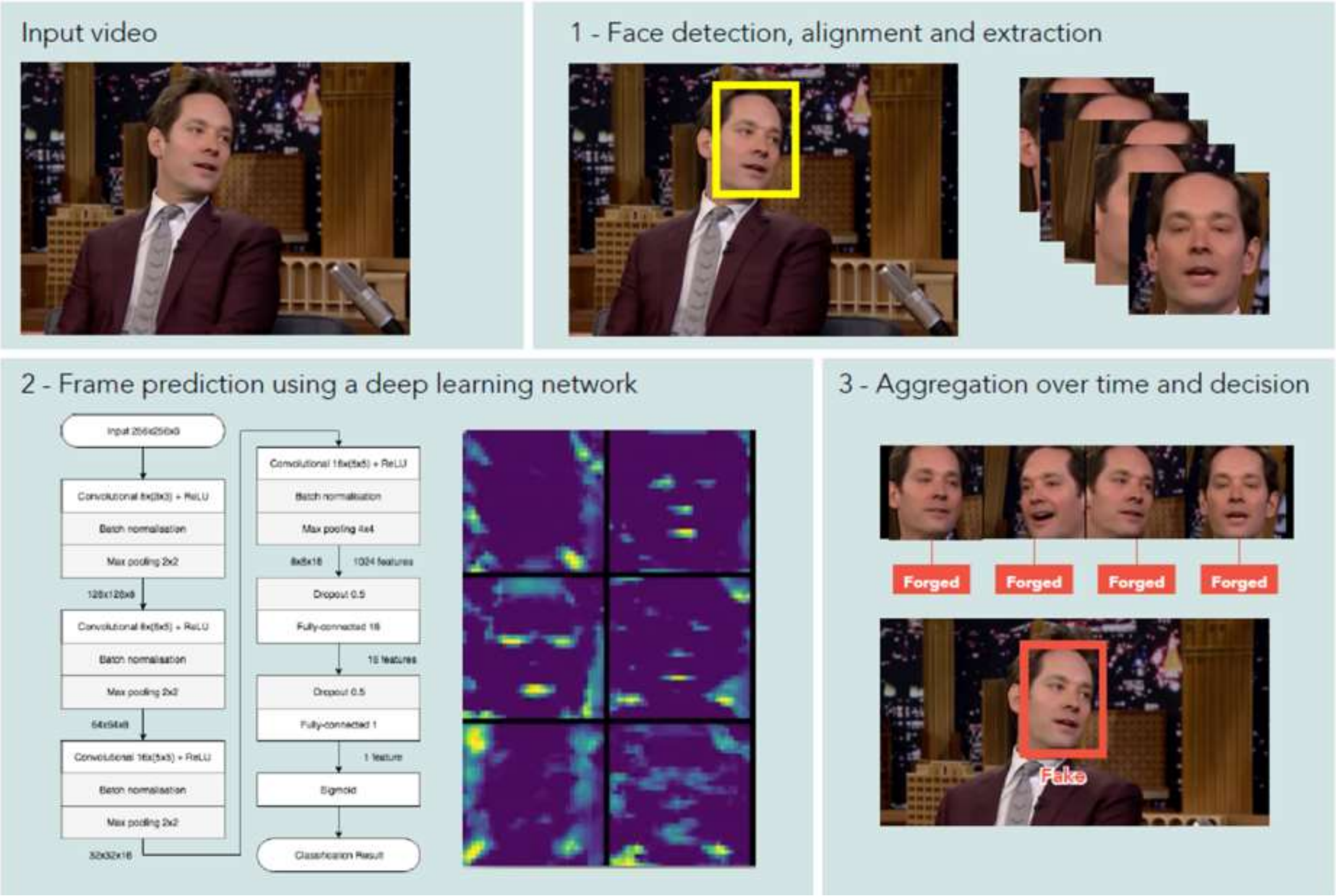


Faceswap – GAN[2]
Image: shaoanlu

Outline

- Introduction, Generating Fake Media Using Human-Related Information
- Methods for Generating Fake Media Based on Faces
- Methods for Detecting Fake Media Based on Faces
- Advanced Fake Media Generation and Detection Methods
- Toward Countering Infodemics (JST CREST FakeMedia, NII SynMedia Center)

Mesonet: simple, but the world first fake facial video detector

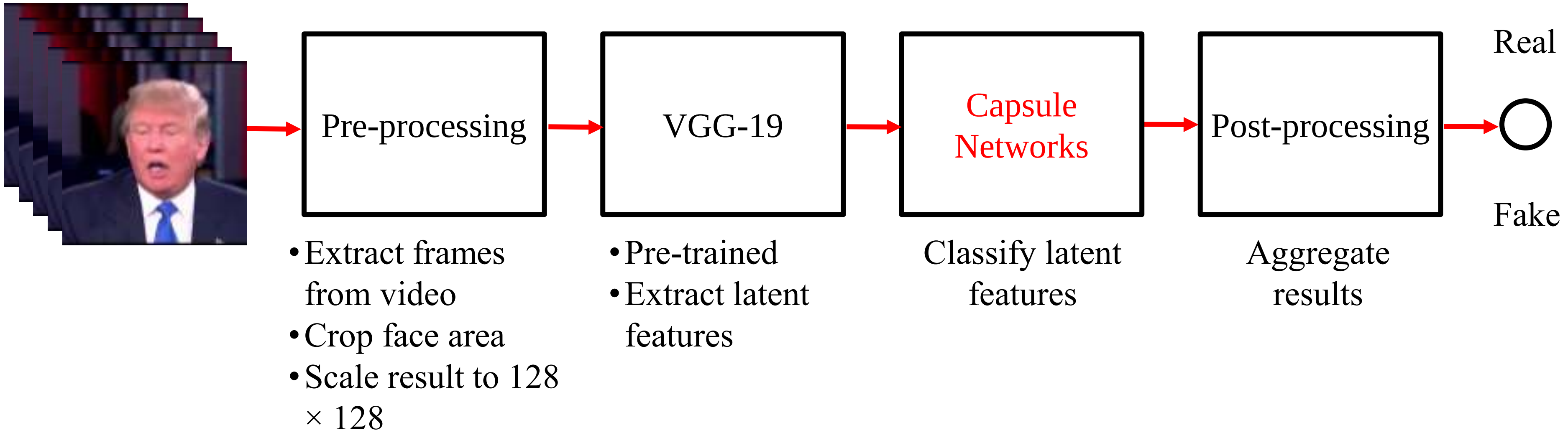


D. Afchar, V. Nozick, J. Yamagishi, and I. Echizen, " MesoNet: a Compact Facial Video Forgery Detection Network, " Proc. of the IEEE International Workshop on Information Forensics and Security (WIFS 2018), pp.1-7, December 2018 (number of citations: 704)

Fake Facial Video Detector using Capsule Networks

- Media forensics has become a timely and important topic due to significantly increased risks of realistic fake videos (deepfakes).

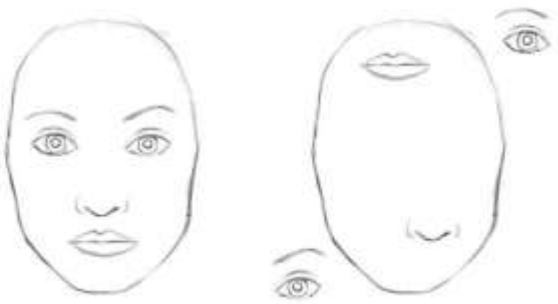
- Combine VGG19 with Capsule Network as a countermeasure



Huy H. Nguyen, Junichi Yamagishi, Isao Echizen, “Capsule-Forensics: Using Capsule Networks to Detect Forged Images and Videos” ICASSP 2019 (number of citations: 406)

Why capsule networks?

- **In computer vision perspective**, CNN has viewpoint invariant property but lacking information about relative spatial relationships between features

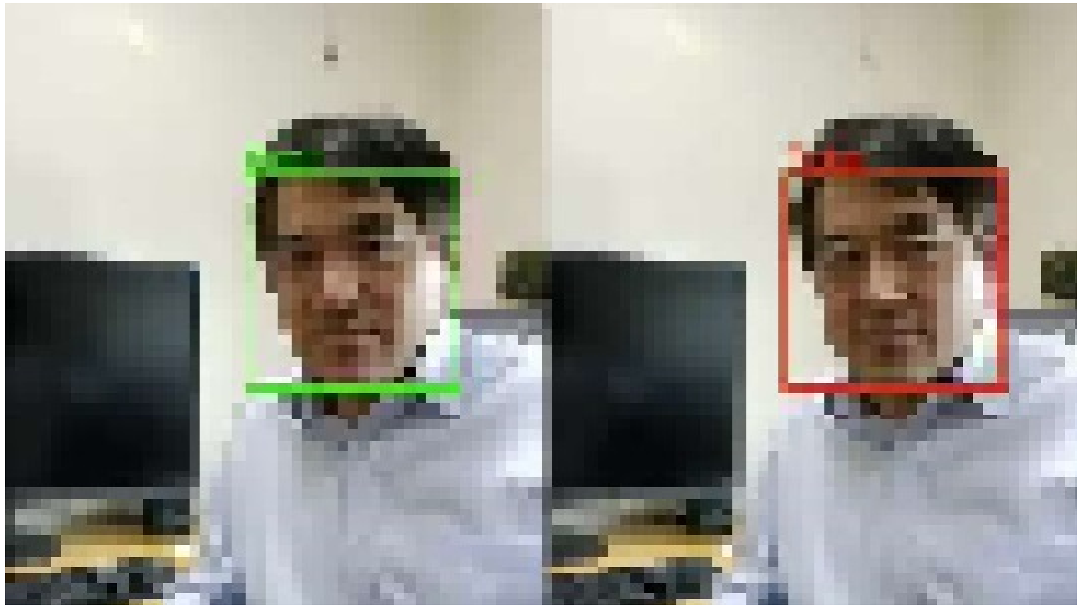


- Capsule networks have several capsules, each capsule is a CNN learning some specific representations (spoofing artifact or irregular noise in digital image forensics).

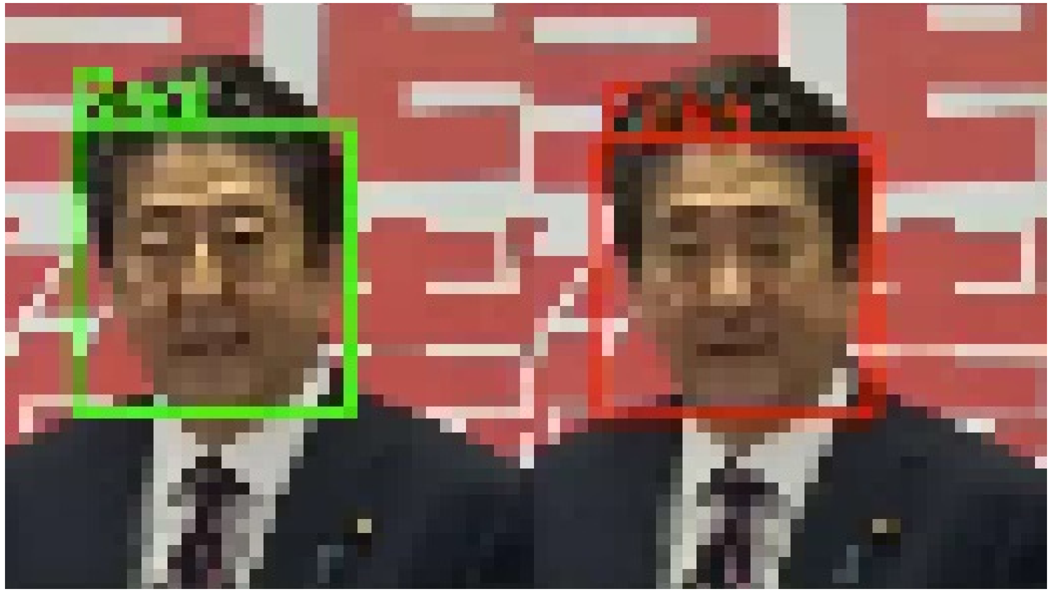
- The agreements between low-level capsules decide the activations of the high-level capsules.

Huy H. Nguyen, Junichi Yamagishi, Isao Echizen, "Capsule-Forensics: Using Capsule Networks to Detect Forged Images and Videos" ICASSP 2019 (number of citations: 406)

Detection results (Faceswap)



Swap faces using deepfake!

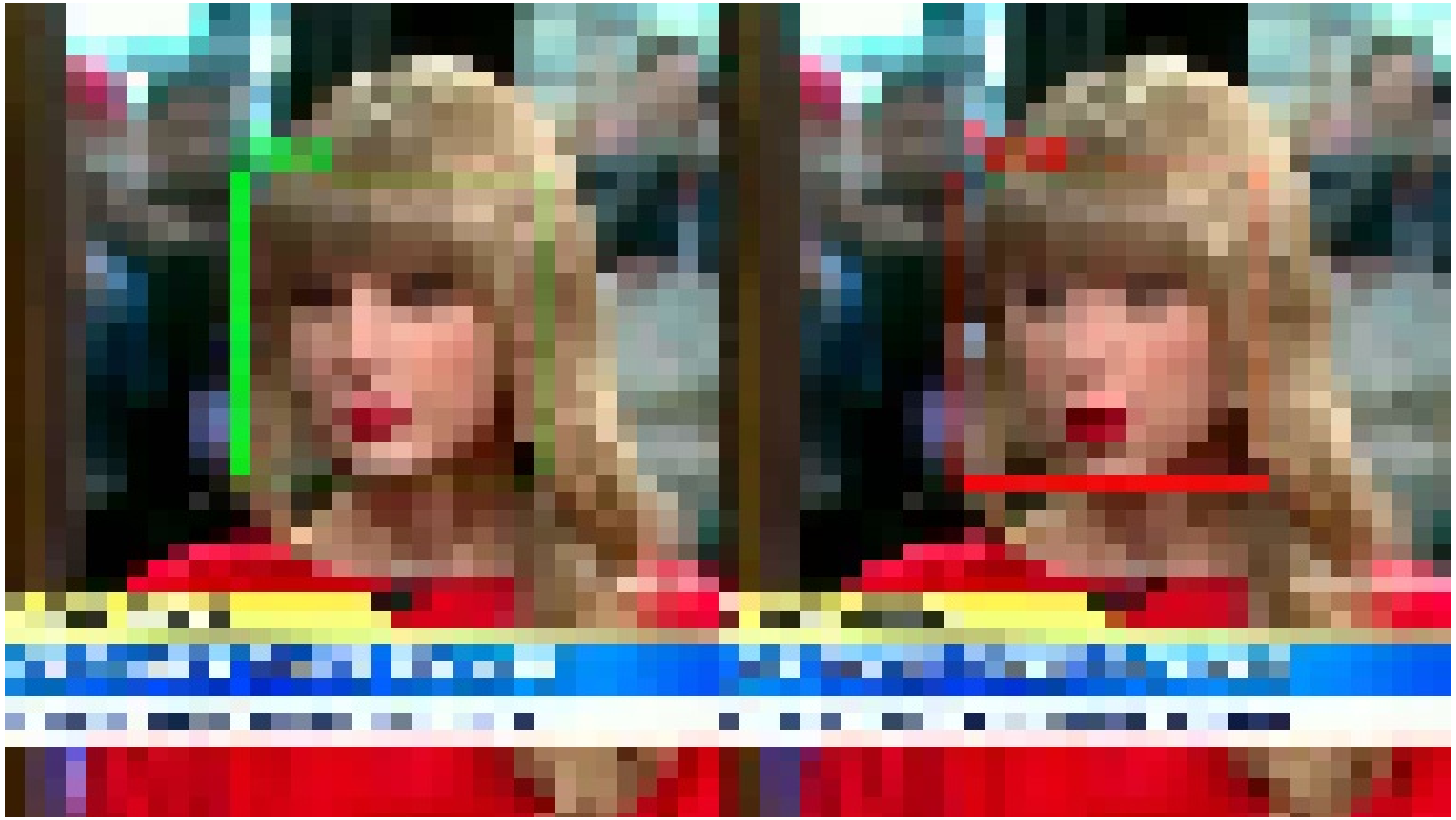


Our Deepfake dataset

	Real (frame)	Forged (frames)
Train	4,600	6,525
Dev	511	725
Eval	2,889	4,259

EER: 1.42%

Detection results (Face2Face)



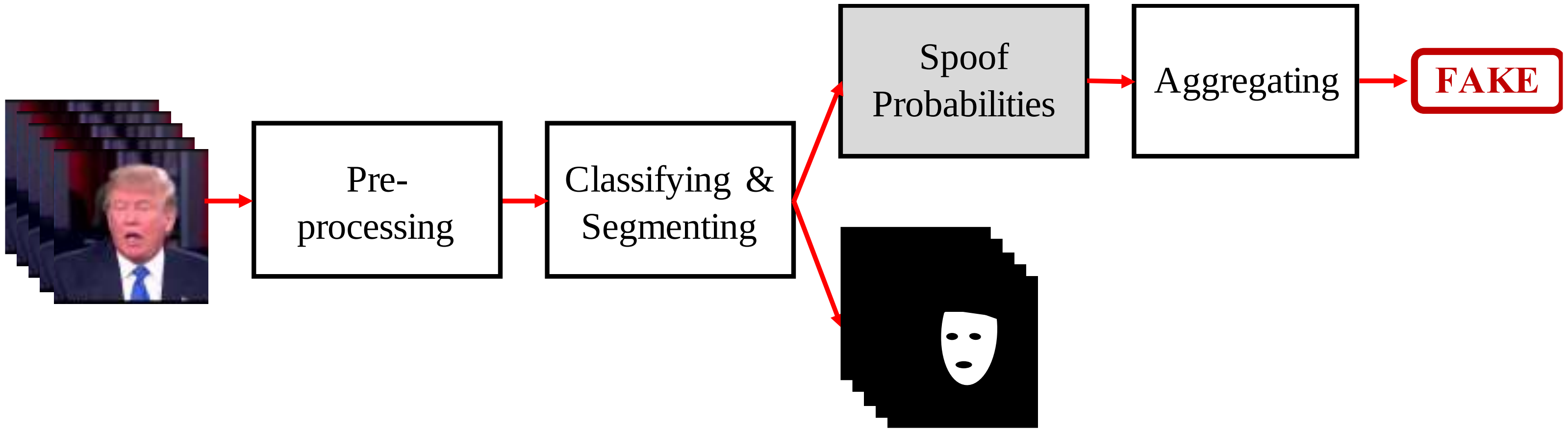
FaceForensics dataset

	Real (frame)	Forged (frames)
Train	7,040	7,040
Dev	1,500	1,500
Eval	1,500	1,500

EER
- No compression: 0.67%
- Light compression: 2.67%
- Strong compression: 17.0%

Joint Fake Facial Video Detection and Segmentation

- Multi-task learning: Combine **classification** task and **segmentation** task



- **Shape** of segmentation mask could reveal clue about **type** of **manipulation method.**

Real

Deepfakes
(rectangular mask)

Face2Face
(smooth mask)

FaceSwap
(polygon-like mask)

Huy H. Nguyen, Fuming Fang, Junichi Yamagishi, Isao Echizen, “Multi-task Learning For Detecting and Segmenting Manipulated Facial Images and Videos"Proc. of the BTAS 2019,8 pages, September 2019 (number of citations: 247)

Huy H. Nguyen, Fuming Fang, Junichi Yamagishi, Isao Echizen, “Multi-task Learning For Detecting and Segmenting Manipulated Facial Images and Videos"Proc. of the BTAS 2019,8 pages, September 2019 (number of citations: 247)

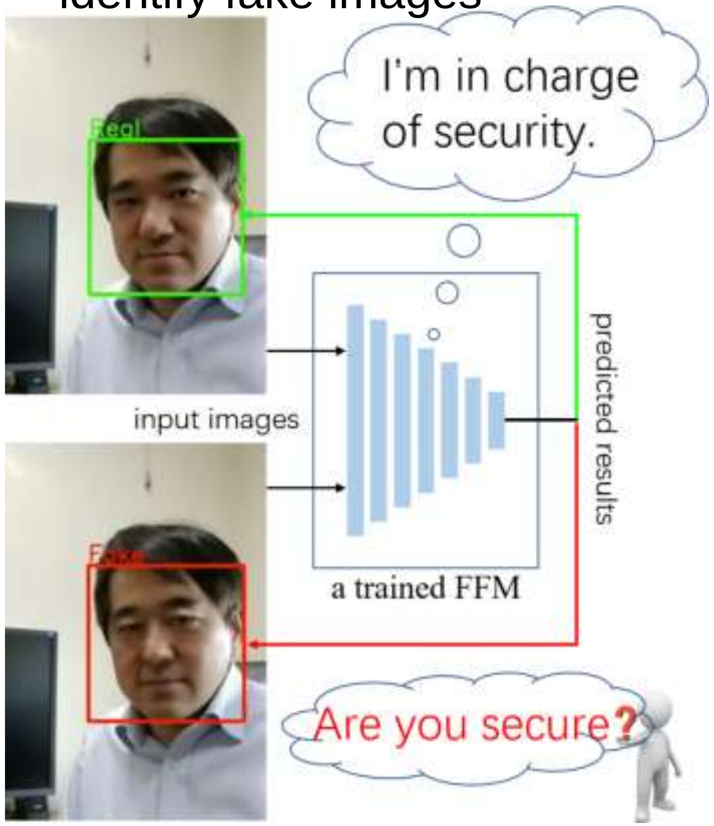
Outline

- Introduction, Generating Fake Media Using Human-Related Information
- Methods for Generating Fake Media Based on Faces
- Methods for Detecting Fake Media Based on Faces
- Advanced Fake Media Generation and Detection Methods
- Toward Countering Infodemics (JST CREST FakeMedia, NII SynMedia Center)

Adversarial attacks against fake video detection / segmentation

Background/Motivation

- DNN-based forgery forensics models (FFMs) are used to identify fake images

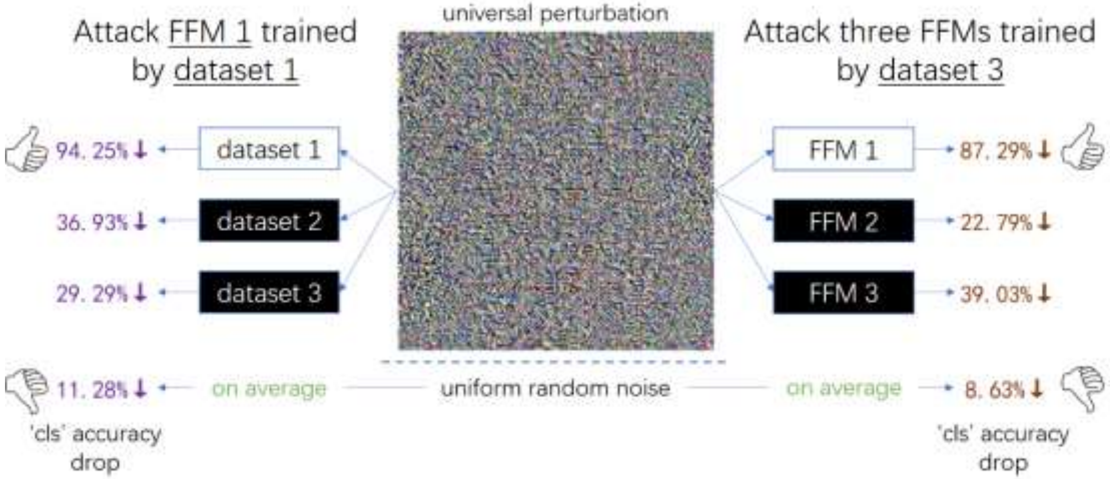


- Check security of three FFMs against adversarial attacks

Details

- Individual attack using gradient's information

- Universal attack based on over-firing

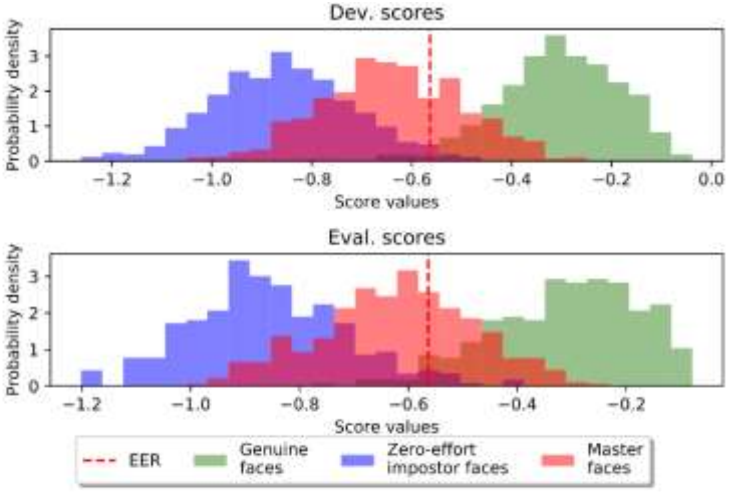


[1] R. Huang, F.M. Fang, H.H. Nguyen, J. Yamagishi, and I. Echizen, "Security of Facial Forensics Models Against Adversarial Attacks," Proc. of the IEEE International Conference on Image Processing (ICIP) 2020, 6 pages, (Oct. 2020)

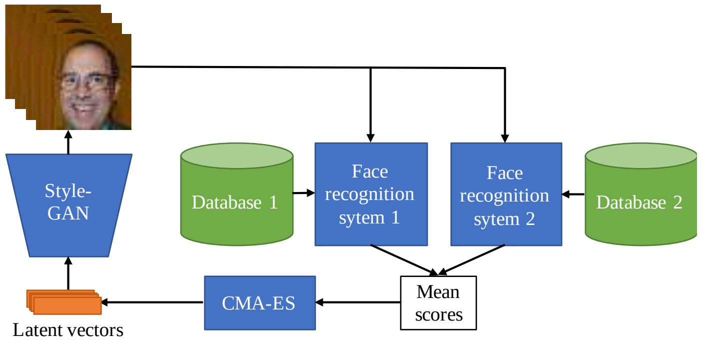
Master Face Attacks on Face Recognition Systems

Background

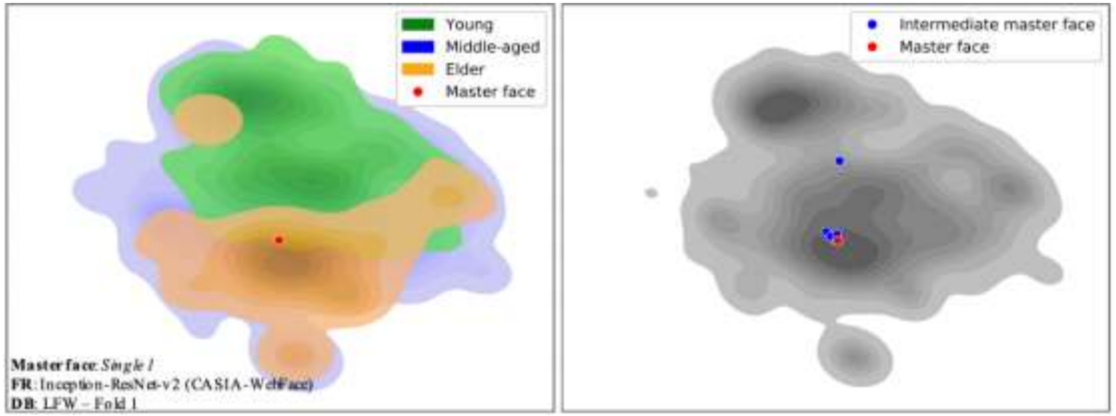
The **first work** to generate a **master face** (or a wolf face) which matches with multiple faces by a face recognition system.



Proposed Method



Results



Location of the master face in the latent space of a face recognition system

Master face

Master face and all matched faces with different genders, races, and appearances

H. H. Nguyen, J. Yamagishi, I. Echizen, and S. Marcel, "Master Face Attacks on Face Recognition Systems," IEEE Transactions on Biometrics, Identity and Behavior (IEEE TBIOM), 2022 .

OpenForensics: Multi-Face Forgery Detection and Segmentation In-The-Wild

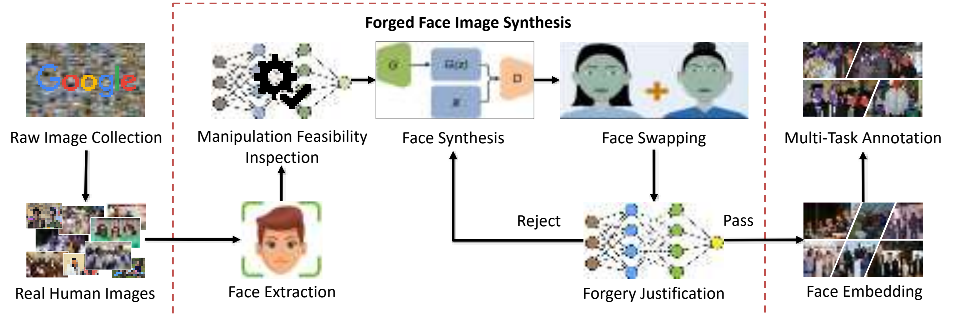
Background

- It is extremely difficult to point out forged faces among many faces in natural scenes.

Contributions

- Address new tasks of multi-face forgery detection and segmentation in-the-wild
- Present new dataset: 115k images with 334k faces
- Provide benchmark suite

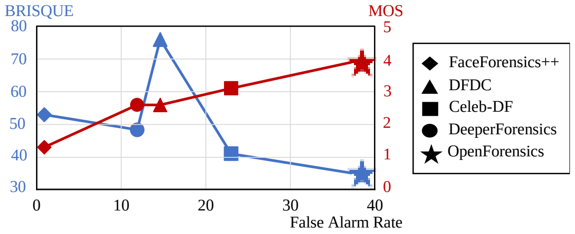
Dataset Generation

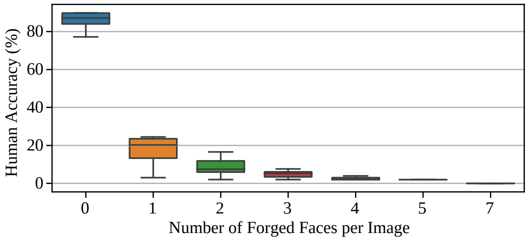


❖ **Test-Challenge set with data augmentation:**

User Study

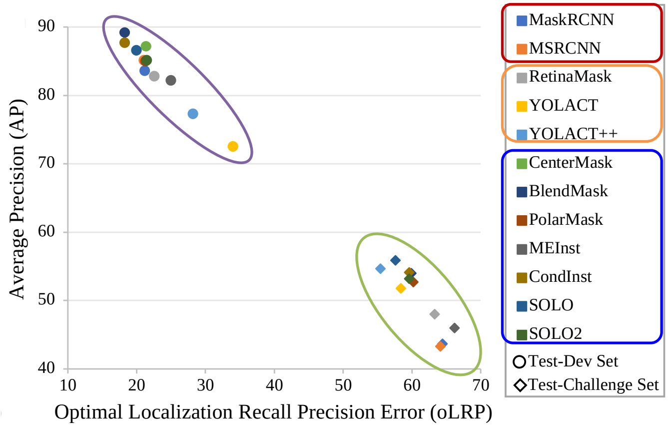
- 3,000 images (5 datasets) was used in experiments
- 200 participants (80 experts and 120 non-experts)





- OpenForensics can trick human (highest justification error) with highest realism
- More fake faces cause more missed detection

Benchmark



Le T.-N., Nguyen, H. H., Yamagishi, J., & Echizen, I., “OpenForensics: Large-Scale Challenging Dataset For Multi-Face Forgery Detection And Segmentation In-The-Wild”, International Conference on Computer Vision (ICCV), 2021 **(Core A*)**

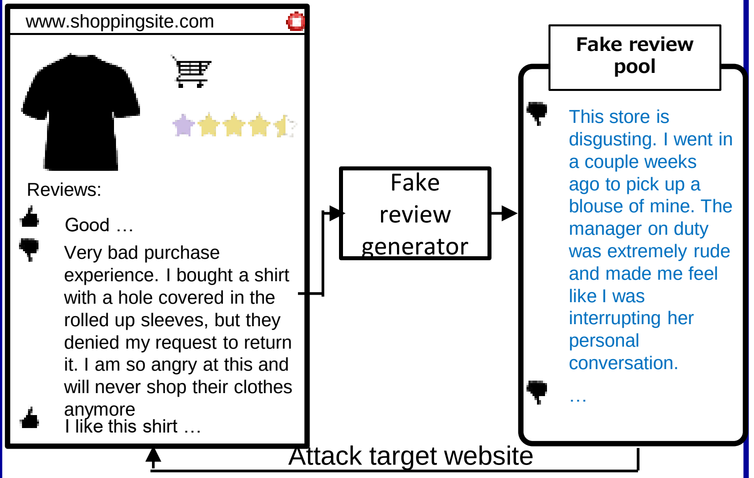
Computer Generated Fake Review Generation / Detection

Background/Aim

- High performance language models have been published
- These models could be used for fake review generation
- We show how natural review can be generated by the up-to-date language models
- We show how accurate the existing fake text detection methods have

Overview

Generation based on existing reviews



Details

Step 1: generation

x → GPT-2 (OpenAI) → x′

x: seed review (selected from website)
x′: generated review

Step 2: Variation

x′ → BERT (Google) → sentiment(**x**)?

Yes: accept
No: reject and discard

Accuracy improvement by fine-tuning with review database

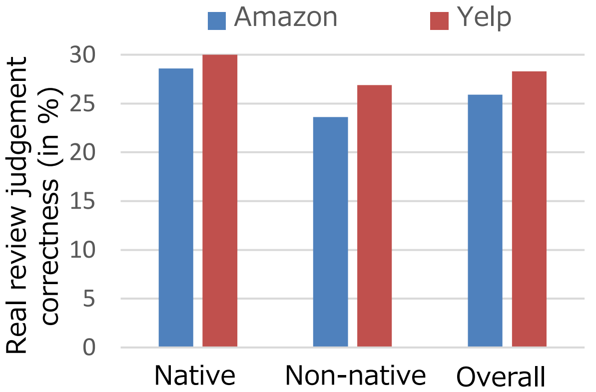
Review database → GPT-2 (OpenAI) → Fine-tuned GPT-2

Review database → BERT (Google) → Fine-tuned BERT

- Experiment

Human evaluation:

- 150 computer-generated and 50 real reviews
- 39 native/ 41 no native subjects
- Chance level: 25%



Detection (Fusion of 3 methods):
Grover(2019), GLTR(2019), and GPT-2PD/RoBerta(2019)
Equal Error Rates [%]

Detector	Amazon	Yelp	Overall
Grover	43.6%	36.9%	40.7%
GTLR	40.9%	35.9%	38.5%
GPT-2PD	**20.9%**	25.8%	23.5%
Grover + GTLR	35.3%	34.6%	34.9%
Grover + GPT-2PD	24.9%	22.2%	23.4%
GTLR + GPT-2PD	25.0%	**19.6%**	**22.5%**
Grover + GTLR + GPT-2PD	25.0%	**19.6%**	**22.5%**

- Feasible to detect the automatically generated reviews, but, not perfect
- Become one of evidences for OpenAI to release the largest GPT-2

[1] David Ifeoluwa Adelani, Haotian Mai, Fuming Fang, Huy H. Nguyen, Junichi Yamagishi, Isao Echizen, Generating Sentiment- Preserving Fake Online Reviews Using Neural Language Models and Their Human- and Machine-based Detection, AINA-2020 April 2020 (number of citations: 57)

Generation of biased news

Background/Aim

- High-performance language models are widely used for language generation tasks and these models are already being used to create fake news.
- An attacker can generate biased news to change political bias of their reader's.
- We show how biased news can be generated using GPT-2 and GROVER models.
- We show the generated news is fluent and the bias in them is clearly visible.

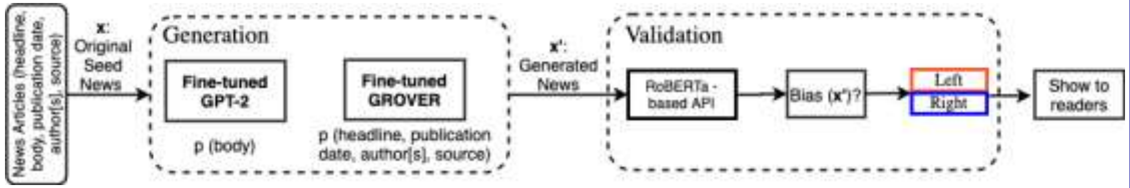
Overview

Threat model

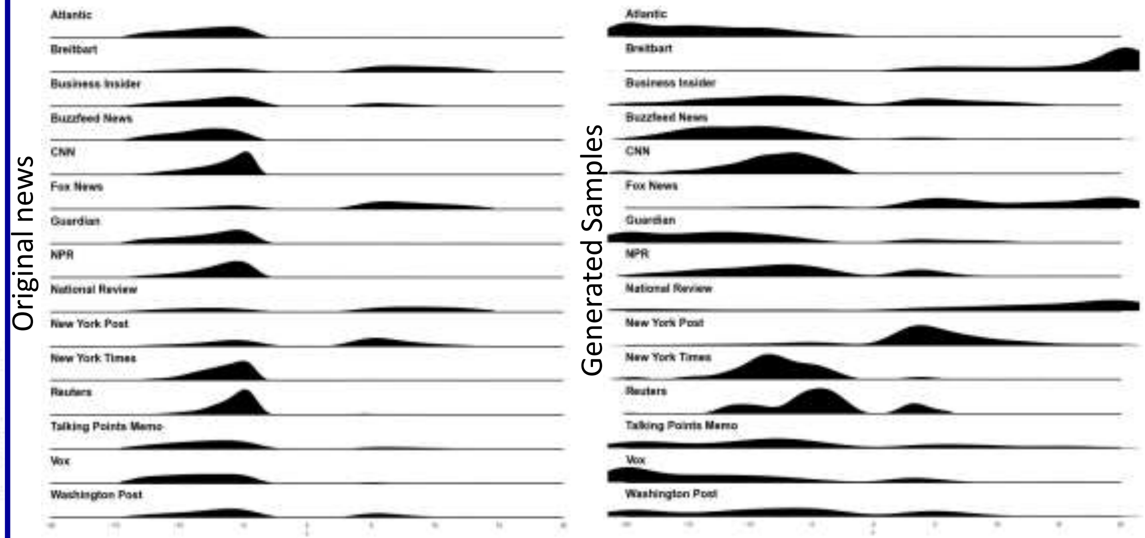
Original news is used as seed by the "Biased News Generator" to generate left or right biased news. Readers are then exposed to the generated biased news to change their original bias (either flip or increase).

Details

Generation Procedure



Bias Distribution



Subjective Fluency Evaluation

Real review subjective judgement with 80 participants correctness (in %)

Model	Native	Non Native	Overall
GPT-2	0.46 (16)	0.50 (23)	0.49 (39)
GROVER	0.43 (16)	0.48 (25)	0.46 (41)

Subjective Bias Evaluation

- Participants marked a clear bias 92% of the times.
- Participants chose bias correctly (between left and right) 63% of the times.

[1] Gupta, S., Nguyen, H. H., Yamagishi, J., & Echizen, I. (2020). Viable Threat on News Reading: Generating Biased News Using Natural Language Models. NLP+CSS Workshop at EMNLP 2020

Outline

- Introduction, Generating Fake Media Using Human-Related Information
- Methods for Generating Fake Media Based on Faces
- Methods for Detecting Fake Media Based on Faces
- Advanced Fake Media Generation and Detection Methods
- Toward Countering Infodemics (JST CREST FakeMedia, NII SynMedia Center)

Fake media (FM) and infodemics

https://www.who.int/emergencies/diseases/novel-coronavirus-2019/advice-for-public/myth-busters

- AI technology evolution and enhancement of computer resources
  - Learn a large amount of biometric information to generate fake media
    - Impersonate a corporate executive with fake voice and exploit cash (2019)
    - Participate in the Zoom conference by pretending to be Elon Musk with a fake face (2020)

- COVID-19 and infodemics
  - “Infodemics” of uncertain information cause anxiety and confusion in society
    - Fake news regarding preventive and therapeutic methods without scientific basis
    - Photographs of city scenes taken from a specific direction with a telephoto-lens camera that gave the impression of a crowded area.

https://nyheder.tv2.dk/samfund/2020-04-26-hvor-taet-er-folk-paa-hinanden-disse-billeder-er-taget-samtidig-men-viser-to

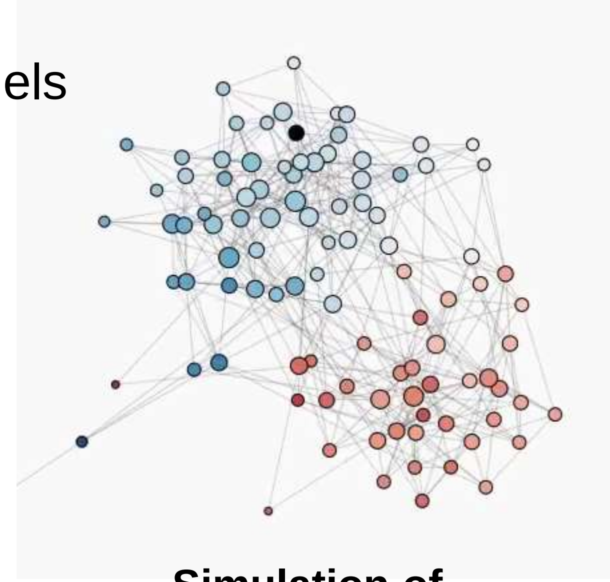
- Attackers use AI to generate fake media and then spread them on social media to create an infodemic
  - Fake media in a broad sense: Deepfake, adversarial examples, and propaganda
  - Intentional occurrence of infodemic and thought guidance of the masses
  - Attack on a specific individual by spreading hoaxes

Social information technologies to counter infodemics (JST CREST, Dec 2020- Mar 2026)

CREST FakeMedia

Toward healthy human-centered cyber society：dealing with various fake media (FM) & decision support

- Advanced FM detection technologies
  - Provide information to users in a format that explains not only FM detection but also the target to be deceived (i.e., persons or AI technology)
- FM detoxification technologies
  - Use detoxified FM as normal media for learning data of machine learning models
- Information technologies that counter infodemics and support diverse decision-making
  - Echo chamber suppression & incorporation of various reliable info. by FM detection / detoxification
- ELSI
  - No law that directly punishes fake media generation
  - "Transparency" is important as to how the platform identifies fake media

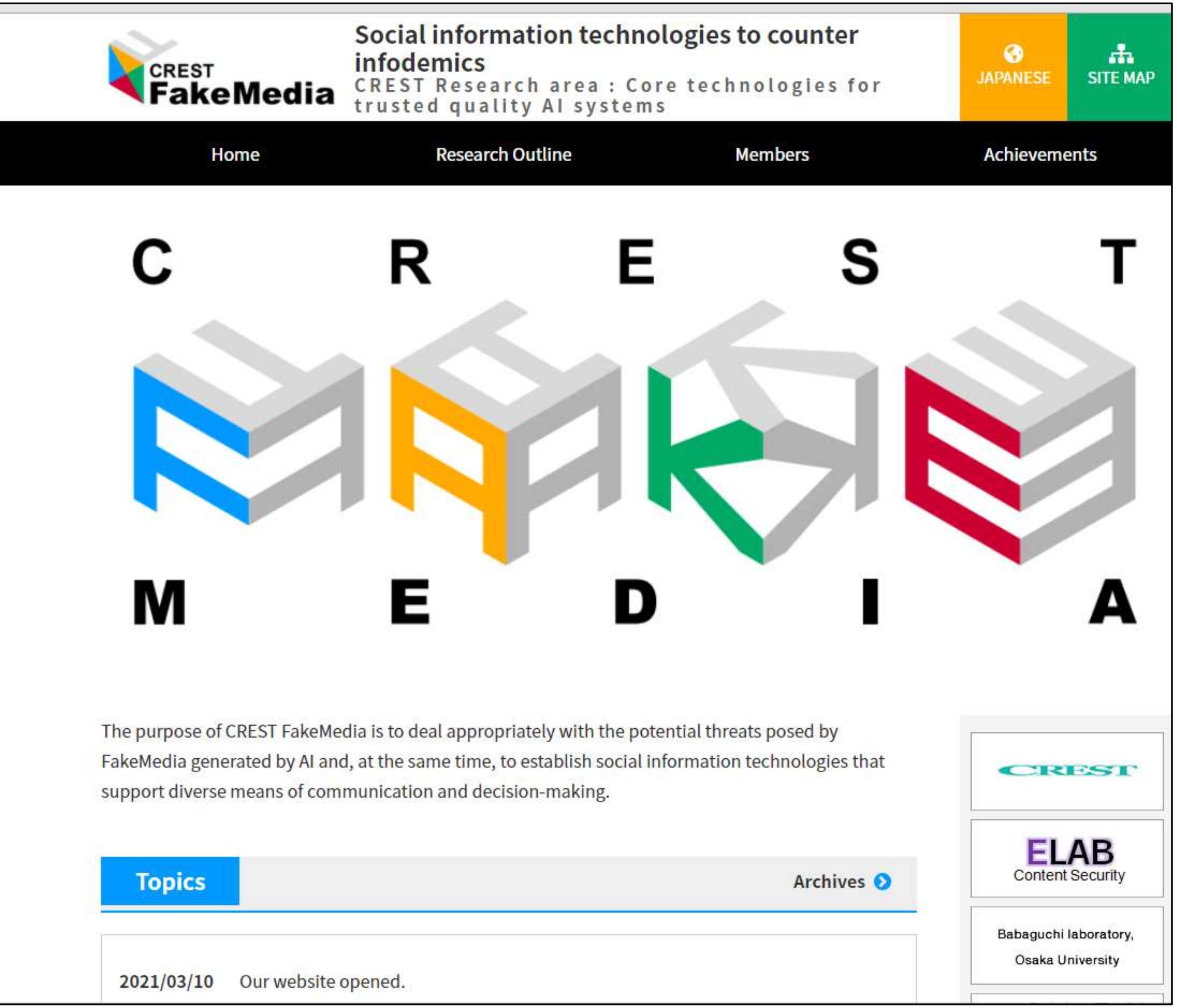


Simulation of echo chamber generation [1]

[1] K. Sasahara, W. Chen, H. Peng, G. L. Ciampaglia, A. Flammini and F. Menczer, Social Influence and Unfollowing Accelerate the Emergence of Echo Chambers, Journal of Computational Social Science, 2020

Website of CREST FakeMedia

Proactively disclose preprints, programs, and datasets

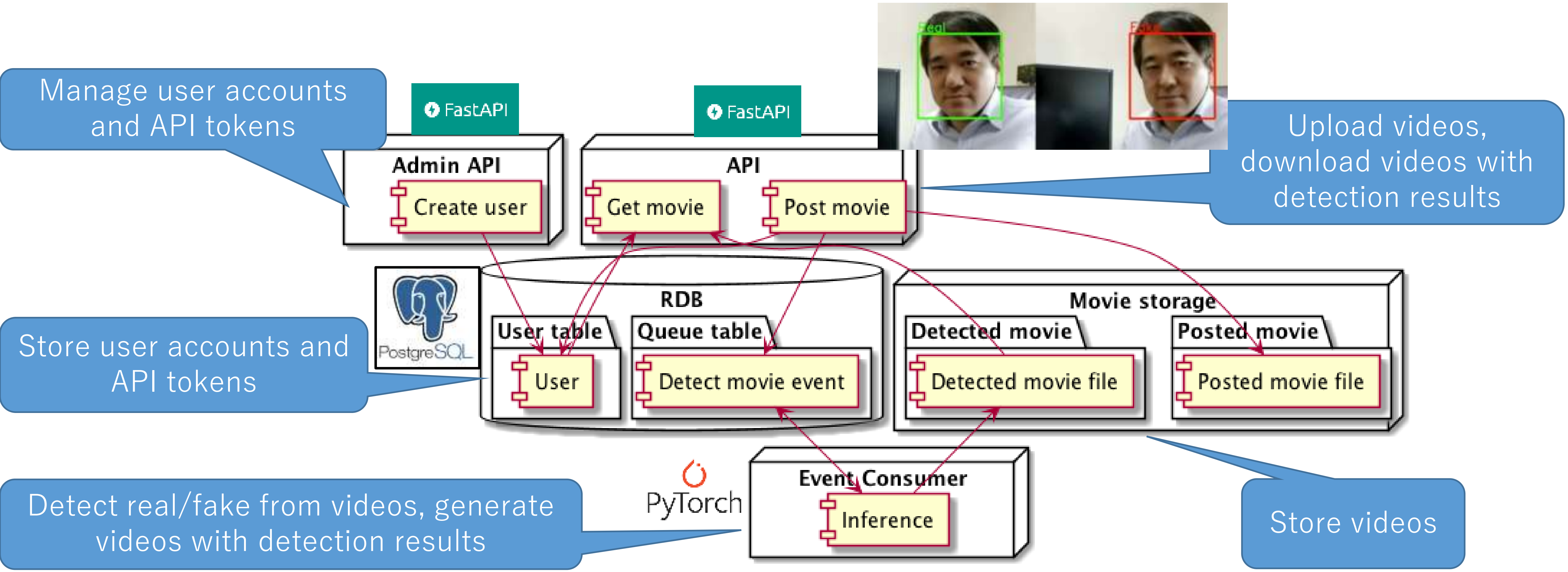


Refereed conference papers

1. Y. Yamasaki, M. Kuribayashi, N. Funabiki, H. Nguyen, and I. Echizen, "A Study of Feature Extraction Based on Denoising Auto Encoder for Classification of Adversarial Examples," APSIPA ASC 2021, December 2021
2. MaungMaung AprilPyone, Hitoshi Kiya, "A Protection Method of Trained CNN Model Using Feature Maps Transformed With Secret Key From Unauthorized Access", APSIPA ASC 2021, December 2021, Preprint
3. Dilrukshi Gamage, Jiayu Chen, and Kazutoshi Sasahara, "The Emergence of Deepfakes and its Societal Implications: A Systematic Review", Conference for Truth and Trust Online, October 2021
4. Sosuke Nishikawa, Ikuya Yamada, Yoshimasa Tsuruoka, Isao Echizen, "A Multilingual Bag-of-Entities Model forZero-Shot Cross-Lingual Text Classification", ACL-IJCNLP 2021 Student Research Workshop (non-archival option), 2021, Link
5. Liangzhi Li, Bowen Wang, Manisha Verma, Yuta Nakashima, Ryo Kawasaki, Hajime Nagahara, "SCOUTER: Slot Attention-based Classifier for Explainable Image Recognition" ICCV 2021, accepted, October 2021, Preprint, code
6. Trung-Nghia Le, Huy H. Nguyen, Junichi Yamagishi, Isao Echizen, "OpenForensics: Large-Scale Challenging Dataset For Multi-Face Forgery Detection And Segmentation In-The-Wild" ICCV 2021, accepted, October 2021, Preprint, presentation video, dataset
7. April Pyone MAUNG MAUNG, Hitoshi KIYA, "TRANSFER LEARNING-BASED MODEL PROTECTION WITH SECRET KEY,", IEEE International Conference on Image Processing, accepted, September 2021
8. Canasai Kruengkrai, Xin Wang, Junichi Yamagishi, "A Multi-Level Attention Model for Evidence-Based Fact Checking", Findings of ACL2021, accepted, August 2021, Preprint, code

AIaaS for automatic detection of fake facial videos – SYNTHETIQ: Synthetic video detector –

- All processes from uploading the video to downloading the video with the detection results can be used as a Web API.
- Easy realization of AI-based web service "AI as a service" by utilizing web API



Global Research Center for Synthetic Media, NII

SynMedia Center

Promote the generation of various media, work to ensure the reliability of media, and conduct research and development for decision-making as an international base for addressing real-world issues.

CREST (Prof. Yamagishi)

VoicePersonae: Speaker identity cloning and protection

Modeling, utilization, and protection of speaker identities by integrating speech synthesis, speech conversion, and speech enhancement

CREST (Prof. Echizen)

CREST FakeMedia

Social information technologies to counter infodemics

Supporting appropriate responses to fake media threats and diverse communication

Global Research Center for Synthetic Media

Synthetic media generation

Speech /image / video processing; natural language processing; computer vision processing

Fake media detection

Digital forensics; information security; privacy protection

Media reliability, Decision-making support

Computational social science; ethical, legal and social implications

Prof. Babaguchi

(Osaka U.)

Prof. Kiya

(TMU)

Prof. Sasahara

(Tokyo Tech.)

Prof. Mizuno

(NII)

EURECOM · Avignon Université · idiap · TUM · universität wien · IIT JAMMU · INDIANA UNIVERSITY

Promote real-world applications through creation of new science and technology fields and research trends, collaborate with domestic and overseas academic institutions, and participate in industry-academia-government collaboration

Website of SynMedia Center

SynMedia Center

Global Research Center for Synthetic Media, National Institute of Informatics

JAPANESE | SITE MAP

Home | About SynMedia Center | Members | Achievements

The Global Research Center for Synthetic Media (SynMedia Center) conducts research and development across various modalities such as face, voice, body, and natural language to generate synthetic media, detect synthetic media (fake media) generated for improper purposes, ensure media reliability, and support decision-making. Our mission is to promote the SynMedia Center as an international base for addressing issues in the real world.

NII National Institute of Informatics

CREST FakeMedia

ELAB Multimedia Security

Topics | Archives

Home > About SynMedia Center

About SynMedia Center

With the evolution of AI technology and the enrichment of computer resources stemming from the ability to acquire a large amount of human-related information such as face, voice, body, and natural language, it is becoming possible to generate synthetic media that can be mistaken for the real thing. Synthetic media is expected to be used in various fields such as communication (e.g., virtual avatars) and entertainment (e.g., rakugo speech synthesis), and it is expected that high-quality synthetic media generation technology will be established. Unfortunately, there is a negative side of synthetic media—attackers can generate and distribute fake videos, fake audio clips, and fake documents for the purposes of fraud, thought control, and public opinion manipulation, and this has become a social problem.

To achieve a healthy human-centered cyber society, the Global Research Center for Synthetic Media (SynMedia Center) conducts research and development across various modalities such as face, voice, body, and natural language to generate synthetic media, detect synthetic media (fake media) generated for improper purposes, ensure media reliability, and support decision-making. Our mission is to promote the SynMedia Center as an international base for addressing issues in the real world.

To achieve a healthy human-centered cyber society, we will promote the generation of various media, work to ensure the reliability of media, and conduct research and development for decision-making as an international base for addressing real-world issues.

Pursuing the utilization and reliability of various media through an interdisciplinary and international system

Promote real-world applications through creation of new science and technology fields and research trends, collaborate with domestic and overseas academic institutions, and participate in industry-academia-government collaboration

Thank you.