NIST Open Media Forensics Challenge
OpenMFC 2022 Workshop

# OpenMFC 2022 Evaluation Program

**Haiying Guan***

Baptiste Chocot[+], Ilia Ghorbanian[-], Lukas Diduch[^], Yooyoung Lee[$], and Christopher Tu[#]

Multimodal Information Group, Information Access Division, ITL, NIST

[*] Computer Scientist at NIST
[+] Foreign Guest Researcher at NIST
[-] Student in the Professional Research Experience Program at NIST
[^] Software Engineer at Dakota Consulting, Inc. / Contractor at NIST
[$] Supervisory Computer Scientist at NIST
[#] Student in the Student Volunteer Program at NIST

# Disclaimer

- Certain commercial equipment, instruments, software, or materials are identified in this article in order to specify the experimental procedure adequately. Such identification is not intended to imply recommendation or endorsement by NIST, nor is it intended to imply that the equipment, instruments, software or materials are necessarily the best available for the purpose.
- The views, opinions and/or findings expressed are those of the author and should not be interpreted as representing the official views or policies of the Department of Defense or the U.S. Government.
- If not specified in the figure title, all images, graphs, and charts are original works created for NIST OpenMFC program or DARPA MediFor Program under NIST IRB #ITL- 0018.

# Acknowledgement


- External collaborators:
  - Prof. Siwei Lyu, Dr. Shan Jia, and Yan Jun at University at Buffalo
  - Prof. Conrad Sanderson
  - U.C. Denver and PAR Government in the MediFor Program
  - Prof. Jennifer Newman, Li Lin, Prof. Yong Guan from Iowa State University, and Prof. Roy A. Maxion from Carnegie Mellon University
- NIST contributors
  - Edmond Golden
  - Noah Dove
  - Jonathan Fiscus
  - Timothee Kheyrkhah
  - Peter Fontana
  - Jesse G. Zhang
  - Daniel Zhou

# OpenMFC 2022 Outline

- What's new? 

- A Comprehensive review
  - OpenMFC program overview
  - OpenMFC evaluation design and 2022 tasks
  - OpenMFC 2022 datasets
  - OpenMFC 2022 metrics, and results

# New! What's new (1)?

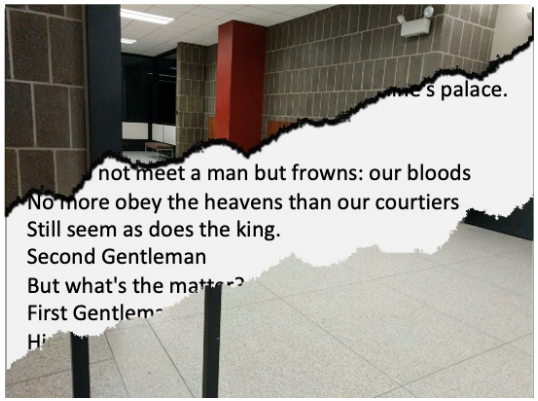
- Reorganized all evaluation tasks
  - Focusing on the deepfake evaluation
- Deepfake evaluation
  - Establishing the new evaluation
  - Generating the new deepfake datasets

(a) Original video (b) Donor video (c) Deepfaked video

Figure: A deepfaked video example using the VidTimit videos and the DeepFaceLab tool.

# New! What's new (2)?

- Two new evaluation tasks
  - Image Splice Manipulation Detection (ISMD)
    - Designed for beginners
    - Dataset:  OpenMFC22_ImageSplice_MD
    - 2000 images, only focusing on splice image detection
  - Steganography Image Detection (StegD)
    - Dataset: OpenMFC22_Image_StegD
    - About 400+ images, focusing on stego image detection
- All OpenMFC evaluation datasets are in the new format
- Updated data release and OpenMFC Registration forms
- New data resource website
- Updated evaluation leaderboards



# What is OpenMFC?

## - Open Media Forensics Challenge

# Media Forensics

- "Media Forensic is scientific study into the collection, analysis, interpretation, and presentation of audio, video, and image evidence obtained during the course of investigations and litigious proceedings" [1]

a. Test Image b. Original Image c. Manipulated Mask

Figure : An Example of Media Forensics

[1]https://artsandmedia.ucdenver.edu/areas-of-study/national-center-for-media-forensics/about-the-national-center-for-media-forensics

# Deepfakes (1)

- Deepfakes are photorealistic images and videos built using GAN (Generative Adversarial Network) techniques originated within the Artificial Intelligence (AI) domain starting from late 2017

  Example: You Won't Believe What Obama Says In This Video! by BuzzFeedVideo[1]
  (manipulated by Jordan Peele using Adobe After Effects and AI face-swapping tool, FakeApp.)

[1] https://www.youtube.com/watch?v=cQ54GDm1eL0

# Deepfakes (2)

- Demo:
- https://thispersondoesnotexist.com/

# Deepfakes (3)

- Generative Adversarial Network (GAN): conceptual illustration

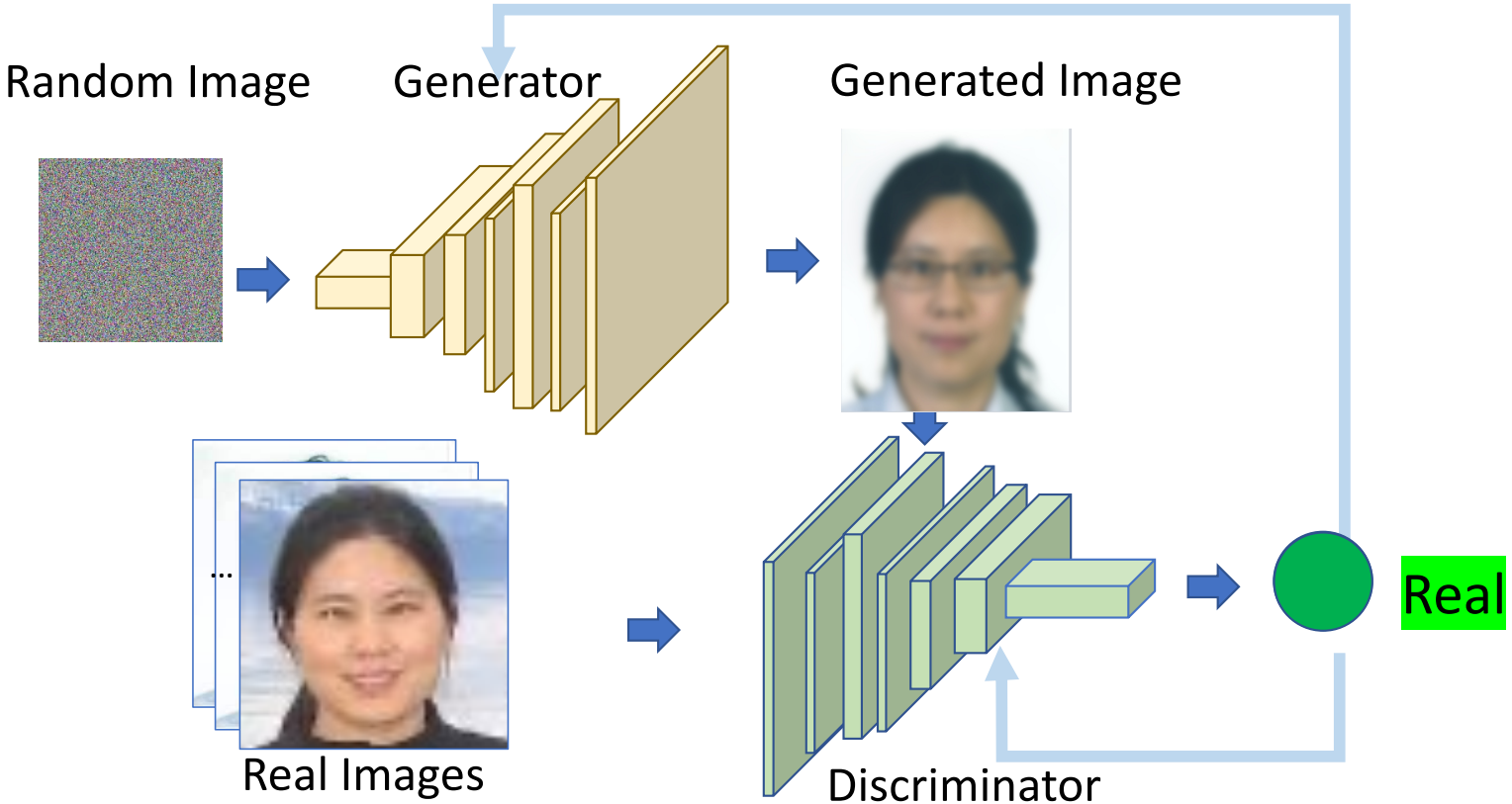



Figure: How Generative Adversarial Network (GAN) works

# OpenMFC Program

- OpenMFC is an evaluation series open to public participants worldwide to support media forensics research and help advance the state-of-the-art imagery (image and video) forensics technologies
  - Online leaderboard benchmark evaluation (derived and updated from DARPA MediFor)
  - Open publicly to all researchers worldwide
  - Free participation (NIST does not provide funds to participants)
  - Free data and resources (with signed agreements)
  - Benchmark evaluation (due to the unbiased, neutral position of our team)
  - A platform for collaboration

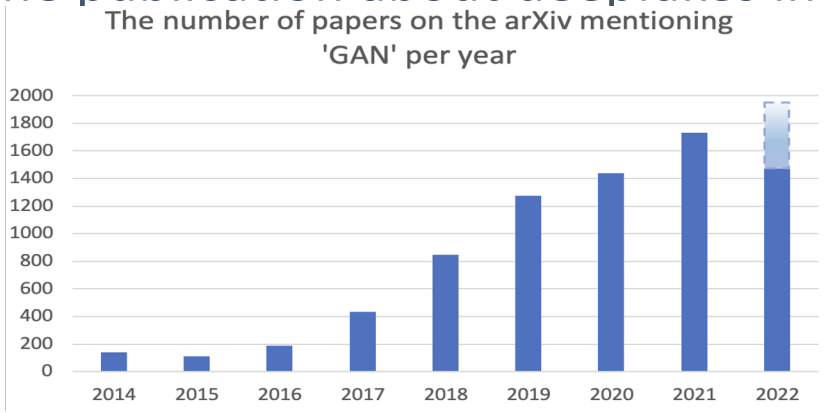
# Why OpenMFC?

# Why work on media forensics?

- "Seeing is not believing"
  - Hundreds of media editing tools (Adobe CC, GIMP, Corel Paintshop Pro, Skylum Luminar, DxO PhotoLab, ON1 Photo RAW, ACDSee Photo Studio Ultimate, Pixlr Editor, Canva, PicMonkey, Snappa, PortraitPro, Fotor, …)
  - CGI, and anti-forensics techniques
- Applications
  - Fake news detection in social media platform
    - Facebook, Twitter, Instagram, Snapchat, and Google
  - Academic misconduct
  - Criminal law and private investigation
  - Geospatial intelligence
  - Cybersecurity
  - Disinformation

# Why focus on deepfakes?

- Malicious media generated by deepfake tools are major threats in impersonation and disinformation
- Deepfakes become ubiquitous, spread across domains such as social media, politics, and cybersecurity, and erode trust. It is quickly becoming one of the most worrying applications of AI for crime or terrorism
- The publication about deepfakes increases in arXiv



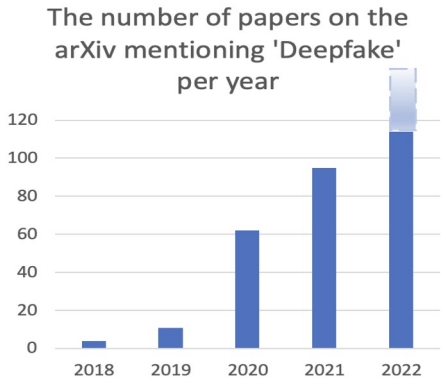




Figure: The number of papers on the arXiv mentioning 'GAN' or 'deepfake' per year as 09/15/2022

Example:
A deepfake of Ukrainian President [1] [2]

[1]https://twitter.com/MikaelThalen/status/1504123674516885507?ref_src=twsrc%5Etfw%7Ctwcamp%5Etweetembed%7Ctwterm%5E1504123674516885507%7Ctwgr%5E%7Ctwcon%5Es1_&ref_url=https%3A%2F%2Fwww.npr.org%2F2022%2F03%2F16%2F1087062648%2Fdeepfake-video-zelenskyy-experts-war-manipulation-ukraine-Russia

[2]https://www.npr.org/2022/03/16/1087062648/deepfake-video-zelenskyy-experts-war-manipulation-ukraine-russia

# Why OpenMFC Evaluation Series?

- Benchmark evaluation with performance report tested against variability of realistic scenarios
- Arms Race: between the media forensics and deepfakes with anti-forensics
- Data Drift: the old detection systems fail to capture advancements of the new deepfake tools
- One-shot vs. recurring evaluation: A series of recurring evaluations are more desirable than a one-shot evaluation

# OpenMFC Objective and Strategy

- Objective
  - Stimulate the research in media forensics
  - Understand the state-of-the-art
  - Provide continuous cross-year comparisons reports
  - Help researchers to improve their system with system performance analysis report
  - Bridge the gap between laboratory and field evaluation
- Strategy
  - Collaboration instead of competition: nature of the media forensics
  - Group the media forensics researchers and build a strong connected community

# How to design the OpenMFC?

# Evaluation Design Challenges

- What to evaluate?
  - Not too easy and not too hard
  - To evaluate a comprehensive system brings challenges in designing a unified, flexible evaluation framework
- What resource to use?
  - Lack of benchmark datasets: human post-annotation doesn't work well
  - Different media forensics applications/technologies need different evaluation data
  - Hundreds if not thousands of factors (e.g. manipulation methods)
  - Problems of using existing data for the algorithm/system development
- How to evaluate?
  - How to handle the continuous, dynamic changes of technologies (e.g. forensic vs. anti-forensic)?
  - How to adapt to the new technologies (e.g. deepfakes)?
  - How to evaluate intelligent system or semantic system?
- What we can get from the evaluation?
  - Baseline performance, state-of-the-art, and cross-year performance comparison

# Media Forensics Analytics Technologies: Incomplete Survey

- Copy Move
- Geometric-based Cropping Detector
- Lighting
  - Gradient-Based Illumination Description
  - Light inconsistency on faces
- Face-based tamper detection
  - Facial Expression
- Pixel-based tamper detection
  - JPEG Compression Detection
  - JPEG Dimples
  - Noiseprint
  - Resampling Anomaly
- Color Phenomenology
- Holistic approaches
- Splice detection
- GAN/Deepfakes/AI-Synthesized detection
- Inconsistency detection
  - Audio visual speaker identity inconsistency
  - Audio visual lip out of synchrony
  - Audio visual scene inconsistency detector
  - Light inconsistency
  - Weather / location
  - Codec inconsistency
  - Noise inconsistency
- Video
  - Frame duplication
  - Frame drop
- ENF (Electric Network Frequency)
- Rebroadcast
- Camera verification
- Provenance filtering and graph building

# Media Forensics Evaluation Challenges

- What is the difference between media forensics and other research topics?
  - Intrinsic property
    - A comprehensive system instead of a single technology
  - Open topics: manipulations emerging and change all the time
    - Fast evolution
    - Dynamic upgraded
  - Forensics vs. Anti-forensics
    - No traditional steady improvement curve
    - Prediction: who is going to win in the end?
  - Semantics instead of facts

# Evaluation Task Design Strategy

single input (detection) → pair inputs (verification) → multiple input (association)

## Single File Authenticity

Manipulation Detection:
Is the image/video manipulated?

Localization:
Where is the image/video manipulated?
- Spatial
- Temporal
- Temporal-spatial

## Authenticity in Context

### Image Pair Authenticity

Splice Detection:
Does image1 contain some of image2?

Localization:
- Where in image1 was image2 content spliced?
- Where in image2 is the splice donor?

### Image+ Image Collection

Provenance Filtering:
Find related images

Provenance Graph Building:
Construct a phylogeny graph of related images

### File+Camera

Camera Verification:
Was an image/video taken by a known camera?

### File+Event

Event Verification:
Was an image capture during a known event?

# New! OpenMFC22 Evaluation Tasks

- Manipulation Detection (MD)
  - Imagery: Image or Video
  - Detection: to detect if the media has been manipulated
  - [Optional] Localization: to spatially localize the manipulated region
  - Subtasks: Image_MD (IMD), **ImageSplice_MD (ISMD)**, Video_MD (VMD)
- Deepfake Detection (DD)
  - Imagery: Image or Video
  - To detect whether an image/video was GAN manipulated or deepfaked.
  - Subtasks: Image_DD (IDD), Video_DD (VDD)
- **Steganography Image Detection (StegD)**
  - Imagery: Image
  - To detect whether an image was a steganography image

# Image Manipulation Detection and Localization

## System Input

Image(s) + (Metadata)

Probe image

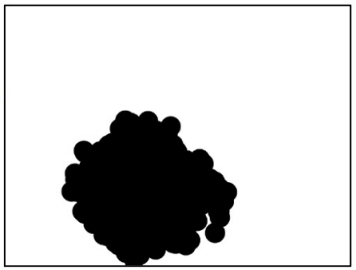
**Image Detection and Localization Analytic System**

## System Output

Confidence score

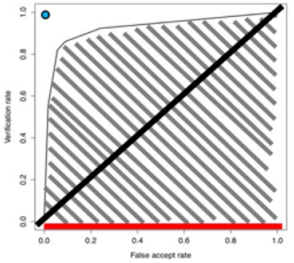
97.86

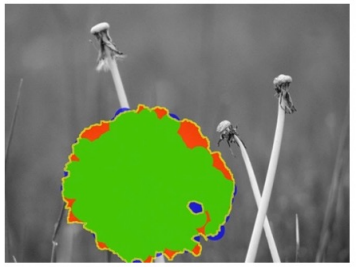
System output mask



## Metrics

Receiver Operating Characteristic (ROC)
Area Under the Curve (AUC)
Correct Detection (CD) at False Alarm Rate 5%



Manipulated image
Matthews Correlation Coefficient (MCC)

# What data to use?

# Evaluation Dataset Design Challenges

- Evaluation objective
  - Lab algorithm evaluation vs. in-the-field evaluation (real-world applications)
- Media forensics analytics is an open task:
  - Emerging software and tools (GAN, Deepfakes, CGI, etc.)
  - Anti-forensics
- Curse of dimensionality
  - Large testing space
    - Media space (image/video/audio, camera/scanner)
    - Manipulation space (manipulator, manipulation operations and software)
    - Anti-forensic technology space
  - The combinatorics of one dimension
    - Suppose a 2-Factorial, single operation experimental design
      - 17,500 images = 70 Operations * 2 levels * 125 examples
      - Not realistic (manipulators routinely stack manipulations)
    - The average graph depth in MFC19 was ~4
      - $6.0*10^9$ images = $70^4$ Operations * 2 levels * 125 examples

# Manipulation Reference Collection Challenges

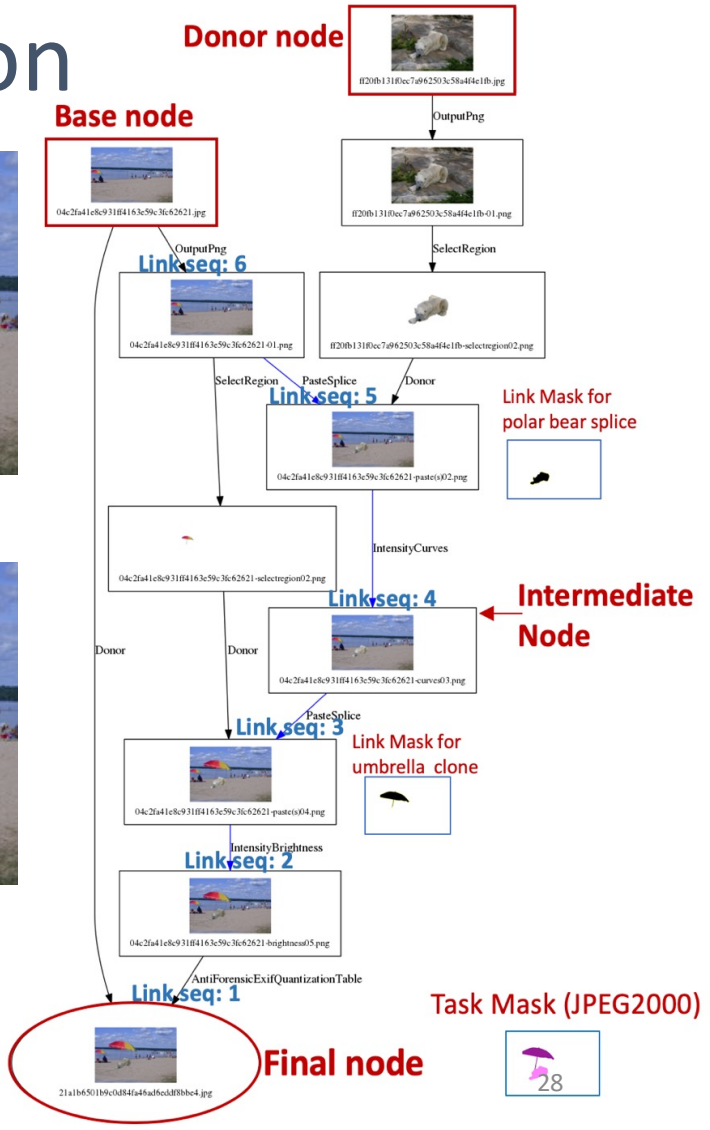
- What reference ground-truth data to collect
  - Time, labor, cost, usage, value
- Effective evaluations require knowledge:
  - If the media was manipulated
  - What editing tool was used
  - Who did the manipulation
  - What is the original media
  - What operations were used
  - How the media was manipulated
  - Where the manipulation occurred
  - Semantics of the manipulation: malicious vs. benign
- How
  - Post manipulation interpretation is nearly impossible

# Concepts in the Data Annotation

- Journal
  - Directed graph
- Node
- Link
- High Provenance (HP)
  - Original image
- Manipulated image
- Base Image – start node
- Final Image – end node
- Donor Image – donor node
- Probe – test sample

Manipulated image

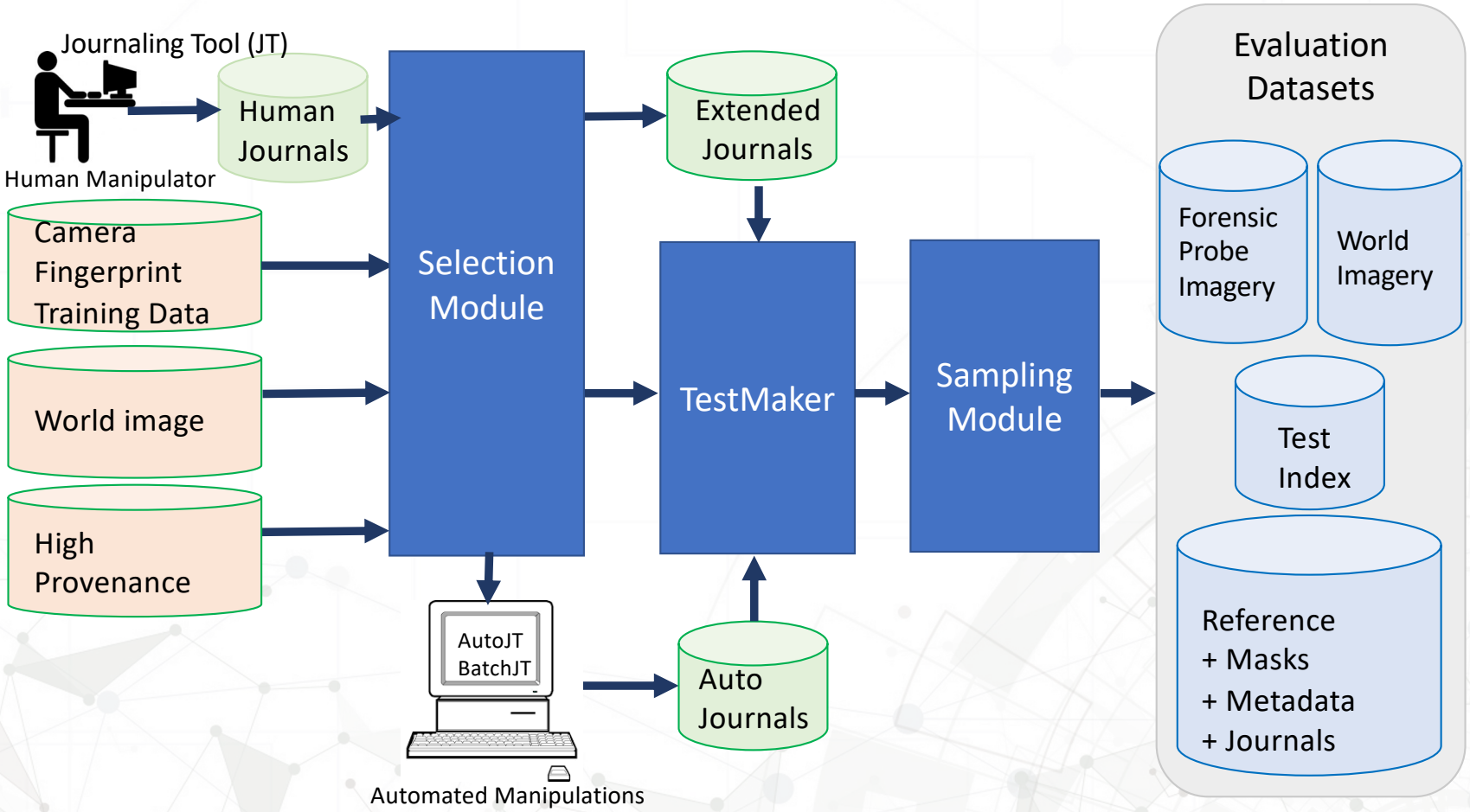
Original image/Base Image



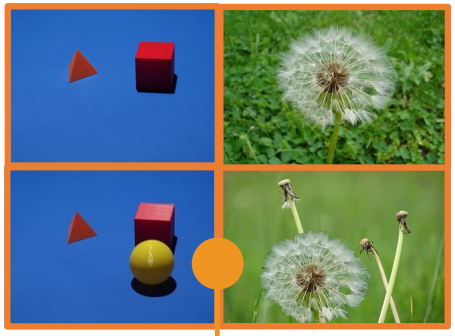
# Manipulation Reference Collection Approaches

- Separate analytics team from data generation team to avoid eval. bias/data bias
  - Drive the analytics teams from their comfortable zone
- Human + machine for data collection and production
  - Human manipulation (realistic)
  - Automatic manipulation (reduce cost)
  - Extended manipulation (special study)
- Journaling Tools (collaborated with PAR Government Systems)
  - Manipulation journaling tool (JT)
    - Record the manipulation step by step with a graph
    - Automate collection of manipulation region mask
  - Automatic journaling tool (Auto-JT)
  - Extended journaling tool (Extended-JT)
- Real-world simulation: social media laundering (UC. Denver)

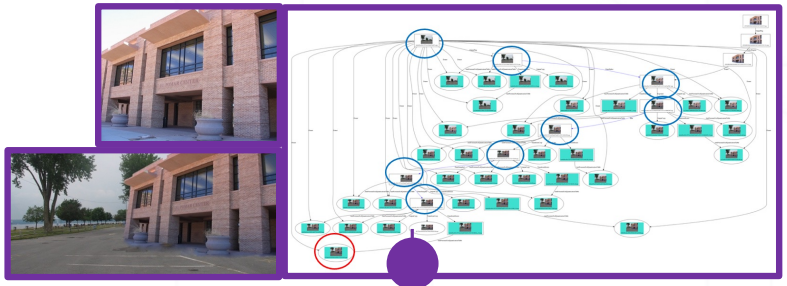
# Evaluation Dataset Production Infrastructure



# OpenMFC Evaluation Dataset (1)



NC16

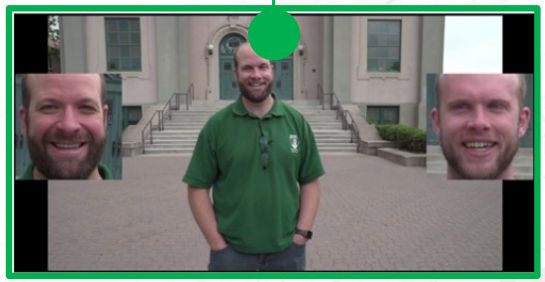


NC17

MFC19

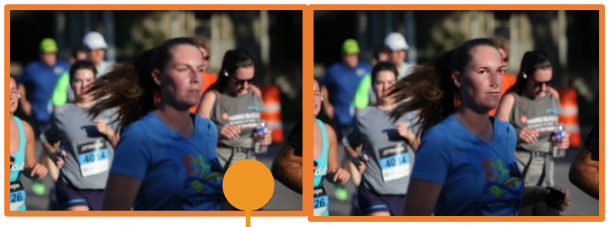



MFC18



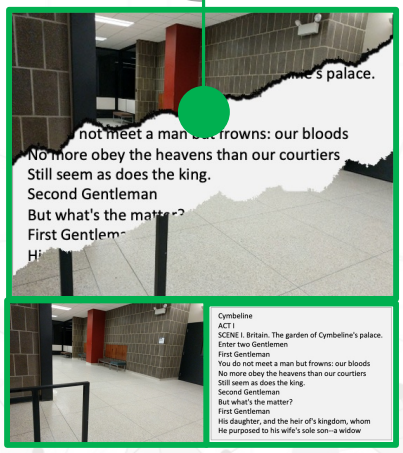
MFC20

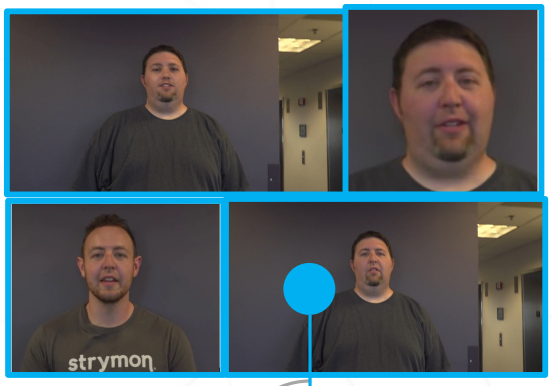
# New! OpenMFC Evaluation Dataset (2)



## OpenMFC GAN Image

## OpenMFC StegD




## OpenMFC Deepfake Video

## OpenMFC Image Splice

# OpenMFC 2022 Evaluation Dataset

Table: OpenMFC2020 Evaluation datasets

| OpenMFC 2022 Evaluation Dataset | OpenMFC 2022 Task | Corresponding MFC Dataset Name | Media Type | Media Number | Journal # | Create Date |
|---|---|---|---|---|---|---|
| **OpenMFC20_Image_MD** | IMD | MFC19 EP1 Image | Image | 16K | 1383 | 2022 |
| **OpenMFC20_Video_MD** | VMD | MFC19 EP1 Video | Video | 1530 | 163 | 2022 |
| **OpenMFC20_Image_DD** | IDD | MFC18 GAN FULL Image | Image | 1.3K | 267 | 2022 |
| **OpenMFC20_Video_DD** | VDD | MFC18 GAN Video | Video | 118 | 19 | 2022 |


Table: OpenMFC2022-2023 Evaluation datasets

| OpenMFC 2022 Evaluation Dataset | OpenMFC 2022 Task | Media Type | Media Number | Create Date |
|---|---|---|---|---|
| **OpenMFC22_SpliceImage_MD** | ISMD | Image | 2K | 2022 |
| **OpenMFC22_Image_StegD** | StegD | Image | 480 | 2022 |
| **OpenMFC22_Image_DD** | IDD | Image | - | 2023 |
| **OpenMFC22_Video_DD** | VDD | Video | - | 2023 |

# OpenMFC Development Datasets

Table: OpenMFC released datasets

(Highlighted: the evaluation sets; Grayed: the development sets in DARPA MediFor MFC)

| MFC Released Dataset | Media Type | MFC Dev./Eval. | Media Number | Journal # | Create Date |
|---|---|---|---|---|---|
| **Kick Off (NC16) Image** | Image | Dev. | 1.1K | 400 | Jul-16 |
| **NC17 Dev Image** | Image | Dev. | 3.5K | 398 | Mar-17 |
| **MFC18 Dev1 Image** | Image | Dev. | 5.6K | 197 | Jan-18 |
| **MFC18 Dev2 Image** | Image | Dev. | 38K | 411 | Feb-18 |
| **NC17 EP1 Image** | Image | Eval. | 4K | 406 | Jun-17 |
| **MFC18 EP1 Image** | Image | Eval. | 17K | 758 | Mar-18 |
| **NC17 Dev Video** | Video | Dev. | 212 | 25 | Mar-17 |
| **MFC18 Dev1 Video** | Video | Dev. | 116 | 9 | Jan-18 |
| **MFC18 Dev2 Video** | Video | Dev. | 231 | 21 | Feb-18 |
| **NC17 EP1 Video** | Video | Eval. | 360 | 34 | Jun-17 |
| **MFC18 EP1 Video** | Video | Eval. | 1028 | 114 | Mar-18 |

# New! OpenMFC 2022 Data Release Website

- Evaluation and Development Data
  - OpenMFC Registration Form
  - OpenMFC Data Release Form

**Present OpenMFC Evaluation Data**

OpenMFC participants need to download the OpenMFC datasets listed below. The OpenMFC Evaluation data only contain image probe and indexes files. No reference groundtruth data is provided during evaluation period. OpenMFC dataset contains the same data as MFC2019 Image/Video Evaluation data and MFC2018 GAN Challenge full image data. The only difference is the dataset name.

**Manipulation Detection (MD) Task**

- **Image Manipulation Detection (IMD):** OpenMFC20_Image_MD (previous MFC19 Image Data) - Download_tarSized_OpenMFC20_Image_MD.sh
- **Image Splice Manipulation Detection (ISMD):** OpenMFC22_SpliceImage_MD - Download_tarSized_OpenMFC22_ImageSplice_MD.sh
- **Video Manipulation Detection (VMD):** OpenMFC20_Video_MD (previous MFC19 Video Data) -Download_tarSized_OpenMFC20_Video_MD.sh

**Deepfake Detection (DD) Task**

- **Image Deepfake Detection (IDD):** OpenMFC20_Image_DD (previous MFC18 GAN Full Image Data) - Download_tarSized_OpenMFC20_Image_DD.sh
- **Video Deepfake Detection (VDD):** OpenMFC20_Video_DD (previous MFC18 GAN Video Data) - Download_tarSized_OpenMFC20_Video_DD.sh

**Steganography Detection (StegD) Task**

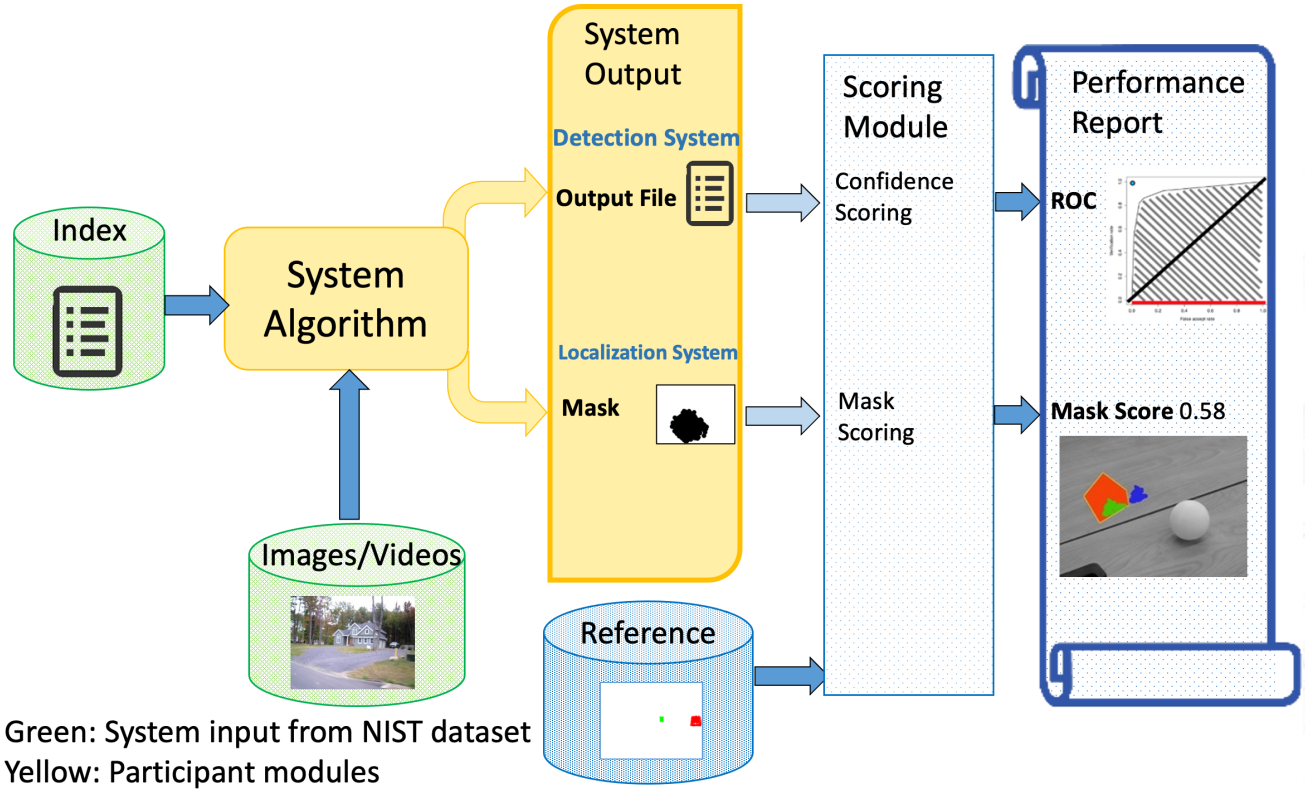
- **Steganography Detection (StegD):** OpenMFC22_Image_StegD - Download_tarSized_OpenMFC22_Image_StegD.sh

**Past Media Forensics Challenge(MFC) Evaluation Data**

MFC19 is used in the OpenMFC 2022 Evaluation. MFC20 will be used for the future OpenMFC Evaluation. MFC18, MFC19, and MFC20 data have the same format. MFC18 reference groundtruth data is provided below.

# How to evaluate?

# OpenMFC: How to evaluate (1)

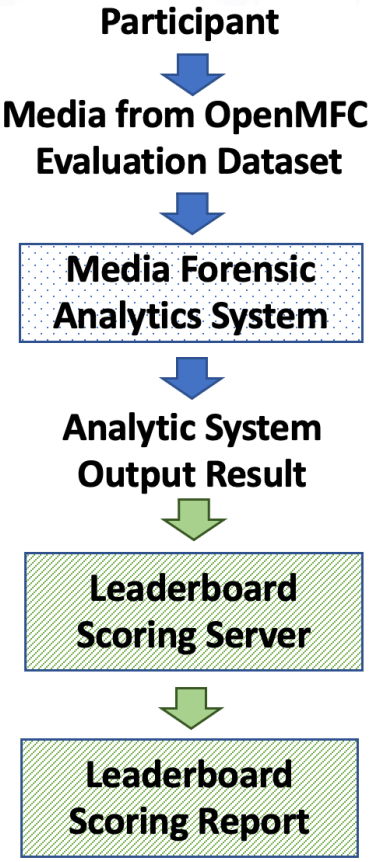



Green: System input from NIST dataset
Yellow: Participant modules
Blue: NIST Evaluation modules

Figure: OpenMFC Evaluation Pipeline

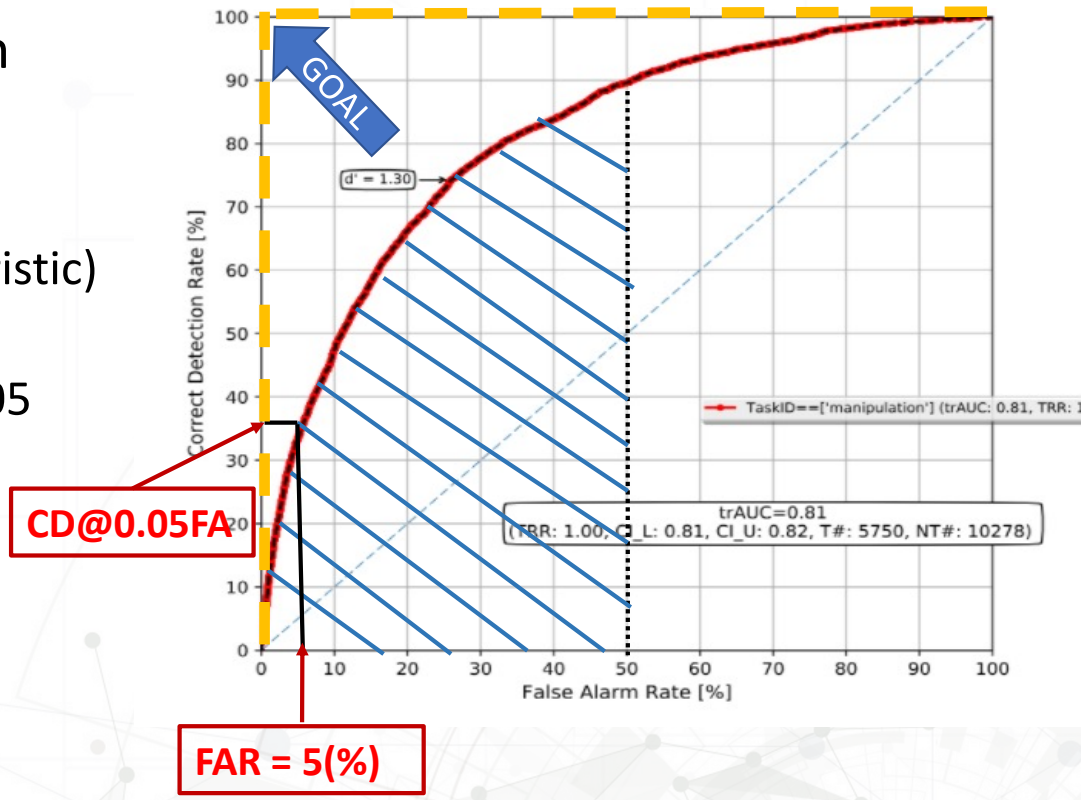
# OpenMFC: How to evaluate (2)

- OpenMFC Take-home evaluation
  - NIST releases test data to participants
  - Participant submits system output
  - Evaluation website provides a leaderboard to report evaluation results



# Detection System Evaluation Metrics

- Evaluate the accuracy of a system output (e.g., confidence score)
- Evaluation metrics
  - ROC (Receiver Operating Characteristic)
  - AUC (Area Under Curve)
  - CD (Correct Detection) @ FAR = 0.05



# What we found?

# OpenMFC IMD dataset: OpenMFC20_Image_MD

- 16 thousands images for MFC19 Eval.; Over 80+ operations

| Name | Definition | Counts |
|---|---|---|
| **Splice** | Any operation that takes a region from a donor media and pastes it into a probe | 2342 |
| **Clone** | Pixels are sampled from the image and pasted back in different area of the image | 1268 |
| **Splice/Clone** | Pixels are pasted within or between the images | 3005 |
| **Crop** | Outer pixel regions from a probe image are removed | 579 |
| **Resize** | Image dimensions from a probe image are changed | 653 |
| **Intensity** | A range of intensity pixel values is changed | 2269 |
| **Antiforensic** | Any techniques that erase processing history of image manipulations | 5055 |
| **Antiforensic-PRNU** | Any techniques that use PRNU | 1304 |
| **Antiforensic-CFA** | Any techniques that use CFA | 200 |
| **Social Media** | Any techniques that use social media related operations | 348 |
| **Global Blur/Smooth** | Any techniques that use a low-pass filter (globally) to remove outlier pixels (e.g., noise) | 62 |
| **Local Blur/Smooth** | Any techniques that use a low-pass filter (locally) to remove outlier pixels (e.g., noise) | 1143 |
| **GAN** | Any operations that use GAN-based techinques locally/globally | 530 |
| **NonGAN-CGI** | Any operations that use non-GAN CGI | 309 |
| **Distortion** | Deformation of images | 918 |
| **Remove** | Remove a set of pixels. | 833 |
| **Face Manipulation** | Any manipulation done to a face. | 22 |
| **All** | All data without selective scoring | 5750 |

# MFC results on OpenMFC20 _Image_MD

- 47 analytic systems from 13 teams
- Highest AUC: 0.866
- CD@FAR=0.05: 0.456

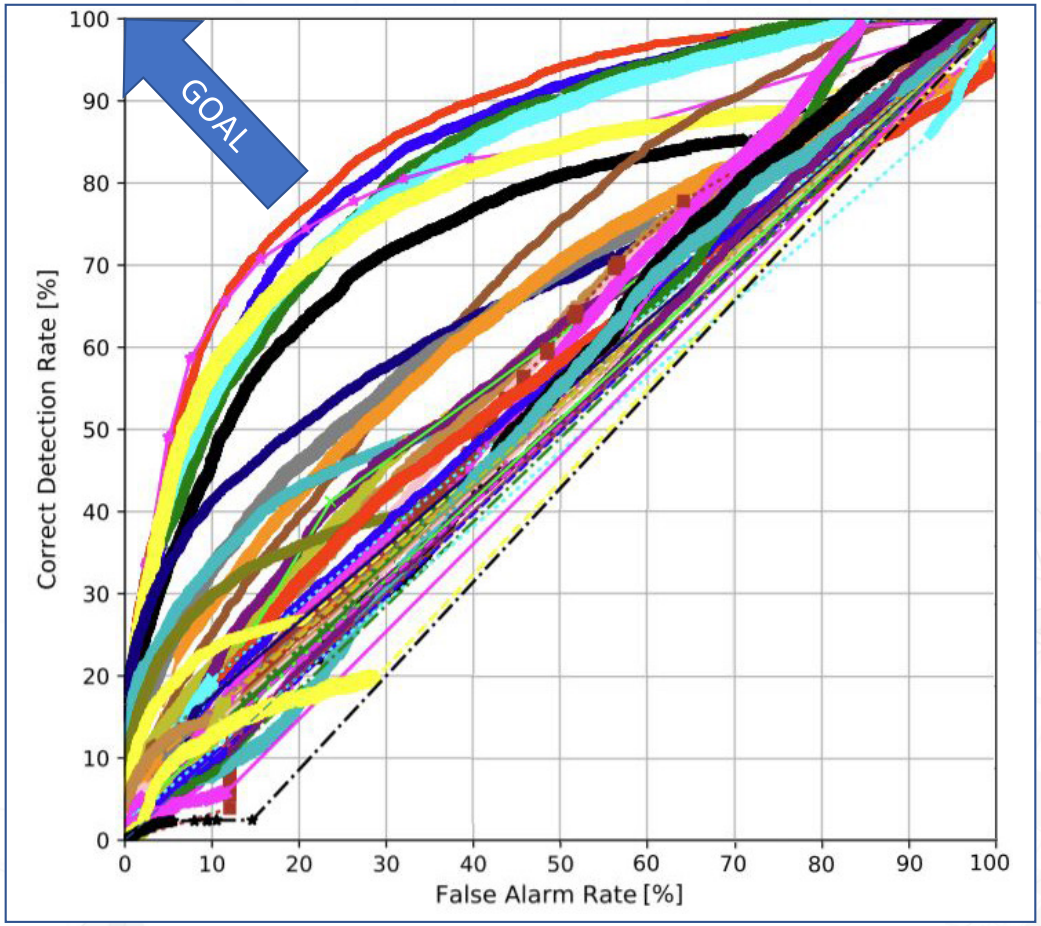
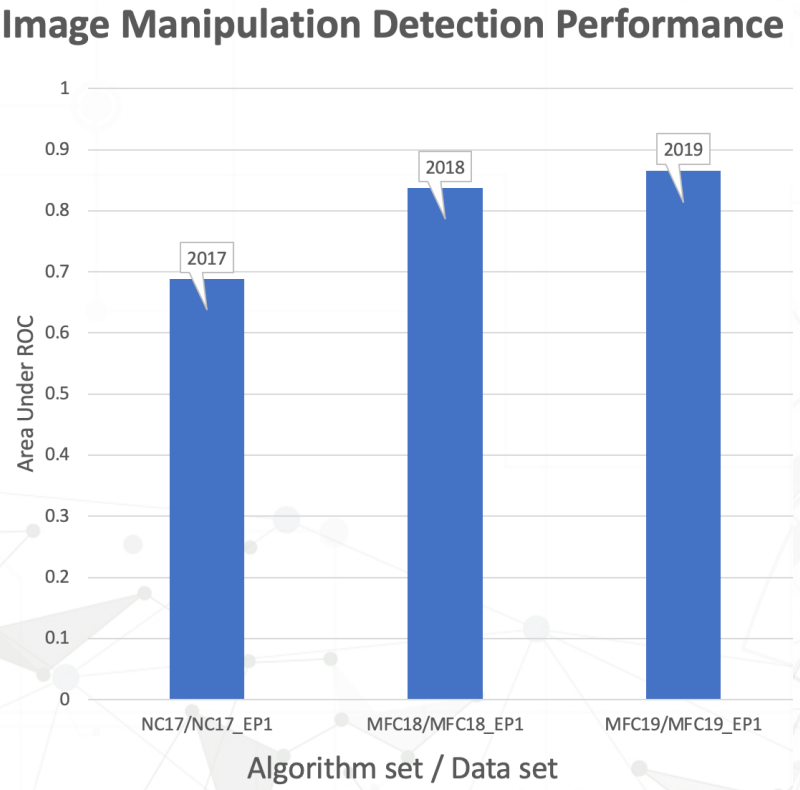



Figure: MFC results on OpenMFC20 _Image_MD

# MFC Cross Year IMD Performance Comparison

- Not apples to apples comparison: dataset difficulty increases each year



# NIST OpenMFC IMD Leaderboard

https://mfc.nist.gov/#pills-leaderboard

Table : OpenMFC IMD system performance (AUC)

**Image Manipulation Detection (IMD)**

**IMD-IO (Image Only)**

*Updated: 2022-11-23 09:18:07 -0500*

Previous 1 Next

| RANK | SUBMISSION ID | SUBMISSION DATE | TEAM NAME | SYSTEM NAME | AUC | CDR@0.05FAR | AVERAGE OPTIMAL MCC |
|---|---|---|---|---|---|---|---|
| 1 | 63 | 2021-06-07 11:26:58 | Team1 | System1 | 0.993707 | 0.972 | |
| 2 | 10 | 2020-11-05 21:53:02 | Team2 | System2 | 0.616186 | 0.071351 | |
| 3 | 67 | 2021-06-08 00:51:16 | Team2 | System3 | 0.5 | 0.05 | |
| 4 | 81 | 2021-06-26 00:31:16 | Team2 | System3 | | | 0.0553688699305928 |

Showing 1 to 4 of 4 entries

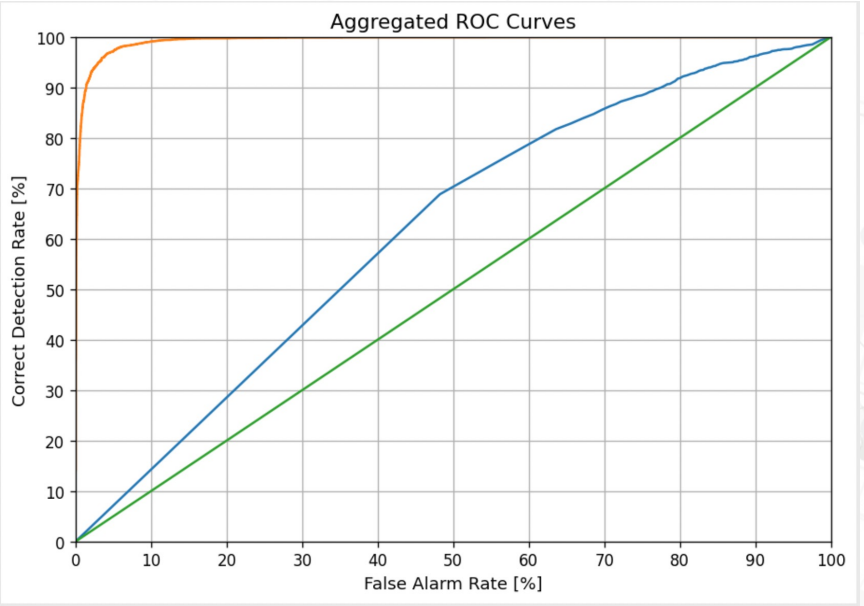


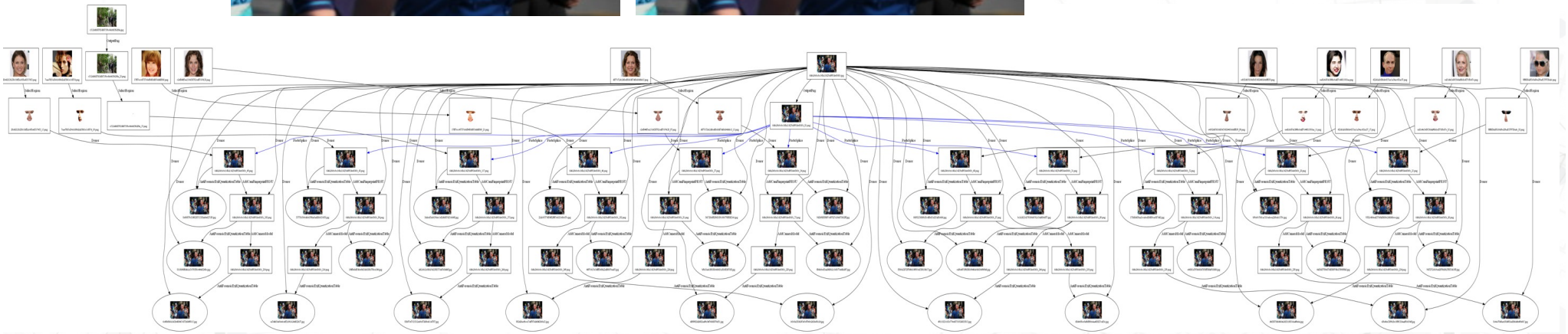

Figure : OpenMFC IMD system performance (ROC)

# OpenMFC IDD

- Task: To detect whether an image was GAN manipulated or deepfaked
- OpenMFC IDD dataset: OpenMFC20_Image_DD
- 1,340 images for MFC18 GAN Image Detection Eval.
- Manipulations:
  - Face swap – GAN vs. real face
  - GAN Fill, GAN Erase (no face)
  - CFA camera model (no face)
- Reference definition
  - Manipulated by GAN operation (with/without other operations) – IsTarget = 'Y'
  - Manipulated by other operations (not GAN) – IsTarget = 'Y'
  - NotManipulated – IsTarget = 'N'

# A face swap journal example for IDD[1]

11 donor
33 manipulated

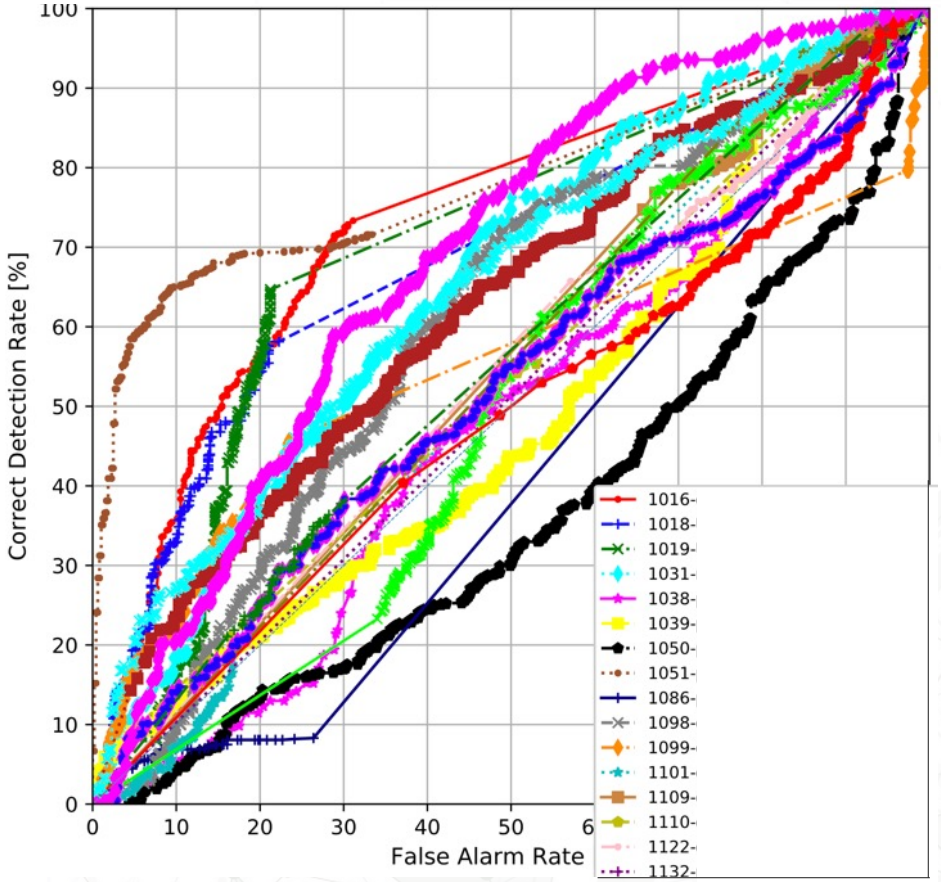


[1]The journal was created by PAR Government

# OpenMFC Submission Files: IDD

- Task:  IDD
- Submission file name : e.g., OpenMFC20_Image_DD_ SystemID_TeamID_Date.tar.gz
  - 1341 rows
  - Example: the first several rows of the submission file
    - ProbeFileID|ConfidenceScore|ProbeStatus
    - 002d809170f2b761ea8136369bddbb1c|0.360968997|Processed
    - 002e5f4b081c3b41d93e1414f40a588d|0.354781056|Processed
    - 004b354f88f064e2ac9ff555abf3ee94|0.11301657|Processed
    - … …

# MFC results on OpenMFC20_Image_DD



# OpenMFC2022 Leaderboard: IDD

https://mfc.nist.gov/#pills-leaderboard

**Image Deepfakes Detection (IDD)**

*Updated: 2022-11-23 09:18:03 -0500*

Previous 1 2 Next

| RANK | SUBMISSION ID | SUBMISSION DATE | TEAM NAME | SYSTEM NAME | AUC | CDR@0.05FAR | AVERAGE OPTIMAL MCC |
|---|---|---|---|---|---|---|---|
| 1 | 90 | 2021-07-10 09:56:14 | Team2 | System4 | 0.689716 | 0.207018 | |
| 2 | 93 | 2021-07-30 18:13:11 | Team2 | System4 | 0.683956 | 0.187135 | |
| 3 | 75 | 2021-06-14 04:12:30 | Team3 | System5 | 0.554261 | 0.012865 | |
| 4 | 52 | 2021-06-01 15:41:20 | Team3 | System5 | 0.547125 | 0.009357 | |
| 5 | 82 | 2021-06-23 09:21:26 | Team2 | System6 | 0.500033 | 0.051077 | |
| 6 | 36 | 2021-05-06 07:32:03 | Team2 | System4 | 0.5 | 0.05 | |
| 7 | 91 | 2021-07-21 15:53:47 | Team2 | System4 | 0.478445 | 0.009357 | |
| 8 | 86 | 2021-07-07 06:59:01 | Team2 | System4 | 0.41193 | 0.00117 | |
| 9 | 87 | 2021-07-07 07:04:14 | Team2 | System4 | 0.403957 | 0.004678 | |
| 10 | 89 | 2021-07-08 04:20:47 | Team2 | System4 | 0.398674 | 0.003509 | |

Showing 1 to 10 of 16 entries

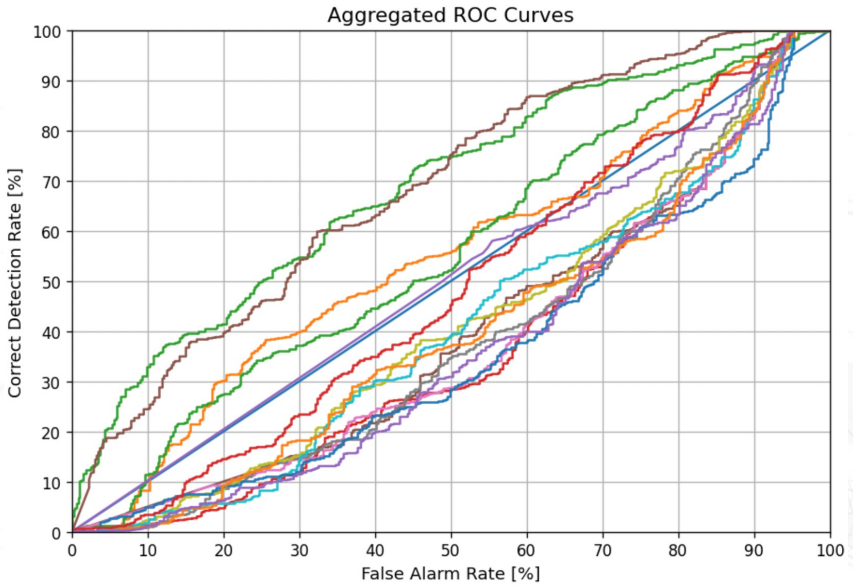



Figure : System performance on the OpenMFC20_Image_DD dataset

- For the IDD task, the best AUC score is 0.69, compared with 0.79 in MFC18.

# OpenMFC VDD

- Task: To detect whether a video was GAN manipulated or deepfaked.
- OpenMFC VDD dataset: OpenMFC20_Video_DD
- 118 videos for MFC18 GAN Video Detection Eval.
- Manipulations:
  - Deepfakes
  - Frame drop
  - GAN Inpainting
- Reference definition
  - Manipulated by GAN/deepfake operation (with/without other operations) – IsTarget = 'Y'
  - Manipulated by other operations (not GAN/deepfake) – IsTarget = 'Y'
  - NotManipulated – IsTarget = 'N'

| NIST Data Sets | Major operations | Probe |
|---|---|---|
| MFC18 Eval. GAN Video<br><br>(total 118 probes) | PAR deepfakes | 18 |
| | Denver deepfakes | 40 |
| | Michigan GAN drop frame | 20 |
| | PAR GAN Inpainting | 20 |
| | World video | 20 |

# OpenMFC20_Video_DD

- The videos were collected and manipulated by UC Denver's team


85edfcdfed8e2101a6a78936bc4be733.mp4

04572fb8d0156500fb82415ce1ed55aa.mp4

# Analysis and Discussion

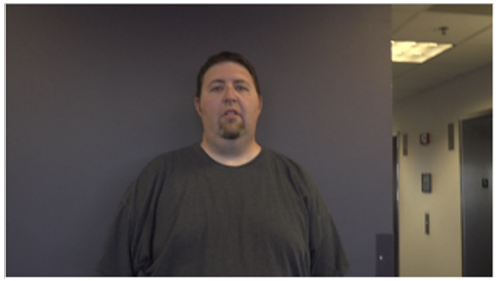
- Media Forensics is still in the early stage
- Performance GAP: many existing self-evaluated research papers show very high detection accuracy, but our tests on media show that detectors have severe limitations, preventing them from being used in real world applications. Self-evaluation lab reports do often not reflect performance against variability of realistic scenarios
- Recurring evaluation is necessary: A series of recurring evaluations are more desirable than a one-shot evaluation
  - Arms Race: between the media forensics and deepfakes/anti-forensics
  - Data Drift: the old detection systems fail to capture advancements of the new deepfake tools

# What is the future plan?

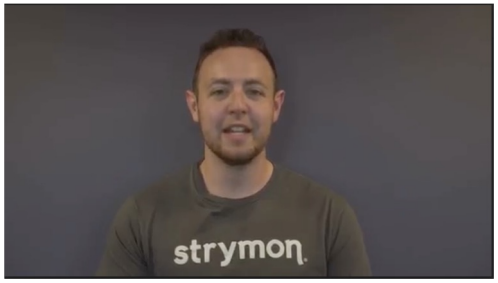
# OpenMFC Deepfake Datasets (Tomorrow's talk)

- Data Source: (1) VidTIMIT from Prof. Sanderson; (2) MFC videos
- Tools: (1) Open tool: DeepFaceLab; (2) Close-source tool: Celeb-DF provided by Univ. at Buffalo


(a) Original video

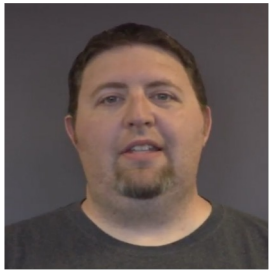


(b) Donor video

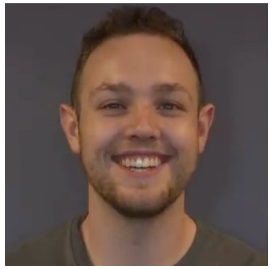



(c) Deepfaked video

Figure: A deepfaked video example using the MFC videos with the Celeb-DF face swap tool.



(a) Original (target) video



(b) Facial expression donor (source) video

(c) Face reenact video of the similar facial expression

Figure: A face reenact video example using the MFC videos with the Celeb-DF face reenact tool.

# Summary

- The OpenMFC datasets have been release to over 700+ peoples over 200 organizations and 26 countries and regions worldwide (continuously increasing).
- OpenMFC 2021 Workshop (Dec. 7-9, 2021)
  - https://mfc.nist.gov/workshop
  - https://mfc.nist.gov/workshopagenda
- OpenMFC 2022 Workshop (Dec. 6-7, 2022)
  - https://mfc.nist.gov/OpenMFC2022workshopagenda/

- New deepfake evaluation program is under construction
- New OpenMFC deepfake datasets are coming!

# Takeaway

- Media Forensics is still in the early stage
- Media Forensics intrinsically is different from other traditional research topics
- OpenMFC Online website: **https://mfc.nist.gov**
- **Join the OpenMFC program!**

# OpenMFC Website: **https://mfc.nist.gov**

- Dataset Release
  - Signup: NC16 Kickoff
  - Signed two agreements
    - Development data with ground-truth
    - Evaluation data without ground-truth
- NIST OpenMFC leaderboard scoring server
- MediScore
  - Github: https://github.com/usnistgov/MediScore/
- Slack: http://openmfc.slack.com
  - Discussion channel: https://app.slack.com/client/T017MTH6RHT/C017MTH7LRK
- OpenMFC performer Google group: openmfc-performer@list.nist.gov

# Questions?

OpenMFC team: mfc_poc@nist.gov

# Thank You!