

OpenMFC 2021 Workshop NIST Talk Overview

Day 1:

- Opening Remark (Yooyoung Lee)
- Evaluation Program Overview (Haiying Guan)
- Discussion - Evaluation Design (Haiying Guan)

Day 2:

- 2020-2021 Evaluation Result (Haiying Guan)
- Evaluation Platform Infrastructure (Lukas Diduch)
- Discussion - Evaluation Leaderboard & System Analysis (Lukas Diduch and Haiying Guan)

Day 3:

- 2021-2022 Evaluation Plan (Haiying Guan)
- Discussion –2021-2022 Evaluation (Haiying Guan)
- Closing Remark (Jim Horan)

OpenMFC 2021 Workshop Agenda



Day 1: Tuesday Dec. 7, 2021 (EST)

Time Start	Time End	Topic
------------	----------	-------

7:00 AM	11:10 AM	TRECVID 2021 TRECVID 2021 Agenda
---------	----------	--

11:10 AM	11:20 AM	TRECVID/OpenMFC 2021 Break
----------	----------	--------------------------------------

11:20 AM	11:35 AM	OpenMFC 2021 Workshop Opening Remarks Yooyoung Lee, Supervisory Computer Scientist, NIST
----------	----------	--

11:35 AM	12:15 PM	Invited keynote: Local and Holistic Methods to Detect Fake Manipulated Images Lakshmanan Nataraj, Principal R&D Engineer, Trimble Inc.
----------	----------	--

12:15 PM	12:45 PM	OpenMFC Evaluation Program Haiying Guan, Senior Computer Scientist, NIST
----------	----------	--

12:45 PM	12:55 PM	Break
----------	----------	--------------

12:55 PM	1:35 PM	Invited keynote: Generation model attribution of face-swapped Deepfake videos Shan Jia, Postdoctoral researcher, University at Buffalo, State University of New York
----------	---------	--

1:35 PM	1:55 PM	Feedback and Discussion on OpenMFC Program Design NIST OpenMFC team
---------	---------	---

- OpenMFC slack: <http://openmfc.slack.com>
- #OpenMFC channel for discussion
- OpenMFC contact: mfc_poc@nist.gov

OpenMFC 2021 Workshop Agenda



Day 2: Wednesday Dec. 8, 2021 (EST)

Time Start	Time End	Topic
7:00 AM	10:00 AM	TRECVID 2021 TRECVID 2021 Agenda
10:00 AM	10:10 AM	TRECVID/OpenMFC 2021 Break
10:10 AM	10:50 AM	Invited keynote: Fighting AI-synthesized Fake Media Empire Innovation Professor: Siwei Lyu, University at Buffalo
10:50 AM	11:30 AM	Invited keynote: Digital Media: Creating and Preserving Trust Wendy Dinova-Wimmer, Sr. Digital Media Architect, Office of the Public Sector CTO, Adobe
11:30 AM	12:10 PM	Invited keynote: Anti Forensic Attacks Using Generative Adversarial Networks: A New Threat Prof. Matthew Stamm, Drexel University
12:10 PM	12:20 PM	Break
12:20 PM	12:50 PM	OpenMFC 2020-2021 Results Haiying Guan, Senior Computer Scientist, NIST
12:50 PM	1:30 PM	OpenMFC Evaluation Infrastructure Lukas Diduch, Senior Software Engineer, NIST/Dakota Consulting Inc.
1:30 PM	1:45 PM	Feedback and discussion on OpenMFC website, leaderboard, and system analysis tool NIST OpenMFC team

3

OpenMFC 2021 Workshop Agenda



Day 3: Thursday Dec. 9, 2021 (EST)

Time Start	Time End	Topic
7:00 AM	11:10 AM	TRECVID 2021 TRECVID 2021 Agenda
11:10 AM	11:50 AM	Invited keynote: Is to see still to believe in deepfake era? Prof. Jun-Cheng Chen, Research Center for Information Technology Innovation, Academia Sinica
11:50 AM	12:30 PM	Invited keynote: Mobile steganography: Looking to the future Prof. Jennifer L. Newman, Iowa State University
12:30 PM	1:00 PM	Invited keynote: Generating a stegDB dataset for NIST's OpenMFC evaluation program Li Lin, Postdoctoral Research Associate at the Center for Statistics and Applications in Forensic Evidence (CSAFE), ISU
1:00 PM	1:10 PM	Break
1:10 PM	1:25 PM	OpenMFC 2021-2022 Haiying Guan, Senior Computer Scientist, NIST
1:25 PM	1:40 PM	Feedback and discussion on OpenMFC 2022 NIST OpenMFC team
1:40 PM	1:50 PM	OpenMFC 2021 Workshop Closing Remarks Jim Horan, Group Leader, Multimodal Information Group, Information Access Division, ITL, NIST



Open Media Forensics Challenge (OpenMFC) 2021 Workshop

OpenMFC Introduction

Yooyoung Lee,
Haiying Guan, Lukas Diduch , and Ilia Ghorbanian

Multimodal Information Group,
Information Access Division

Day 1, Tuesday, Dec. 7, 2021



Acknowledgement

- NIST contributors
 - Jonathan Fiscus
 - Timothee Kheyrkhah
 - Peter Fontana
 - Jesse G. Zhang
- External collaborators:
 - Prof. Siwei Lyu in University at Buffalo
 - Prof. Jennifer Newman, Prof. Yong Guan, Li Lin from Iowa State University, and Prof. Roy A. Maxion from Carnegie Mellon University

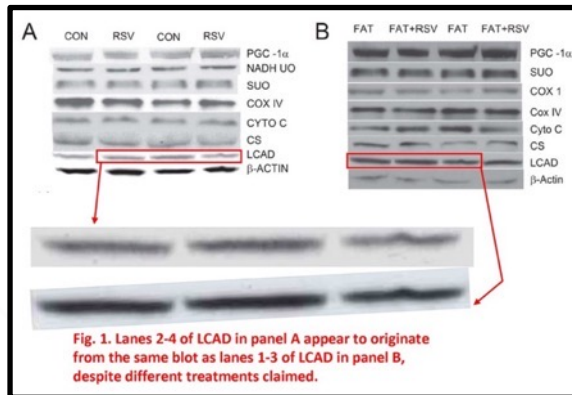
Disclaimer

- 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.
- All images, graphs, and charts are original works created for DARPA MediFor as well as other NIST Programs.

Motivation

Media Forensics is an attempt to determine the authenticity of digital media

Research Fraud¹



Insurance Fraud²



Social Media³



News/Magazines⁴



Public Health &
Safety

Harm Industry

Disinformation
Misconception
Control & Threat



MediFor: Media Forensics

(2017 – 2020)



Sponsor: DARPA MediFor program (PM: Matt Turek)



Definition: determine the authenticity and establish the integrity of visual/audio media



Objective: develop technologies to advance the field of forensics



NIST role: define tasks and metrics, and manage technical evaluations of media forensic technologies

MediFor Evaluation Tasks



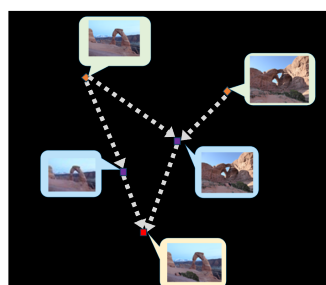
**Image
Manipulation
Detection &
Localization**



**Splice
Detection &
Localization**



**Provenance
Filtering**



**Provenance
Graph
Building**



**Camera
Fingerprint
Verification**



**Event
Verification**



**Video
Manipulation
Detection &
Temporal**

MediFor at A Glance



7
EVALUATION
TASKS



30+
DATASETS



1,200+
SUBMISSIONS



200+
ORGANIZATIONS



20+
COUNTRIES



5+
PUBLICATIONS

Why Was It Challenging?



Large variety of disciplines or domains

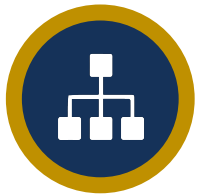
Broad scope



Evaluation design challenges

- Lack of benchmark datasets
- Different data collections and annotations

Large effort and time



Complex scoring protocols

- Holistic, Opt-In, Selective, and special studies

High complexity



Multiple evaluation infrastructures

- Open (take-home) vs Container (sequester)

Participation difficulty



NIST OpenMFC

(2020 – Present)

- Goal: automatically detect and locate manipulations and deepfakes



Image Manipulation Detection and Localization



Video Manipulation Detection



Deepfakes (GAN) Detection

GAN (Generative Adversarial Network)

Details at <https://mfc.nist.gov>



Ongoing Effort



Experiment design and data collection

- Synthetic (GAN-based) data generation
- Comparable real-world data collection



Web-based leaderboard

- Support simplicity & easy to participate



Interactive Dashboard

- Web-based data analysis (data contains rich metadata)
- Research direction for system improvement

Deepfakes Generation

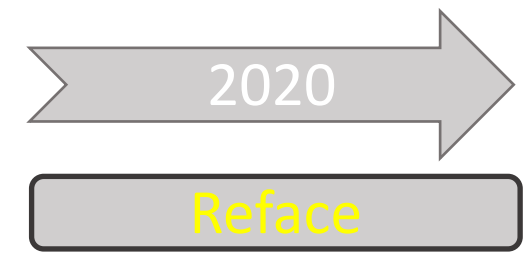
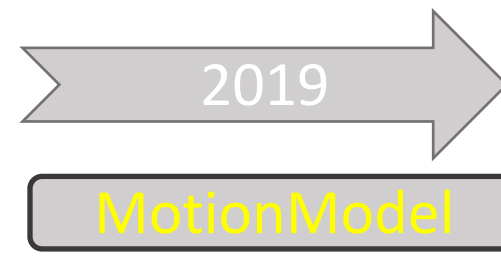
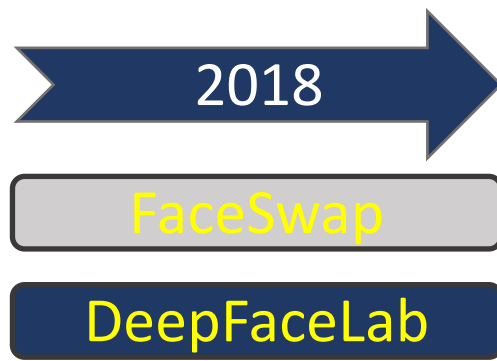
Completed Test

Deepfakes Tools	Release
FaceApp	2017
Deepfakes FaceSwap	2018
DeepFaceLab	2018
First Order Motion Model	2019
Reface	2020

Continued Test

GAN Models	Release
PixelCNN, ProGAN	~2017
SN-GAN, MMD-GAN, Glow	2018
StyleGAN	2019
FSGAN, StyleGAN2	2020
StyleGAN3	2021

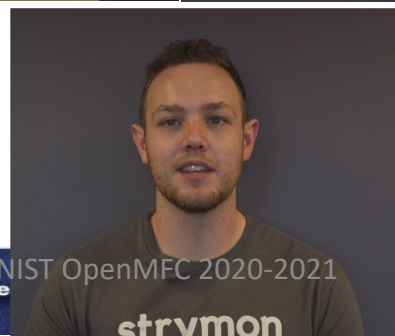
GAN Datasets	Release
CityScapes, ADK20k	2016
CelebA-HQ,	2017
COCO-stuff, VGGFace2	2018
FFHQ	2019
AFHQ v2	2021



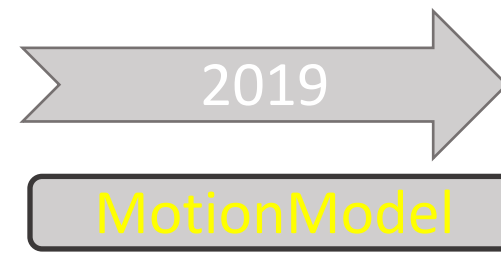
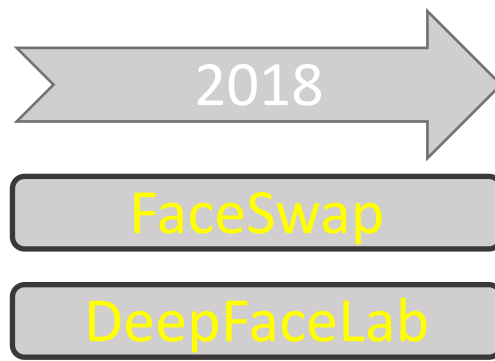
Base



Deepfaked



Donor



Donor

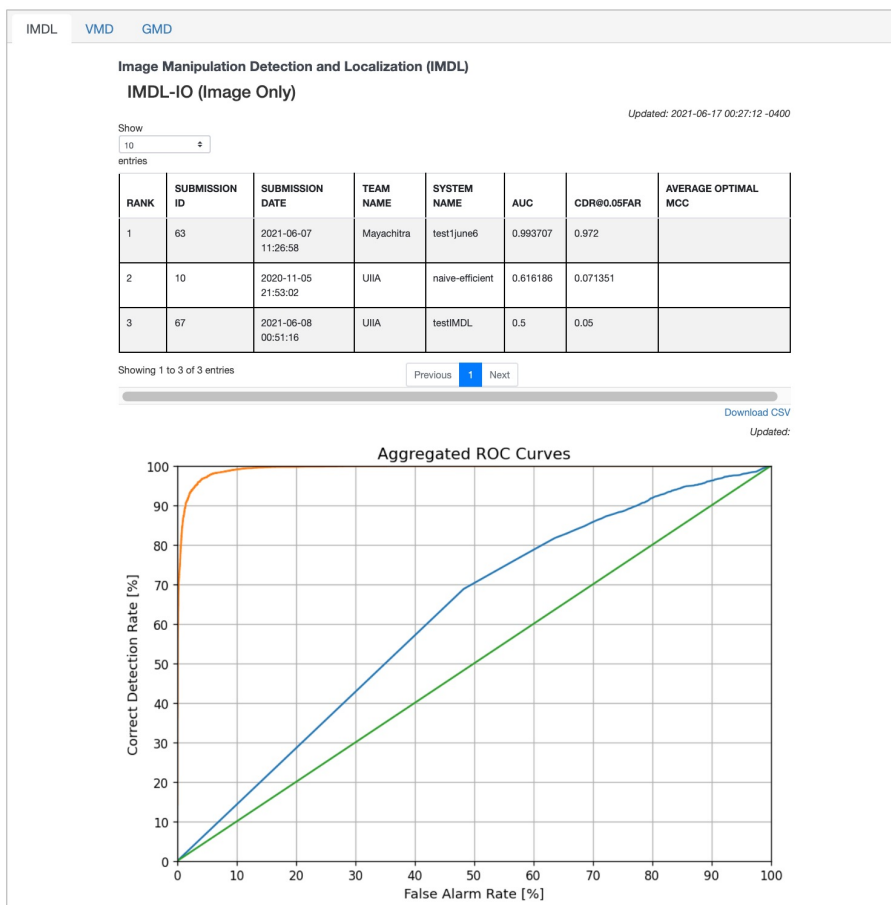


Deepfaked

Web-based Leaderboard

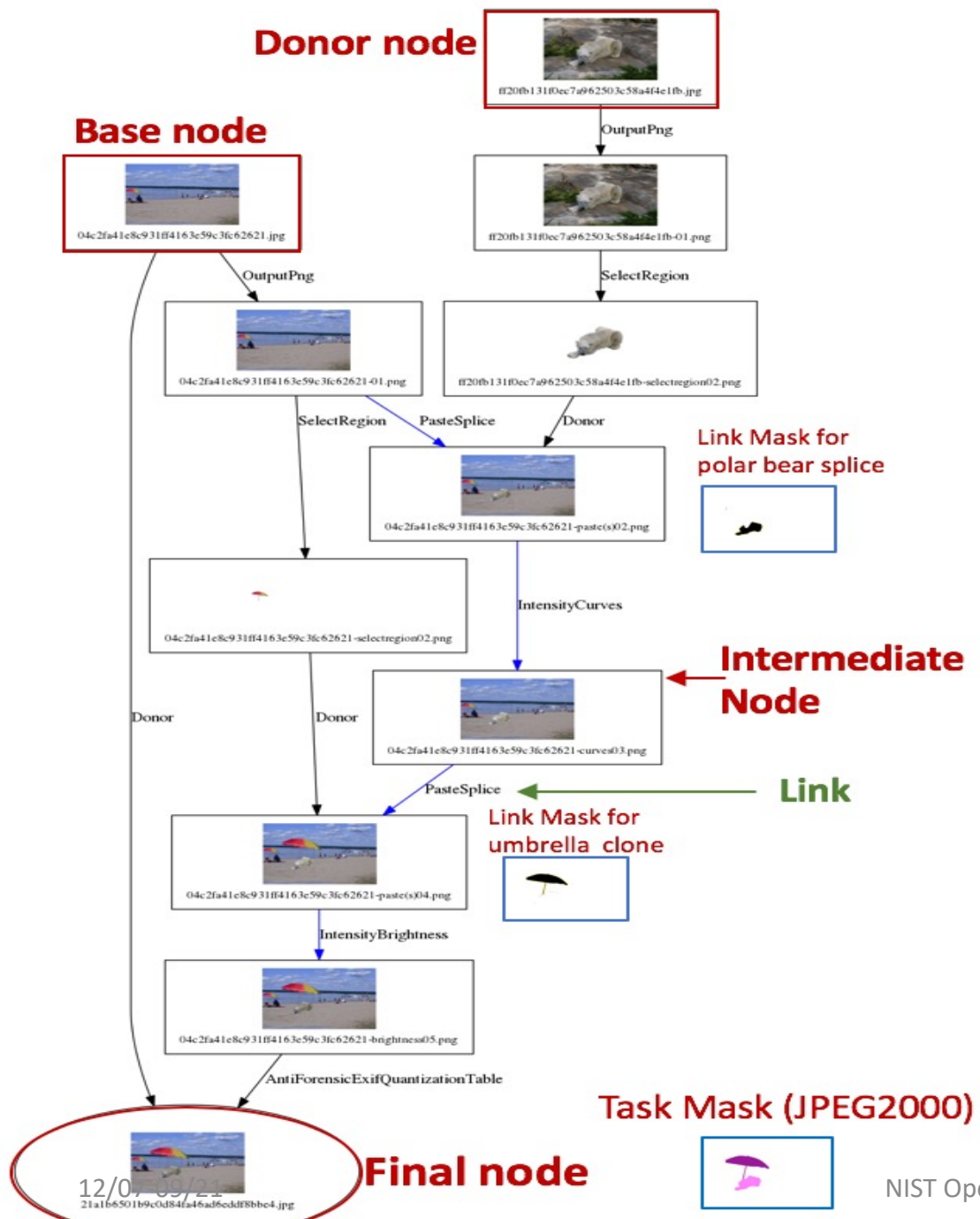
Open Media Forensics Challenge

[OVERVIEW](#) [TASKS](#) [DATA](#) [SCHEDULE](#) **[LEADERBOARD](#)** [SUBMISSION RULES](#) [RESOURCES](#) [CONTACT](#)



Quick Turnaround Leaderboard Evaluation

<https://mfc.nist.gov>



Web-based Media Level Analysis for Validation Set (Provenance)

Interactive Dashboard

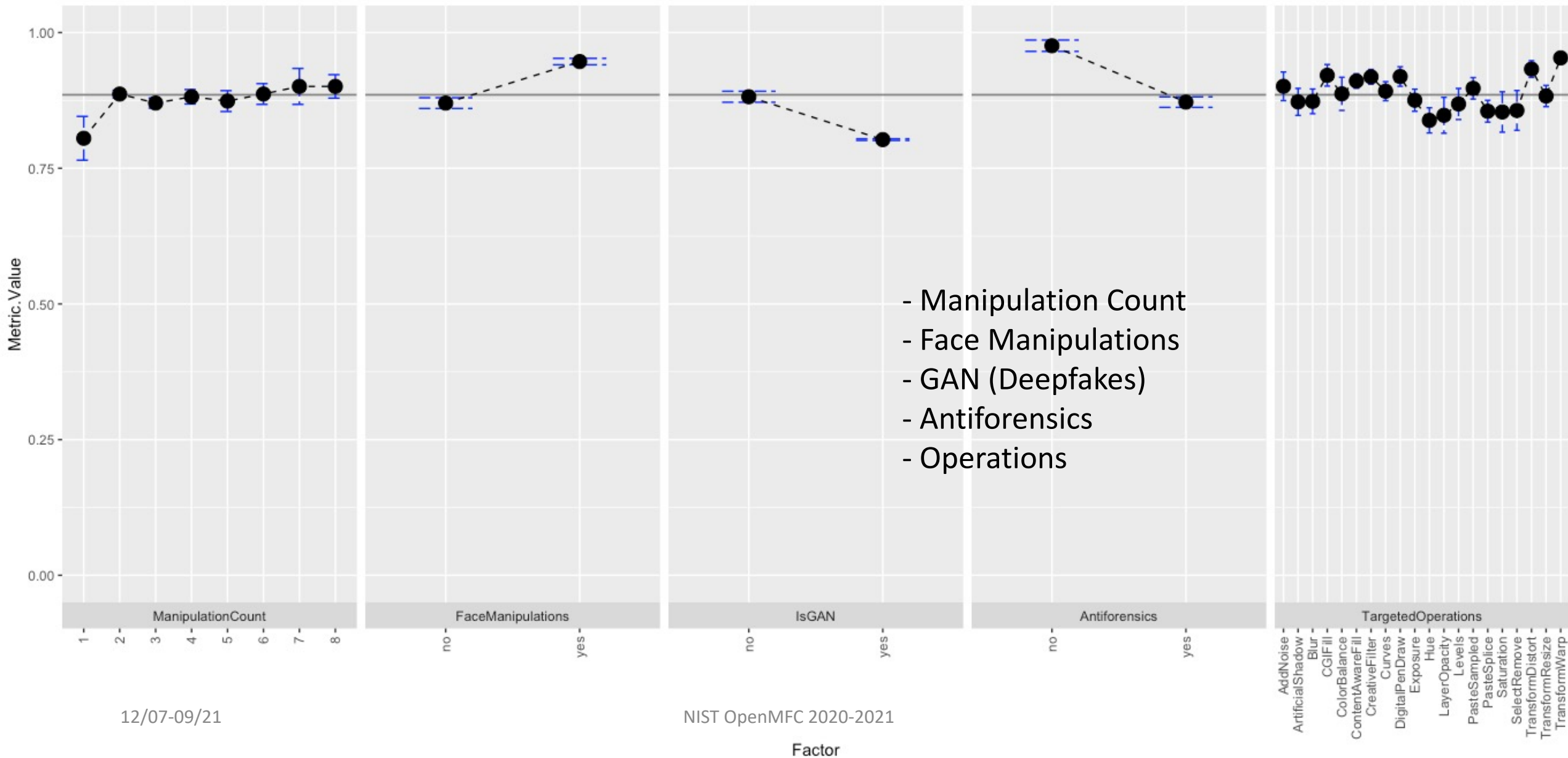
Researchers	What is the accuracy and robustness of a system?
	What are the important factors (and interactions)?
	Which forgery methods are easy/hard to detect?
	How does a system perform across datasets ?

End-Users	How does a system behave in operational environments ?
	What are the optimal settings in my operations?
	How does a system perform on different training data?
	What is the speed efficiency for a system?

Web-based
(& Interactive)
Data Analysis for
both researchers
and end-users

Q: What are important factors that affect system performance?

Main Effects Plot with Error Bars





Our Vision



Expand to “**Consequence Detection**” beyond manipulation detection

- Systematically predicting motivation or intention behind the manipulations/deepfakes
- Categorization & Classification (e.g., violent incitement, vehicle accident)



Contribute to prevent **disinformation and its threat**



Build **collaborations** across sectors and engage **community stakeholders**

NIST AT A GLANCE



3,400+

FEDERAL
EMPLOYEES



5

NOBEL PRIZES



2 CAMPUSES

GAITHERSBURG, MD [HQ]
BOULDER, CO



3,500+

ASSOCIATES



10

COLLABORATIVE
INSTITUTES



400+

BUSINESSES USING
NIST FACILITIES



ManufacturingUSA

NATIONAL OFFICE
COORDINATING 14
MANUFACTURING
INSTITUTES



51

MANUFACTURING
EXTENSION
PARTNERSHIP CENTERS



U.S. BALDRIGE
PERFORMANCE
EXCELLENCE PROGRAM

References

- [1] E.M. Bik, et al. "The prevalence of inappropriate image duplication in biomedical research publications", mBio, 7 (2016)
- [2] <http://www.psblab.org/?p=130>, <https://insurancefraud.org/fraud-stats/>
- [3] <https://brightside.me/creativity-photography/13-famous-photographs-that-are-actually-fake-344410/>
- [4] <https://www.buzzfeednews.com/article/maxseddon/russian-tv-airs-clearly-fake-image-to-claim-ukraine-shot-dow>
- [5] <https://www.faceapp.com>
- [6] arxiv.org/abs/2005.05535
- [7] malavida.com/en/soft-/fakeapp
- [8] www.reddit.com/r/deepfakes
- [9] arxiv.org/abs/2003.00196
- [10] arxiv.org/abs/1812.04948
- [11] <https://hey.reface.ai>
- [12] nirkin.com/-fsgan



OpenMFC Evaluation Program

Haiying Guan

Yooyoung Lee, Lukas Diduch , and Ilia Ghorbanian

Multimodal Information Group,
Information Access Division , ITL, NIST

OpenMFC2021 Workshop : Day 1, Tuesday, Dec. 7, 2021

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.
- All images, graphs, and charts are original works created for DARPA MediFor as well as other NIST Programs.

Acknowledgement

- NIST contributors
 - Jonathan Fiscus
 - Timothee Kheyrkhah
 - Peter Fontana
 - Jesse G. Zhang
- External collaborators:
 - Prof. Siwei Lyu in University at Buffalo
 - Prof. Jennifer Newman, Prof. Yong Guan, Li Lin from Iowa State University, and Prof. Roy A. Maxion from Carnegie Mellon University

OpenMFC Overview Outline

- OpenMFC program overview
 - What, why, who, how
- OpenMFC evaluation design
 - Design challenges, evaluation pipeline, evaluation tasks, evaluation metrics
- OpenMFC 2020-2021 datasets
 - Development dataset, evaluation dataset

Media Forensics: What

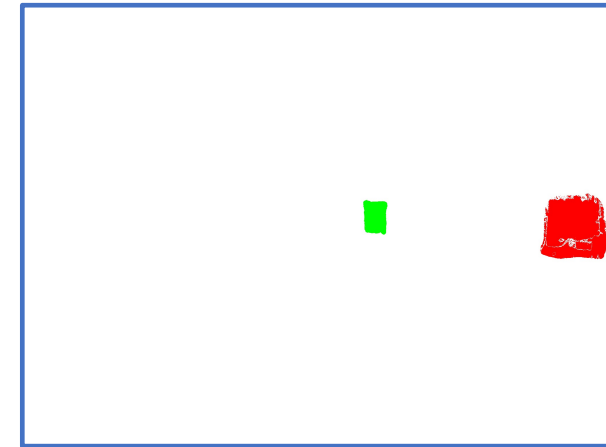
- What is media forensics?
 - “Media Forensic is scientific study into the collection, analysis, interpretation, and presentation of audio, video, and image evidence obtained during the course of investigations and litigious proceedings.”¹



a. Test Image



b. Original Image



c. Manipulated Mask

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

Media Forensics: Why

- “Seeing is not believing”⁴
 - Excellent image editing software
 - Adobe CC, GIMP, Corel Paintshop Pro, Skylum Luminar, DxO PhotoLab, ON1 Photo RAW, ACDSee Photo Studio Ultimate, Pixlr Editor, Canva, PicMonkey, Snappa, PortraitPro, Fotor, ...
 - New advanced technologies
 - Generative Adversarial Network (GAN), Deepfakes¹, CGI, and anti-forensics techniques, ...
- Applications
 - Fake news detection in social media platform^{2,3}
 - Facebook, Twitter, Instagram, Snapchat, and Google
 - Misinformation or disinformation
 - Academic misconduct⁵
 - Criminal law and private investigation, Security⁶
 - Trustworthiness of media content – authentication



Figure Source: Bloomberg Quicktake, 2018¹



Figure Source: EchoFakeD, 2021²

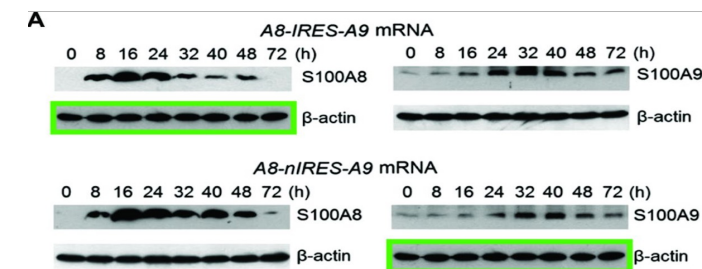


Figure Source: Academic misconduct⁵

¹ <https://www.youtube.com/watch?v=gLoI9hAX9dw>

² <https://doi.org/10.1007/s00521-020-05611-1>

³ <https://www.computer.org/publications/tech-news/research/social-media-verification-assistant>

⁴ H. Farid, "Seeing is not believing," in *IEEE Spectrum*, vol. 46, no. 8, pp. 44-51, August 2009, doi: 10.1109/MSPEC.2009.5186556.

⁵ <https://mbio.asm.org/content/7/3/e00809-16>

⁶ <https://www.cyberscoop.com/social-media-disinformation-represents-security-threat/>

Media Forensics: Challenges (1)

- What is the difference between media forensics and other research topics?
 - Intrinsic property: No single technology fits all
 - 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

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: Challenges (2)

- What is the difference between media forensics and other research topics?
 - Intrinsic property: No single technology fits all
 - Open topic: 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

OpenMFC Program (1)

- 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.
 - Benchmark evaluation is more convincing than self-evaluation
 - Unbiased, neutral position
 - Media forensics analytics is very challenging
 - Keep up with the emerging technologies (GAN, Deepfakes, CGI, anti-forensics, ...)
 - A platform for collaboration
- Open evaluation
 - OpenMFC is open to public, seeks to provide the datasets, build a research platform and inspire research worldwide with leaderboard evaluation, and advance the state-of-the-art of media forensics.
- Closed evaluation
 - DARPA MediFor (<https://www.darpa.mil/program/media-forensics>)
 - DARPA SemaFor (<https://www.darpa.mil/program/semantic-forensics>)

OpenMFC Program (2)

- The NIST OpenMFC is open worldwide. We invite all organizations including past DARPA MediFor Program participants and current DARPA SemaFor Program participants.
- Participation is free. NIST does not provide funds to participants
- NIST team - design the tasks, provides the data, online evaluation platform (real-time report)
- Participant
 - Registration: to register on the website and complete the data license to download the data.
 - Development: to develop the media forensics algorithm/systems.
 - Test: once your system is functional you will be able to upload your outputs to the challenge website and see your results displayed on the leaderboard.

OpenMFC Program: Objective

- OpenMFC Objective
 - Understand the state-of-the-art performance
 - Provide continuously convincing reports: cross-year comparisons
 - Stimulate/Promote the research in media forensics – system performance analysis
 - Help researchers with system performance analysis to improve their system
 - Bridge the gap between lab-report algorithm performance and in-the-field application performance
- Strategy
 - Collaboration instead of competition: nature of the media forensics
 - Group the media forensics researchers and build a strong connected community

Evaluation Design Challenges

- What to evaluate?
 - Not too easy and not too hard
 - Technical methodology varieties brings big challenges in design a unified evaluation framework
- What resources to use?
 - Hundreds if not thousands of manipulation methods
 - Lack of benchmark datasets: human post-annotation doesn't work well.
 - Different technologies need different evaluation data
- What we can get from the evaluation?
 - Baseline performance information
 - State-of-the-art, cross-year performance comparison
- How to evaluate?
 - How to handle the dynamic changes of forensic and anti-forensic technologies?
 - How to adapt the changes and provide the evaluation report in time?

Evaluation Task Design Strategy

- MFC task design

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?

- OpenMFC

single input detection => pair input verification => multi-input analytics integrity

OpenMFC20 Evaluation Tasks

- Image Manipulation Detection and Localization (IMDL)
 - To detect if the image has been manipulated, and then to spatially localize the manipulated region
- Video Manipulation Detection (VMD)
 - To detect if the video has been manipulated
- Image GAN Manipulation Detection (IGMD)
 - To detect GAN-manipulated images (e.g., created by a GAN model, locally/globally modified by a GAN filter/operation, etc.).
- Video GAN/Deepfakes Manipulation Detection (VGMD)
 - To detect GAN/Deepfakes manipulated videos.

OpenMFC20 Evaluation Conditions

- Image Manipulation Detection and Localization (IMDL)
 - Conditions: Image Only (IO) and Image and Metadata (IM)
- Video Manipulation Detection (VMD)
 - Conditions: Video Only (VO) and Video and Metadata (VM)
- Image GAN Manipulation Detection and Localization (IGMDL)
 - Conditions: Image Only (IO)
- Video GAN Manipulation Detection (VGMD)
 - Conditions: Video Only (VO)

Image Manipulation Detection and Localization (IMDL)

System Input

Image(s) + (Metadata)



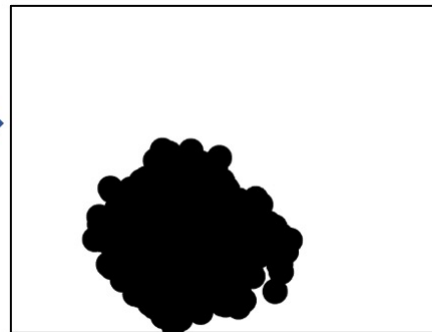
Probe image

**Image
Detection
and
Localization
Analytic
System**

System Output

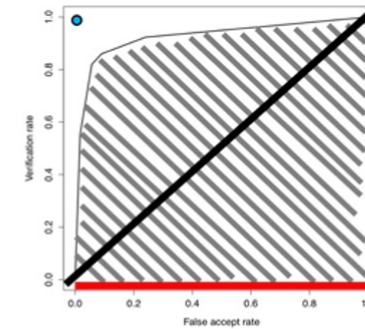
Confidence score
97.86

System output mask

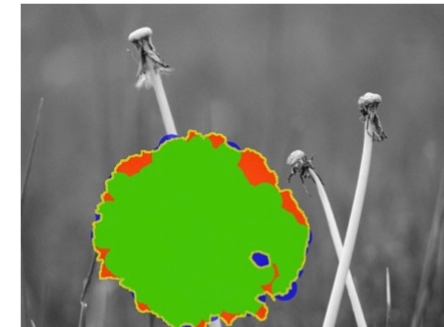


Metrics

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



Manipulated image
Matthews Correlation Coefficient (MCC)



OpenMFC: How to evaluate (1)

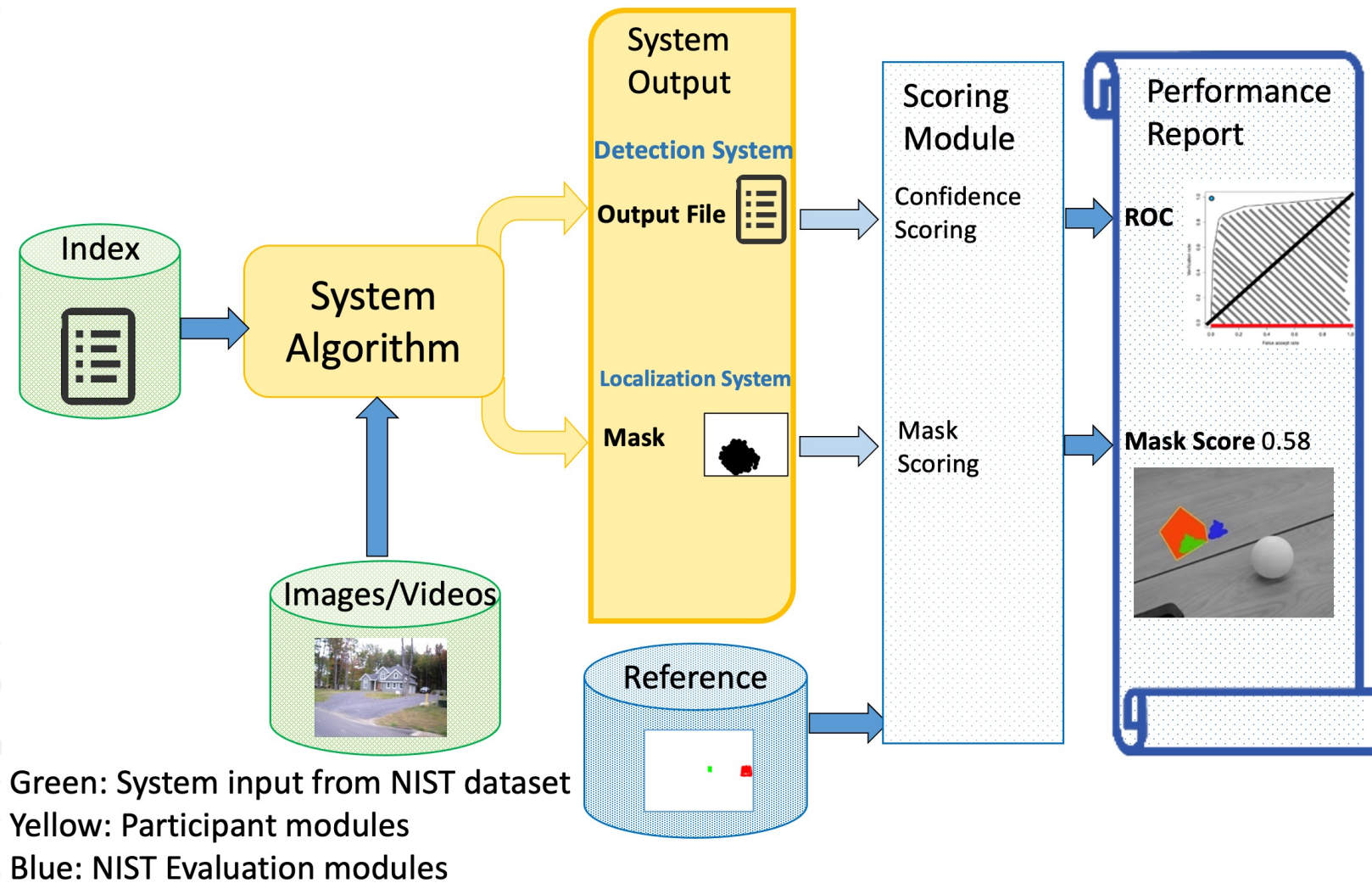
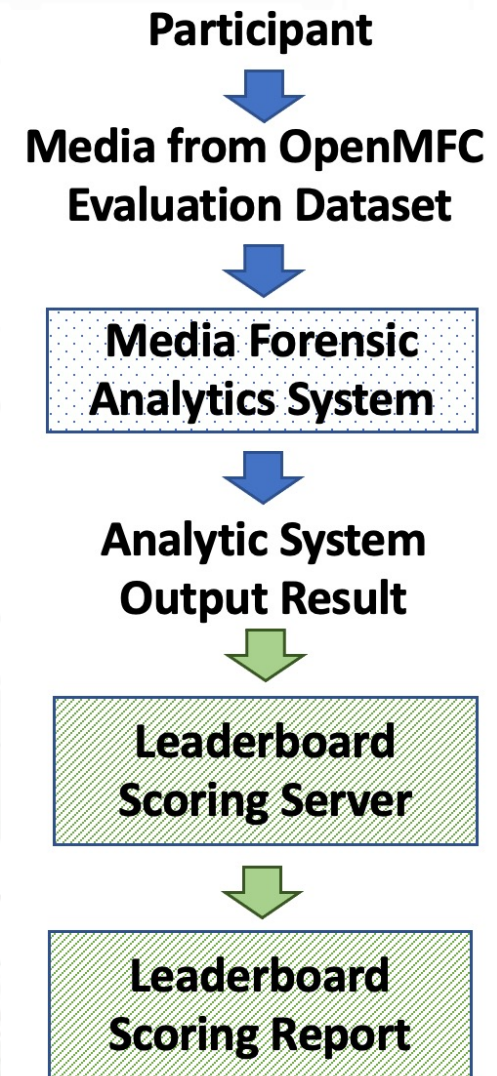


Figure: OpenMFC Evaluation Pipeline

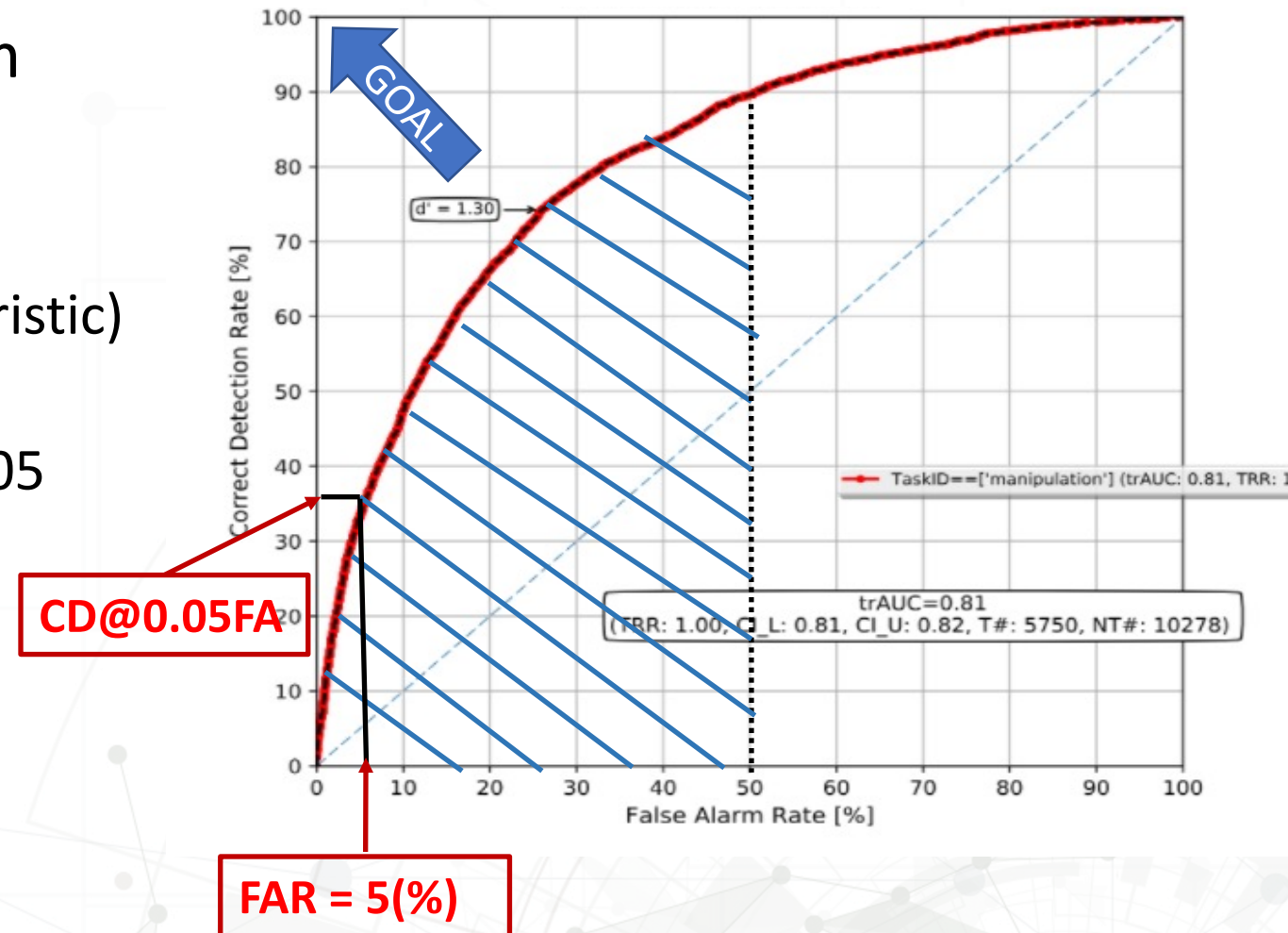
OpenMFC: How to evaluate (2)

- 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



Localization System Evaluation Metrics

- Metrics

- Matthews Correlation Coefficient (MCC)

$$MCC = \frac{TP \times TN - FP \times FN}{\sqrt{(TP + FP) \cdot (TP + FN) \cdot (TN + FP) \cdot (TN + FN)}} \in [-1, 1]$$

- 1 denotes perfect accuracy
 - 0 denotes no correlation
 - -1 denotes perfect inaccuracy.

- Optimum MCC

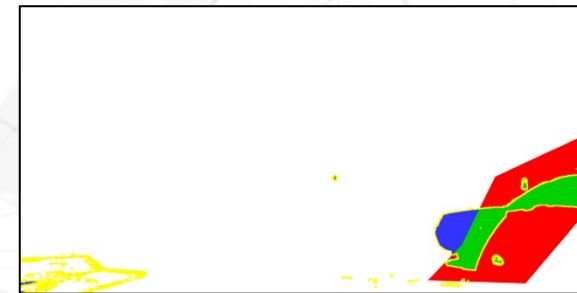
- The MCC at the optimum grey-scale mask threshold
 - Only evaluates on true targets



Probe + ref.
mask overlay



System
output mask



Color-coded
scoring
confusion
matrix

OpenMFC 2020-2021 Evaluation Data

- Evaluation dataset design challenges
- Manipulation reference collection and annotation
- What information you can obtain from the reference data
- Evaluation data production
- OpenMFC evaluation dataset summary

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



Manipulation Reference Collection Approaches

- Separate analytics team from the data generation team
 - 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)



?



Manipulation Journaling Tool



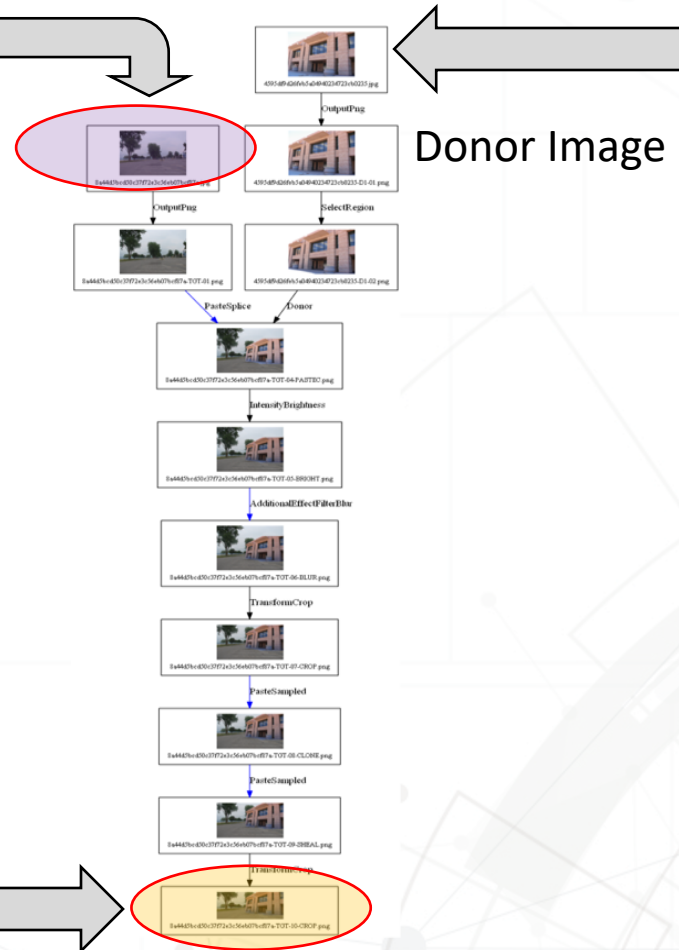
Base Image

High Provenance

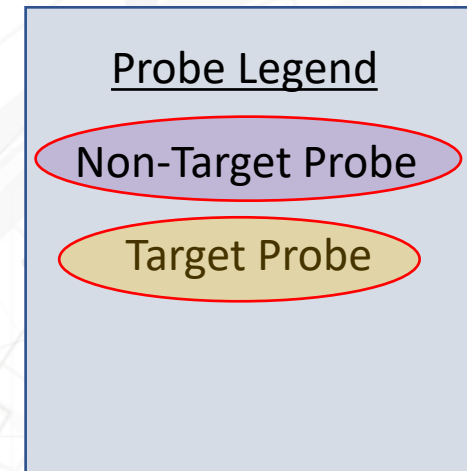


Donor Image

Unknown Provenance



Final Manipulated Image



Manipulation Journaling Tool (Extended Journal)

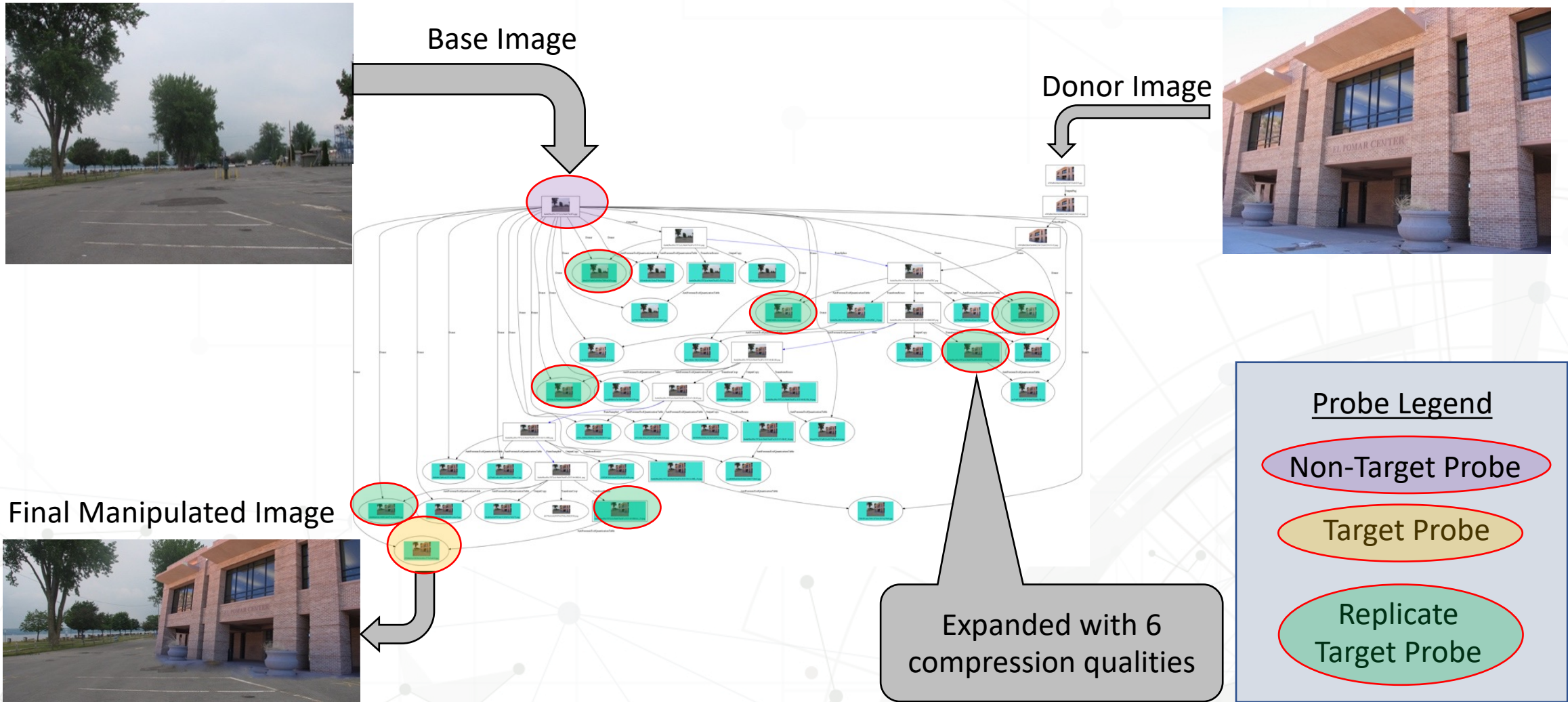


Image Mask for Manipulation Localization

- NC16, NC17 – single layer composite mask
- After MFC18 – multi-layer JPEG 2000 mask
 - Distinct manipulations are recorded in the different layers in JPEG 2000 mask file respectively.
 - Each bit in a byte for a pixel in a single-layer image represents one localizable manipulation.
 - Scoring can thus be extended to specific localizable manipulations in the image.



Manipulated Probe image

12/07-09/21



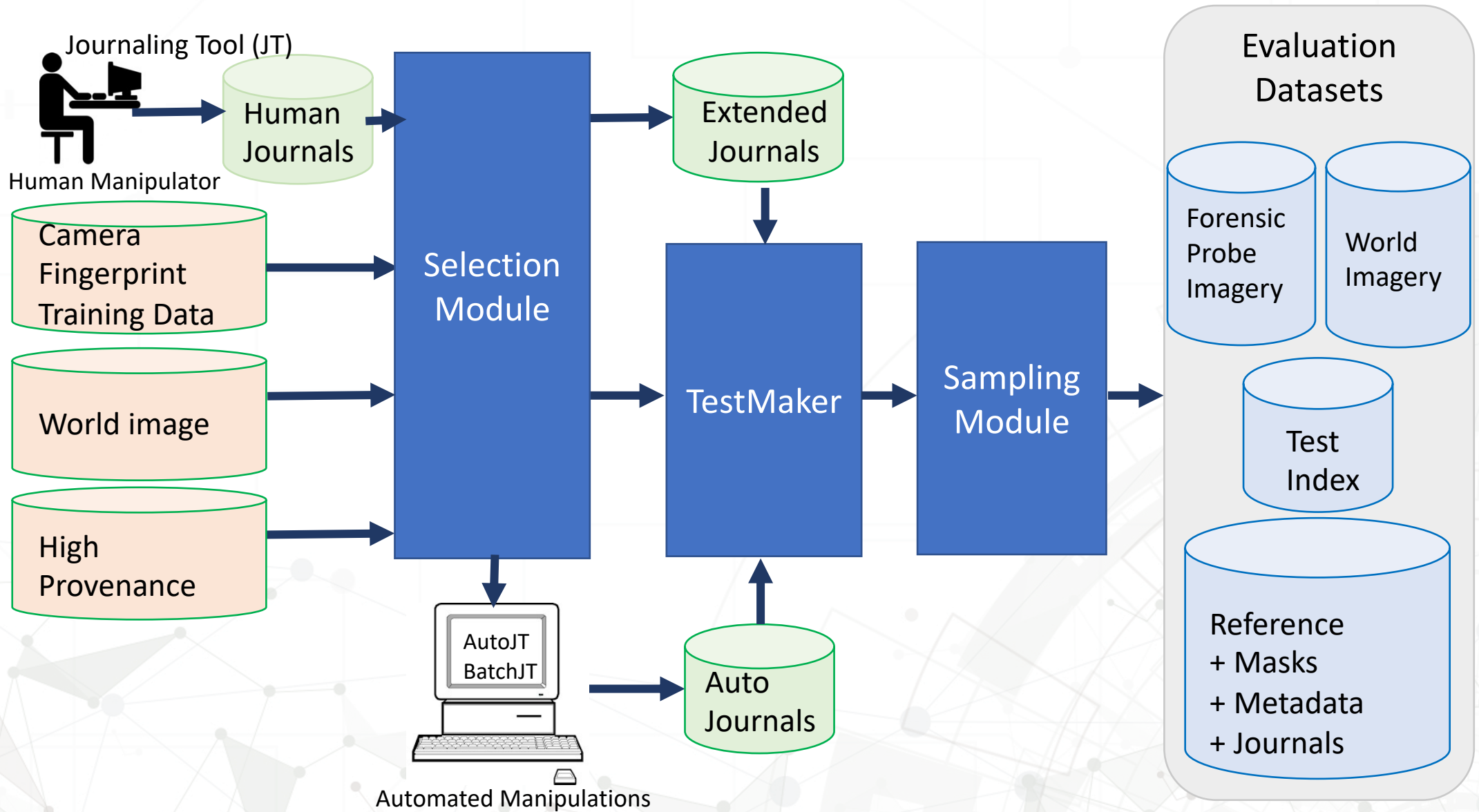
Composite mask

NIST OpenMFC 2020-2021

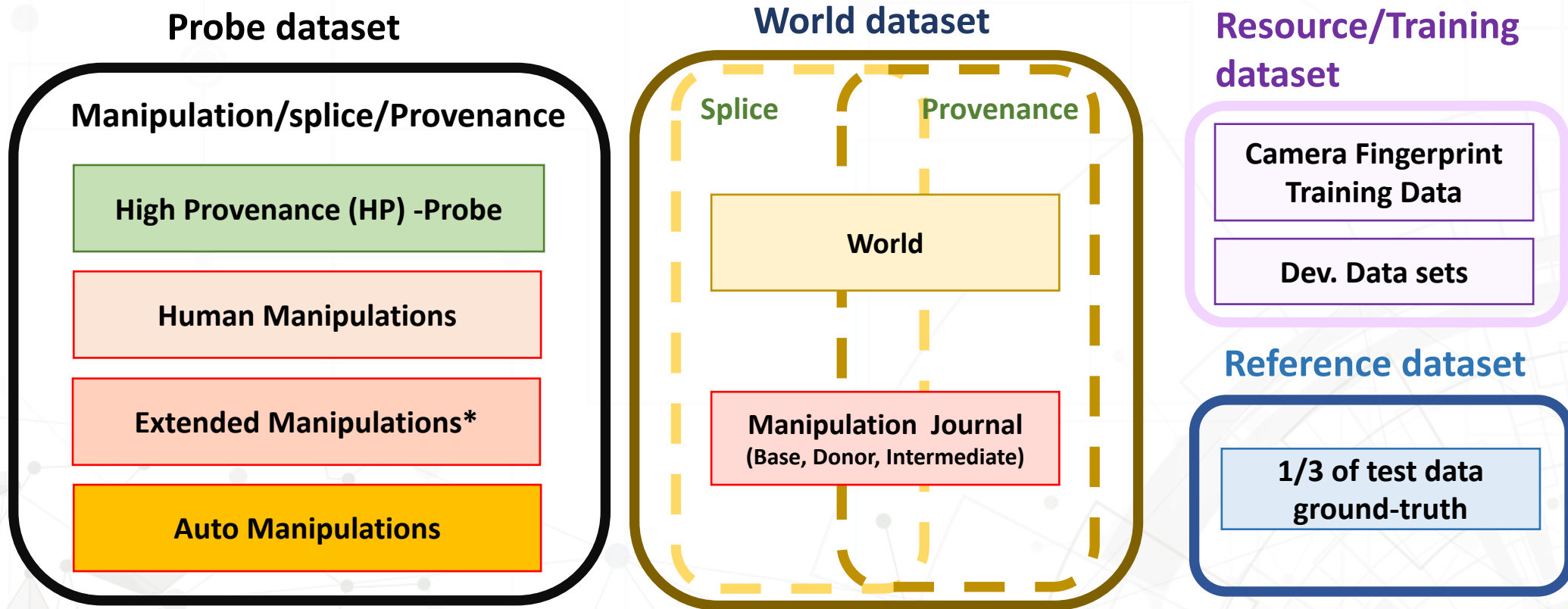
Sequence	Operation	Purpose	Color	Evaluated
5	ContentAwareFill	remove	Green	Y
4	PasteSampled	heal	Cyan	Y
3	PasteSplice	add	Orange	Y
2	Blur		Magenta	Y

An animated representation of the information stored by the JPEG2000. Every region is fully represented. The sequence is listed in descending order for node distance from the manipulated probe and may be distinct from the bit placement in the byte.

Evaluation Dataset Production Infrastructure

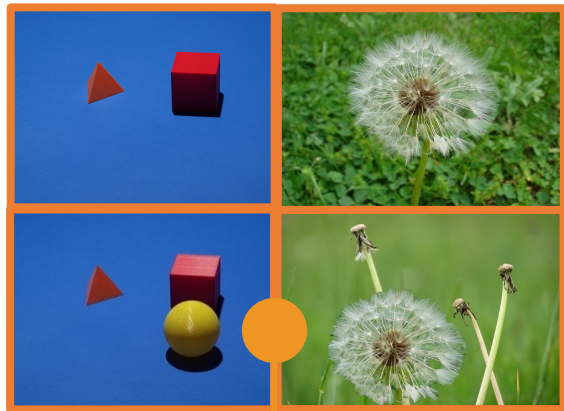


Data Collection and Evaluation Dataset Overview

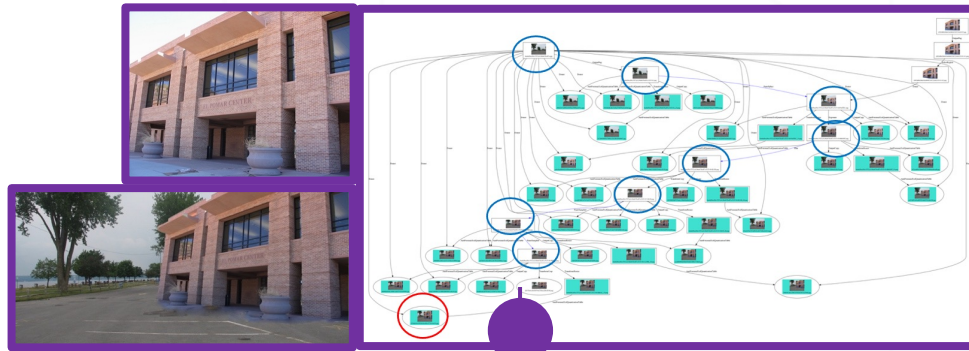


*Extended manipulations: the manipulated images generated by automatic machine manipulation tools in the extended journal described in the previous slide.

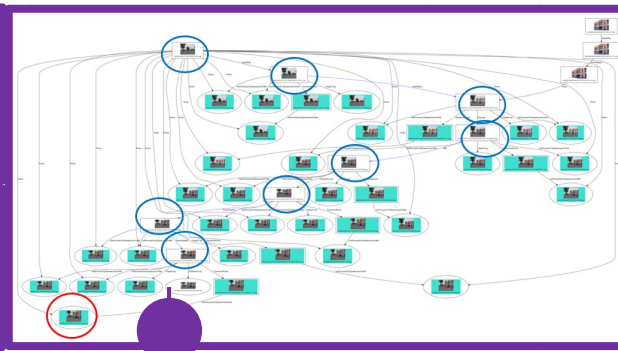
OpenMFC Evaluation Dataset (1)



NC16

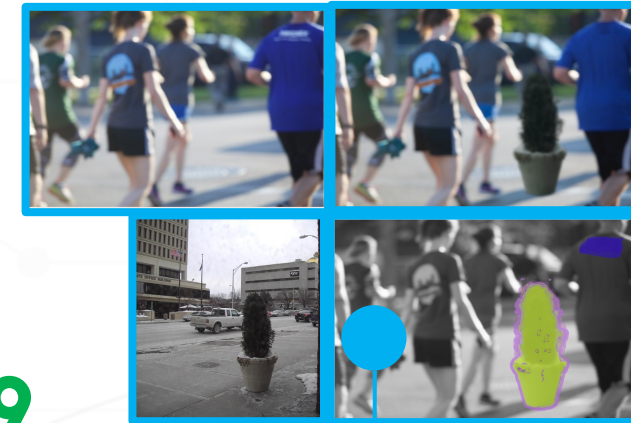


NC17



MFC18

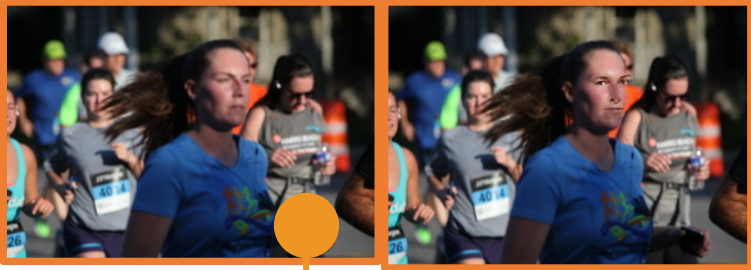
MFC19



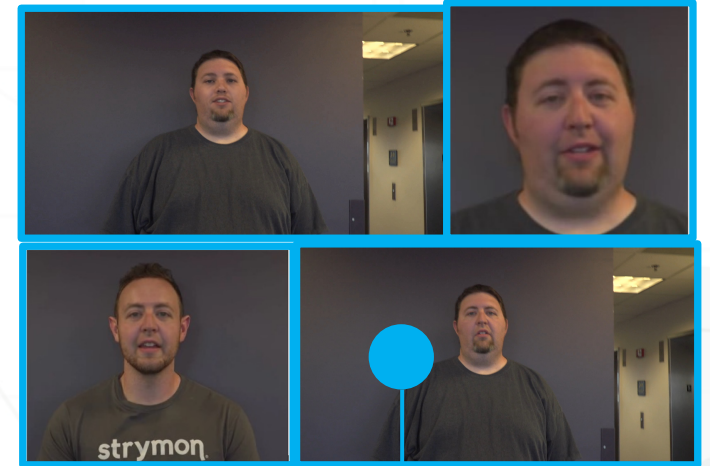
MFC20



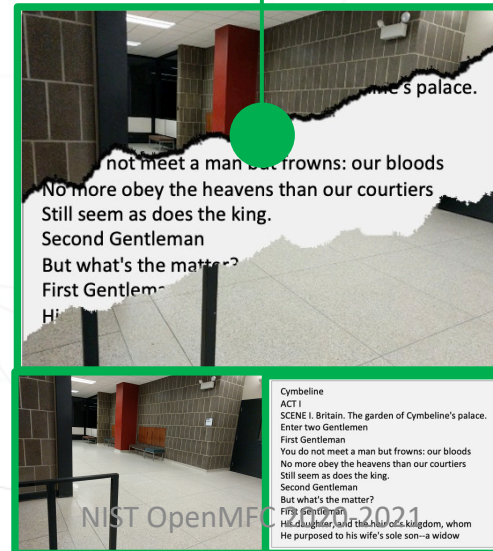
OpenMFC Evaluation Dataset (2)



OpenMFC 2022 StegD²



OpenMFC 2022 GAN Image¹



OpenMFC 2022 Deepfake Video¹

¹ Collaboration with Prof. Siwei Lyu's team

² Collaboration with Prof. Jennifer Newman's team

Large-Scale Benchmark Datasets

- Designed and built over 31 evaluation datasets and 13 development datasets to support 7 evaluation tasks and 2 evaluation challenges in MFC.
- Released the public datasets to over 600 individuals from over 200 organizations and 26 countries and regions worldwide (the number is continuously increasing).

Table: NIST released MFC datasets

(Highlighted: the evaluation sets; Grayed: the development sets)

MFC Released Dataset	Media Type	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
MFC19 EP1 Image	Image	Eval.	16K	1383	Mar-19
MFC18 GAN FULL Image	Image	Eval.	1.3K	267	Apr-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
MFC19 EP1 Video	Video	Eval.	1530	163	Mar-19
MFC18 GAN Video	Video	Eval.	118	19	Jun-18

NIST OpenMFC Resources: <https://mfc.nist.gov>

- MFC open evaluation datasets
 - Signup: NC16 Kickoff,
 - Signed two agreements:
 - NC17 Evaluation Part 1 (EP1),
 - MFC18 EP1,
 - OpenMFC 2020-2021 Evaluation dataset
 - MFC19 EP1 without ground-truth
- NIST OpenMFC leaderboard scoring server
 - <https://mfc.nist.gov>
- MediScore
 - Github: <https://github.com/usnistgov/MediScore/>
- Slack – open for public researcher
 - <http://openmfc.slack.com>
 - Discussion channel: <https://app.slack.com/client/T017MTH6RHT/C017MTH7LRK>
- Participant Google group – OpenMFC performer only
 - openmfc-performer
 - Mailing list: openmfc-performer@list.nist.gov

Takeaway

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

Questions?

OpenMFC team: mfc_poc@nist.gov

OpenMFC Evaluation Summary

- This talk:
 - OpenMFC program introduction
 - OpenMFC 2020-2021 evaluation tasks
 - OpenMFC evaluation metrics
 - OpenMFC evaluation datasets
- Talk in Day 2:
 - OpenMFC 2020-2021 evaluation results
 - OpenMFC evaluation infrastructure
- Talk in Day 3:
 - OpenMFC 2021-2022 tasks
 - OpenMFC 2021-2022 evaluation datasets

The background of the slide is a light gray with a complex, abstract pattern of thin lines and dots, resembling a network or circuit board. There are also larger, faint geometric shapes like rectangles and circles scattered across the background.

Thank You!



Feedback and Discussion:

OpenMFC Program

Haiying Guan
Yooyoung Lee and Lukas Diduch

Multimodal Information Group,
Information Access Division

OpenMFC2021 Workshop : Day 1, Tuesday, Dec. 7, 2021

Discussion Google Document

<https://docs.google.com/document/d/1ALXIkDz3c99rF4fVP1SUoQShsPx4ydcWJDW4Yv9-rk/edit?usp=sharing>

OpenMFC Program Discussion Topics

1. How OpenMFC can improve to promote the research being done in Media Forensics?
2. What are the bottlenecks to research in developing media forensic analytics systems?
3. What existing resources (data/tools) are useful for inclusion in the OpenMFC future evaluation dataset?
4. What other topics/tasks would be useful for future OpenMFC evaluation?
5. What metadata/reference ground-truth would be helpful in furthering your research?
6. What other evaluation metrics would you suggest for the future OpenMFC evaluation?
7. How to design the evaluation to adapt to dynamically changed AI technologies?
8. What is the future of Media Forensics in 5 years, 10 years?

Questions?

OpenMFC team: mfc_poc@nist.gov

Thank You!



NIST Open Media Forensics Challenge

OpenMFC 2020 -2021 Results

Haiying Guan

Yooyoung Lee, Lukas Diduch, and Ilia Ghorbanian

Multimodal Information Group,
Information Access Division, ITL, NIST

OpenMFC2021 Workshop : Day 2, Wednesday, Dec. 8, 2021



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.
- All images, graphs, and charts are original works created for DARPA MediFor as well as other NIST Programs.

Acknowledgement

- NIST contributors
 - Jonathan Fiscus
 - Timothee Kheyrkhah
 - Peter Fontana
 - Jesse G. Zhang
- External collaborators:
 - Prof. Siwei Lyu in University at Buffalo
 - Prof. Jennifer Newman, Prof. Yong Guan, Li Lin from Iowa State University, and Prof. Roy A. Maxion from Carnegie Mellon University

Outline

- OpenMFC 2020-2021 Evaluation Dataset
- OpenMFC 2020-2021 Result and analysis
 - IMDL
 - VMD
 - IGMD
 - VGMD

OpenMFC 2020-2021 Evaluation Tasks

- Image Manipulation Detection and Localization (IMDL)
 - To detect if the image has been manipulated, and then to spatially localize the manipulated region
- Video Manipulation Detection (VMD)
 - To detect if the video has been manipulated
- Image GAN Manipulation Detection (IGMD)
 - To detect GAN-manipulated images (e.g., created by a GAN model, locally/globally modified by a GAN filter/operation, etc.)
- Video GAN/Deepfakes Manipulation Detection (VGMD)
 - To detect GAN/Deepfakes manipulated videos

Six OpenMFC 2020-2021 Leaderboards

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

- Image Manipulation Detection and Localization:
 - Image Only (IMDL-IO)
 - Image + Metadata (IMDL-IM)
- Video Manipulation Detection:
 - Video Only (VMD-VO)
 - Video + Metadata (VMD-VM)
- GAN Manipulation Detection:
 - Image Only (IGMD)
 - Video Only (VGMD)

OpenMFC 2020-2021

Evaluation Datasets

- OpenMFC 2020-2021 IMDL: MFC19 EP1 Image
- OpenMFC 2020-2021 VMD: MFC19 EP1 Video
- OpenMFC 2020-2021 IGMD: MFC18 GAN Challenge Image
- OpenMFC 2020-2021 VGMD: MFC18 GAN Challenge Video

Concepts in the Data Annotation

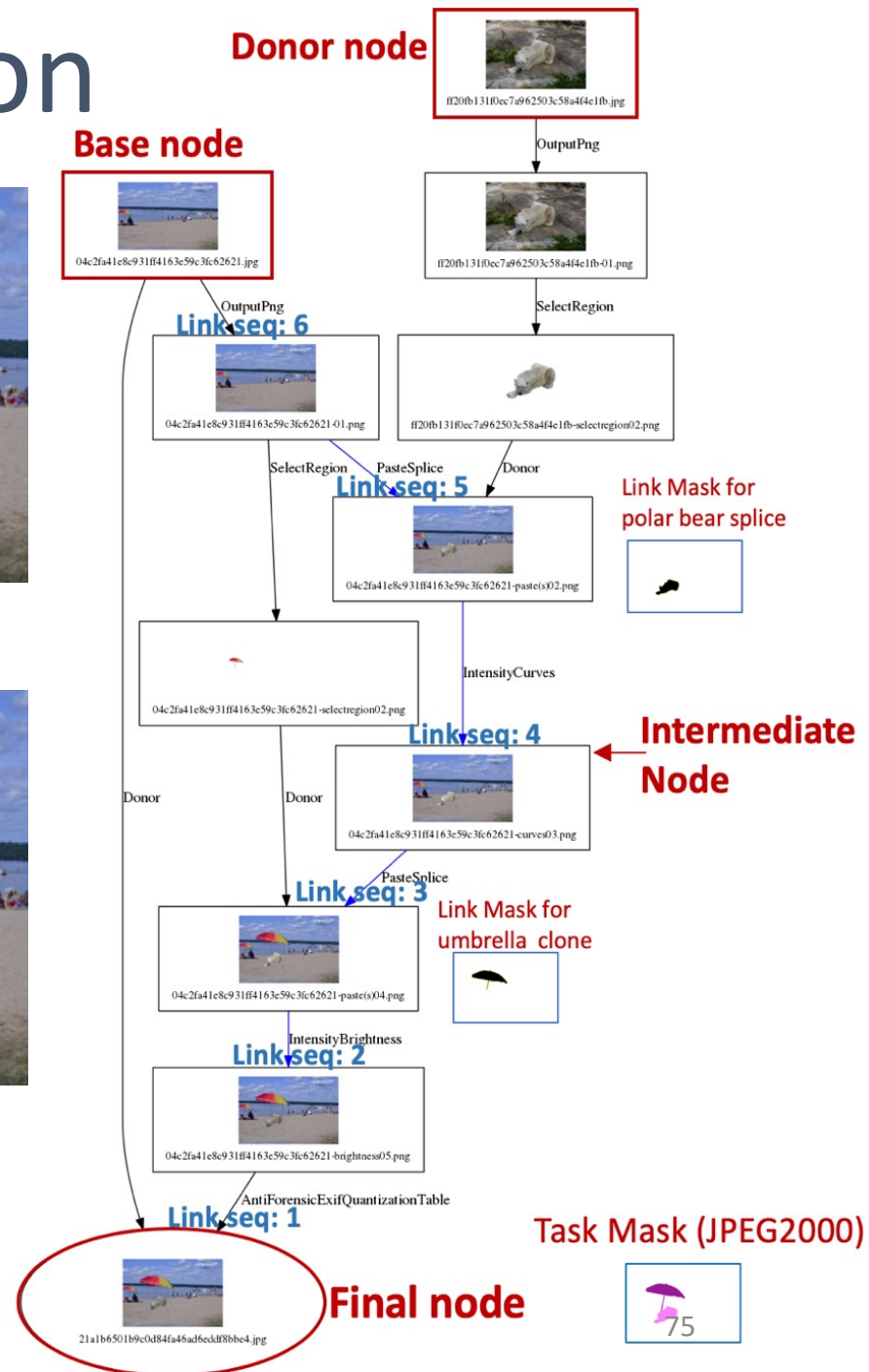
- Journal
 - Directed graph
- Node
- Link
- High Provenance (HP)
 - Original image
- Manipulated image
- Base Image – start node
- Final Image – end node
- Donor Image – donor node
- Probe – test sample



Manipulated image



Original image/Base Image



An Image Journal example



(a) Base Image



(b) Donor Image 1

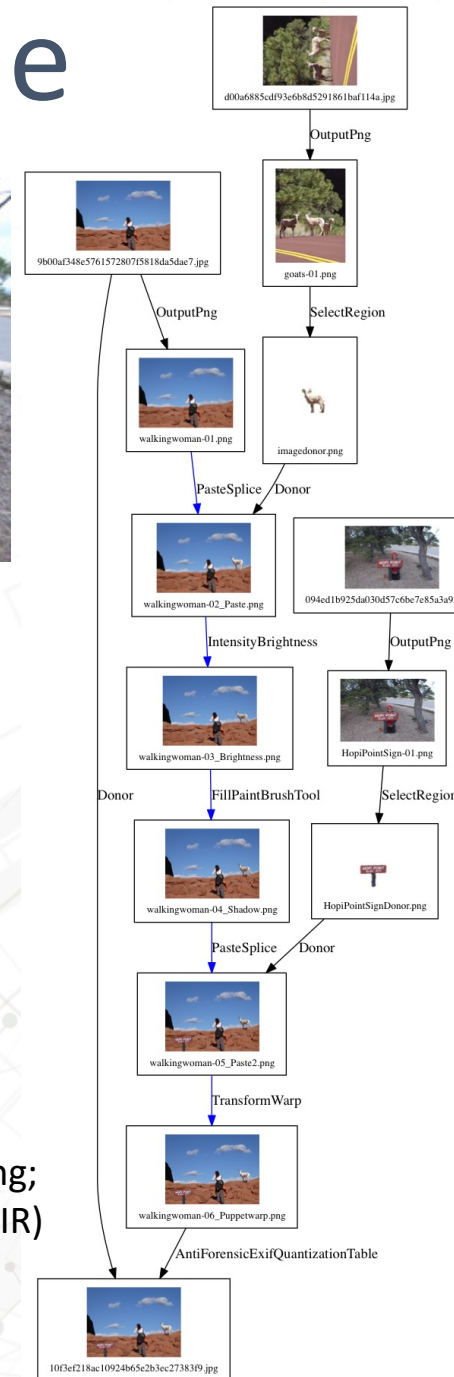


(c) Donor Image 2

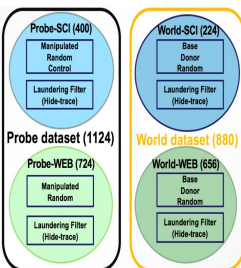
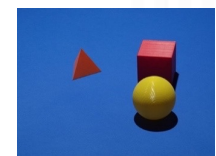


(d) Final manipulated Image

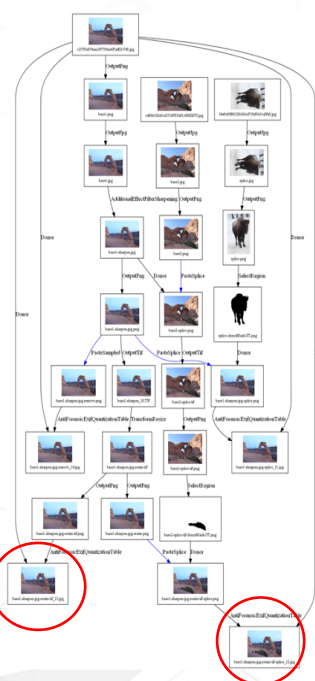
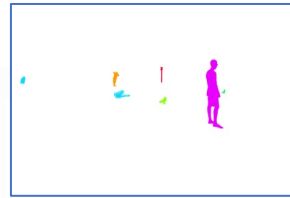
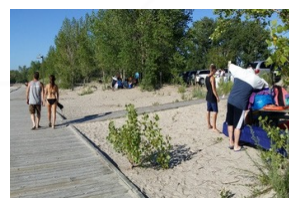
User Guide for NIST Media Forensic Challenge (MFC) Datasets, Guan, Haiying; Delgado, Andrew P.; Lee, Yooyoung; Yates, Amy N.; Zhou, Daniel F., Kheyrkhah, Timothee; Fiscus, Jonathan G., NIST Interagency/Internal Report (NISTIR) Number 8377, July 2021, Available at <https://doi.org/10.6028/NIST.IR.8377>.



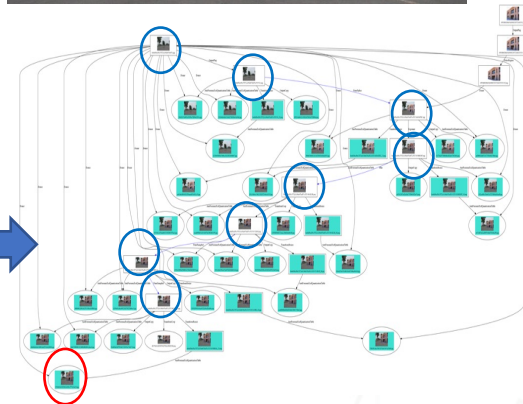
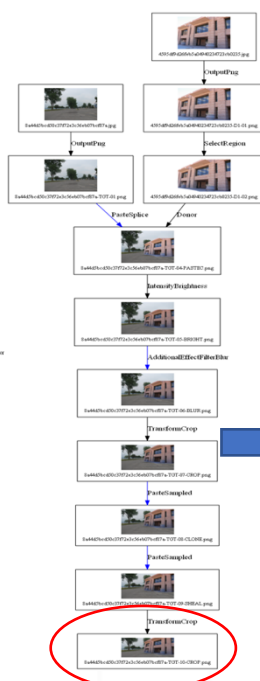
Cross-Year comparisons: MFC Evaluation Dataset History



**Kick-off
2016 Dataset**

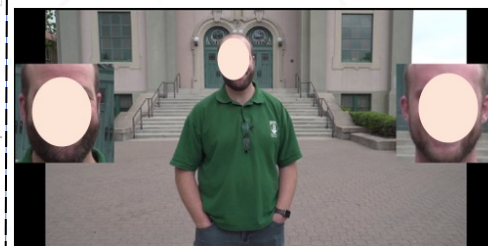
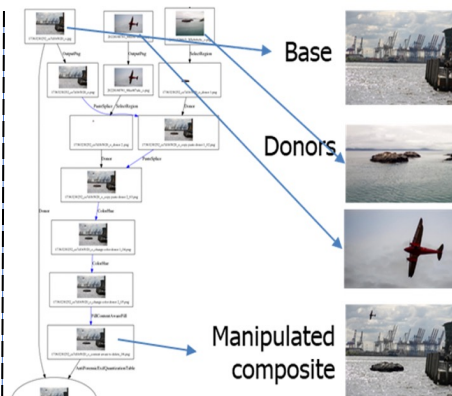


**Provenance
Auto Journaling Tool (JT)
Nimble Challenge 2017**



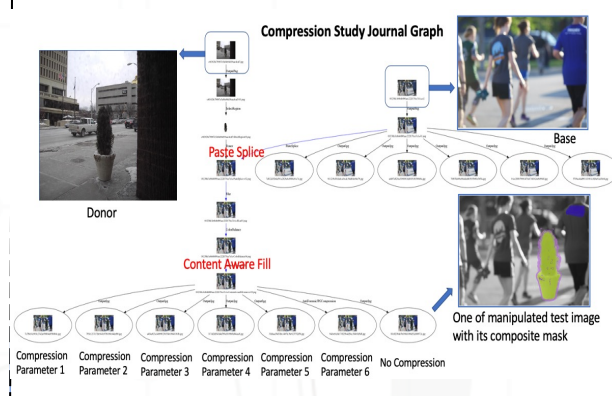
**New Manipulations
(CGI, Recapture, ...)
Extended JT, AutoJT**

MFC 2018



- Camera ID Eval. datasets
- Video Temporal Spatial
- Additional Manipulation Operations (GAN etc.)
- Extended JT, AutoJT

MFC 2019



Special study data

- Compression
- Global Blur
- Single Operation
- Social Media Laundering
- Frame Drop/Dup.

MFC 2020

OpenMFC IMDL Evaluation Dataset Property (1)

- MFC19 Image EP1: 75 manipulation operations
- Unique journal link count – no duplication count using link

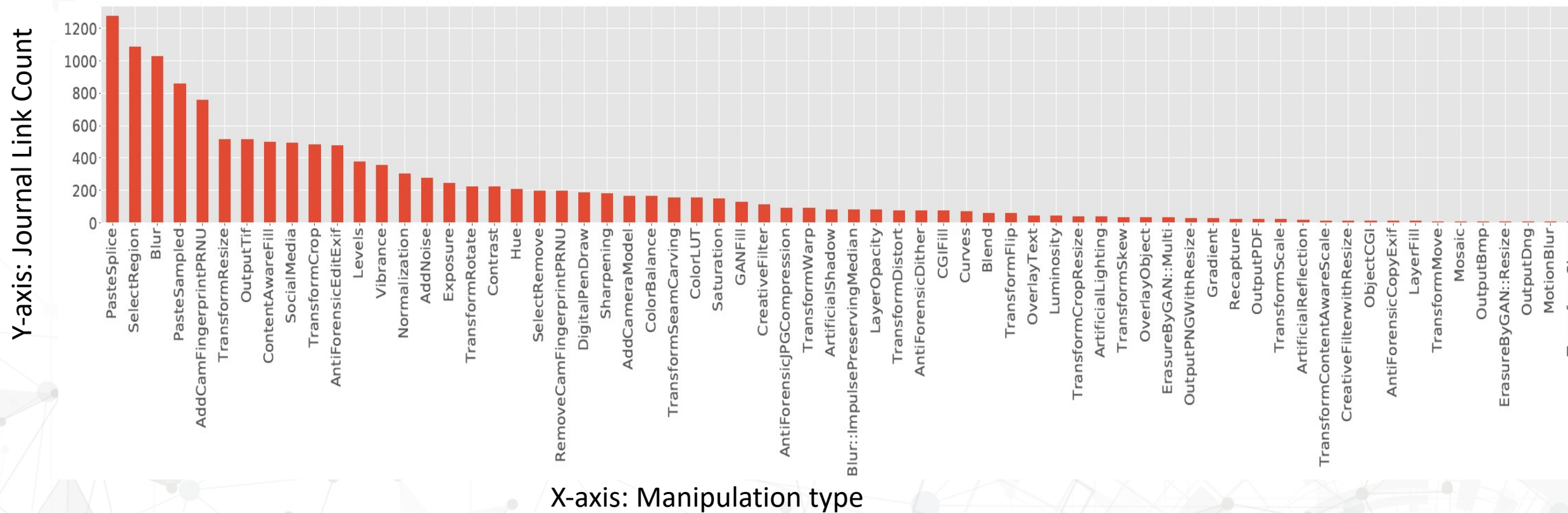


Figure: OpenMFC 2020-2021 image eval. operation histogram of unique journal link (partial)

OpenMFC IMDL Evaluation Dataset Property (2)

- MFC19 Image EP1: 75 manipulation operations
- Test case (probe) count – has duplication counts using probe

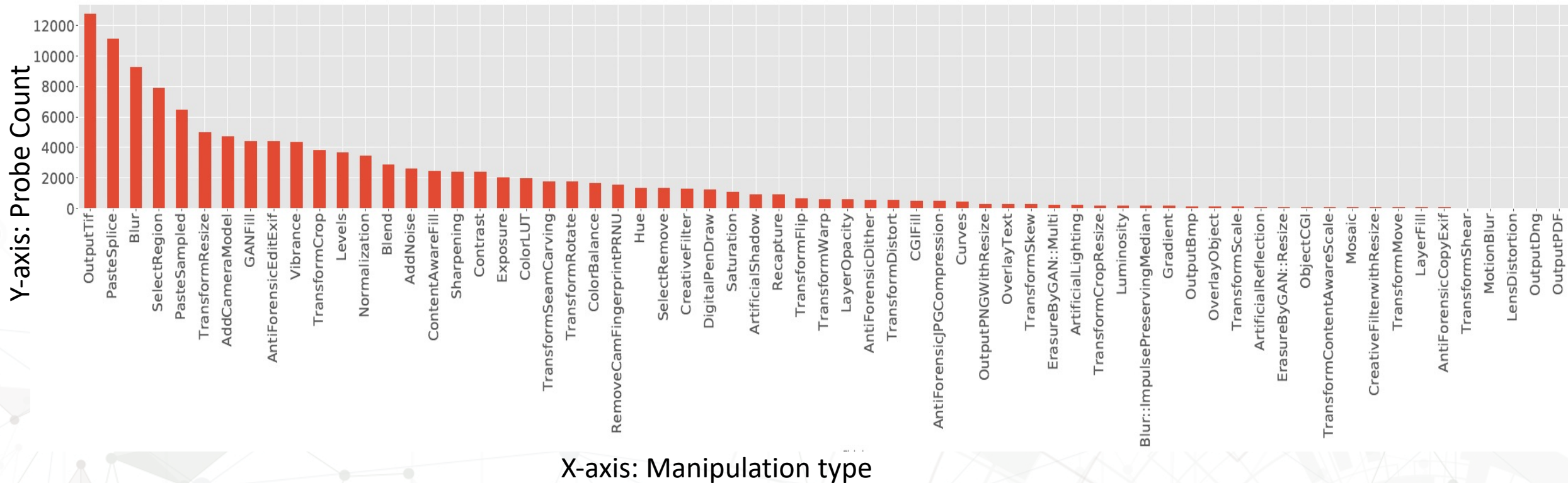


Figure: OpenMFC 2020-2021 image eval. operation histogram of probe count (partial)

OpenMFC 2020-2021 Evaluation Dataset Summary

Table: NIST released OpenMFC datasets

Task(s)	NC16 Kickoff	NC17 EP1	MFC18 EP1	MFC19 EP1	MFC20 EP1
Image	1.1K	4K	17K	16K	20K
Video	-	0.36K	1K	1.5K	2.5K
GAN Image	-	-	1.3K	-	-
GAN Video	-	-	118	-	-

Red Color: indicates the datasets used for in the OpenMFC 2020-2021 evaluation.

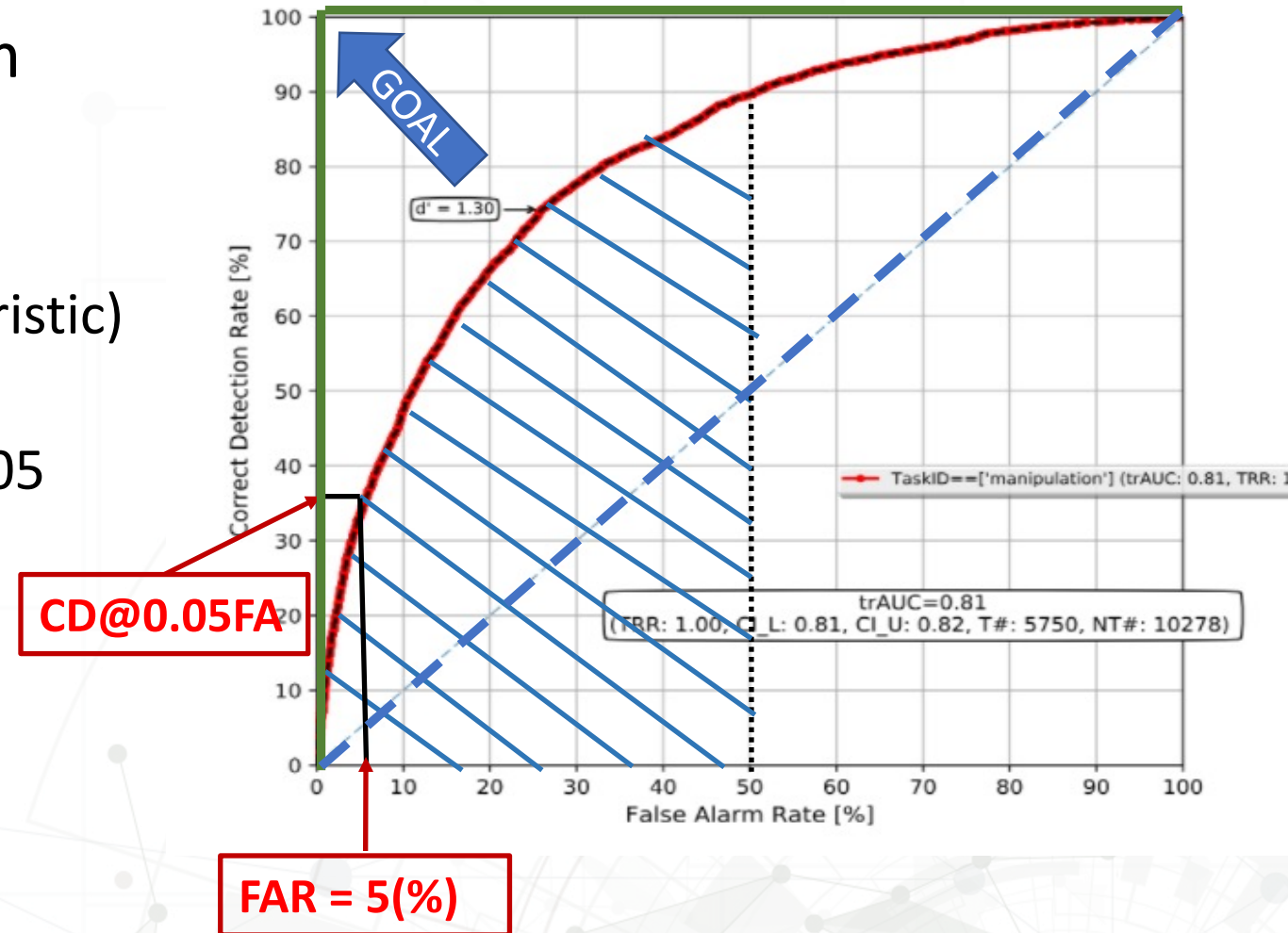
Task:

Image Manipulation Detection and Localization (IMDL)

- IMDL: to detect if the image has been manipulated, and then to spatially localize the manipulated region
- Conditions:
 - Image Only (IO)
 - Image and Metadata (IM)
- Outline:
 - MediFor MFC results
 - OpenMFC participant submission file
 - OpenMFC results
 - Analysis

Detection System Evaluation Metrics

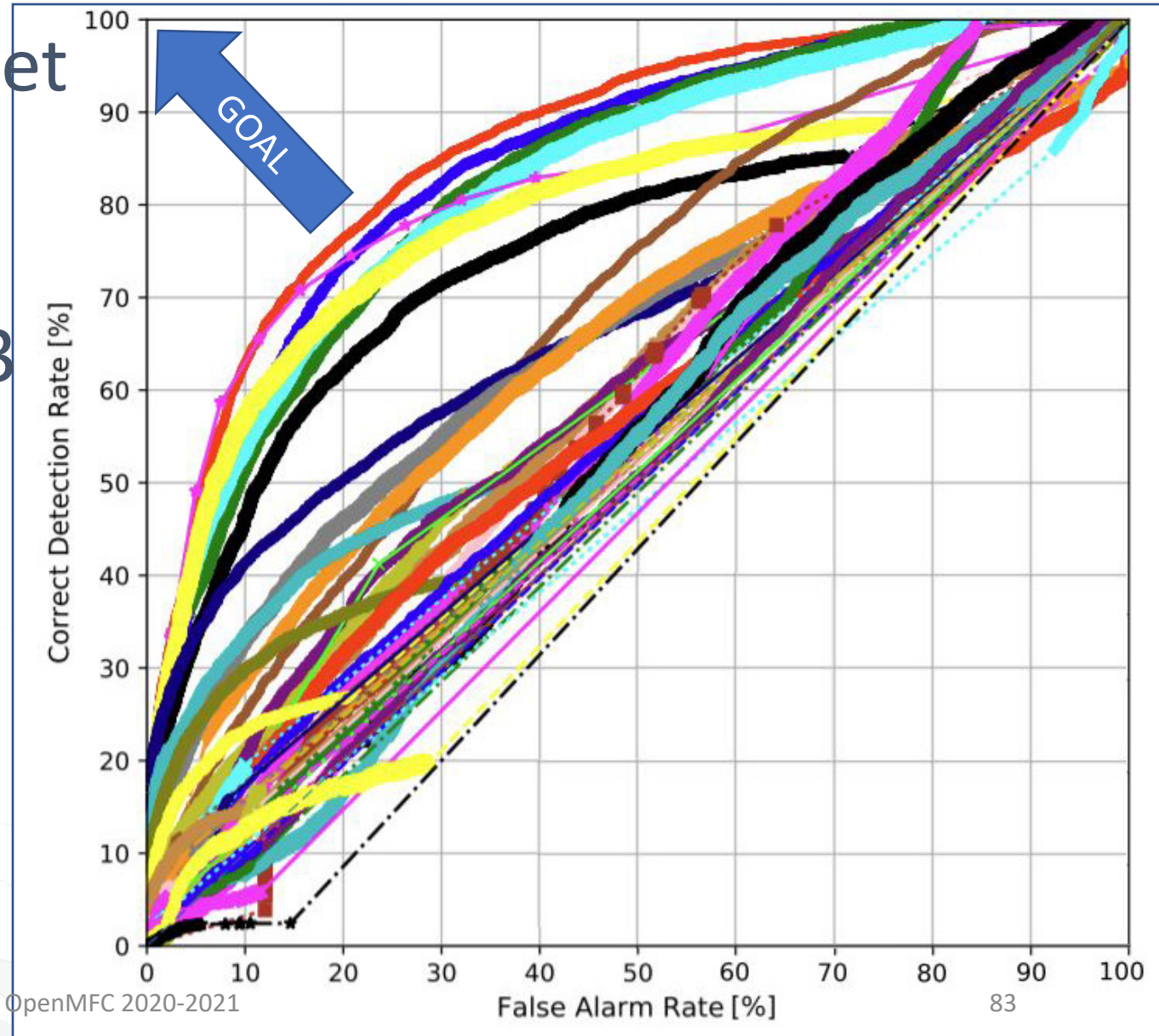
- Evaluate the accuracy of a system output (e.g., confidence score)
- Evaluation metrics
 - ROC (Receiver Operating Characteristic)
 - AUC (Area Under Curve)
 - CD (Correct Detection) @ FAR = 0.05



MFC19 IMD-IO Results

Figure: IMD on MFC19 EP1, full data

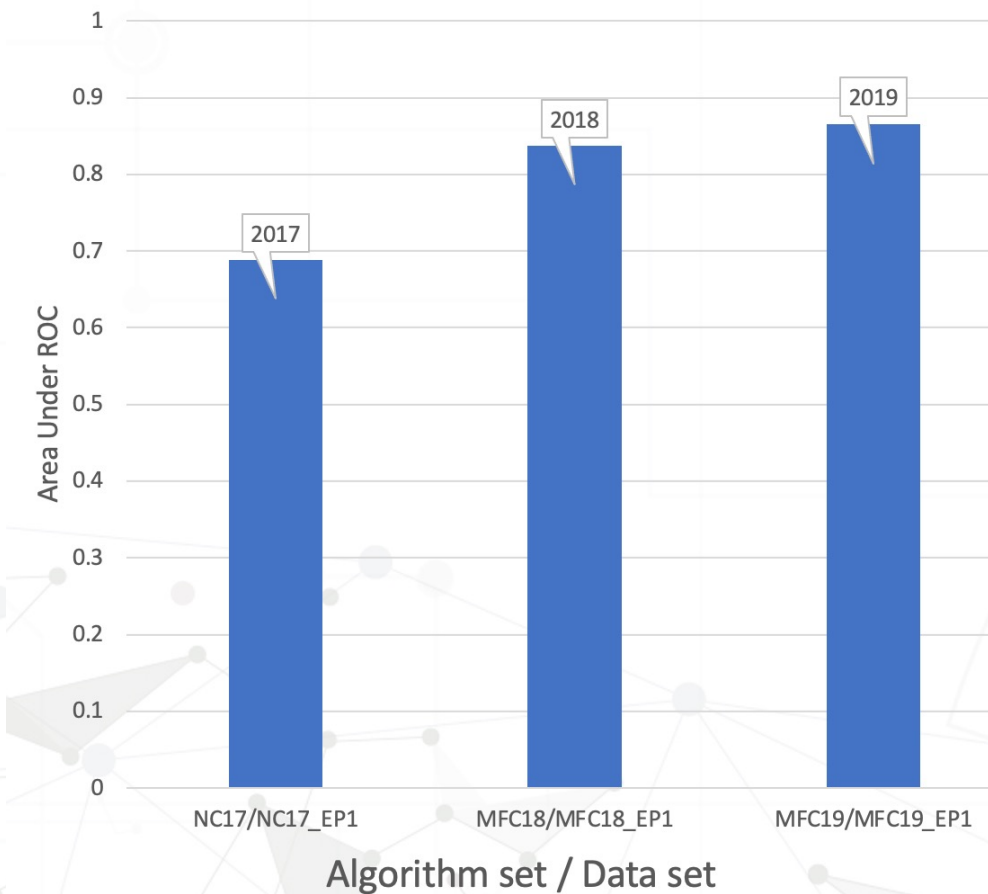
- OpenMFC 2020-2021 dataset
 - MFC19 EP1 dataset
 - 16K probe images
- 47 analytic systems from 13 teams
 - Condition: Image Only (IO)
 - Highest AUC: 0.866;
CD@FAR=0.05: 0.456



MFC Cross Year IMD-IO Performance Comparison

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

Image Manipulation Detection Performance



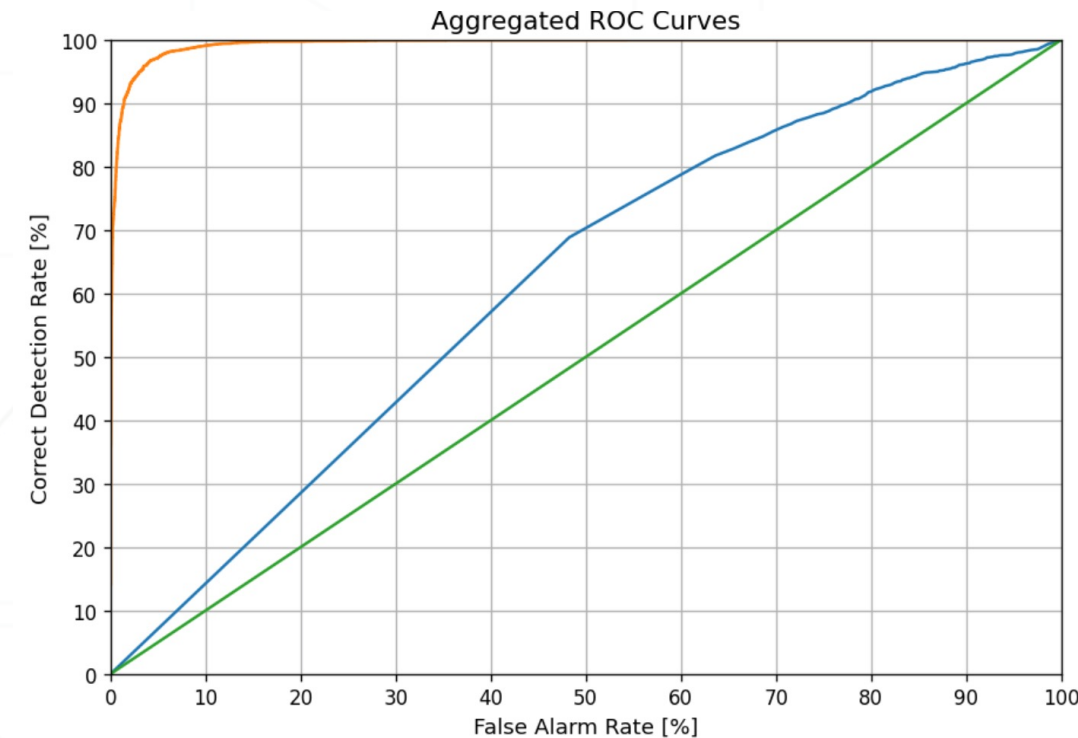
NIST OpenMFC Submission File: IMDL

- Task: IMDL
- Submission file name format:
 - anyfilename.tar.gz
- e.g., OpenMFC20_Image_DJpegDetection_TeamID_Date.tar.gz
 - 16030 rows
 - Example: the first several rows of the submission file
 - ProbeFileID|ConfidenceScore|OutputProbeMaskFileName|ProbeStatus|ProbeOptOutPixelValue
 - 00003e6a1efc7022da825396dc680343|0.29023||Processed|
 - 000d547cd11a767dab2b9bf56991bb0c|0.332853||Processed|
 - 0010d95b9edb40516bf13973cc2b9dbf|0.776651||Processed|
 - 001631ab2323cda12f9eae60c8daf2a8|0.221767||Processed|
 -

NIST OpenMFC Leaderboard: IMDL-IO

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

RANK	SUBMISSION ID	SUBMISSION DATE	TEAM NAME	SYSTEM NAME	AUC	CDR@0.05FAR	AVERAGE OPTIMAL MCC
1	63	2021-06-07 11:26:58	Mayachitra	test1june6	0.993707	0.972	
2	10	2020-11-05 21:53:02	UIIA	naive-efficient	0.616186	0.071351	
3	67	2021-06-08 00:51:16	UIIA	testIMDL	0.5	0.05	
4	81	2021-06-26 00:31:16	UIIA	testIMDL			0.0553688699305928



- For the IMDL-IO detection task, the best public team's AUC is 0.99 compared with 0.87 in MFC19.
- For the IMDL-IO localization task, the best public team's MCC score is 0.055 compared with 0.224 in MFC19.

OpenMFC Performance: Testing vs. Training

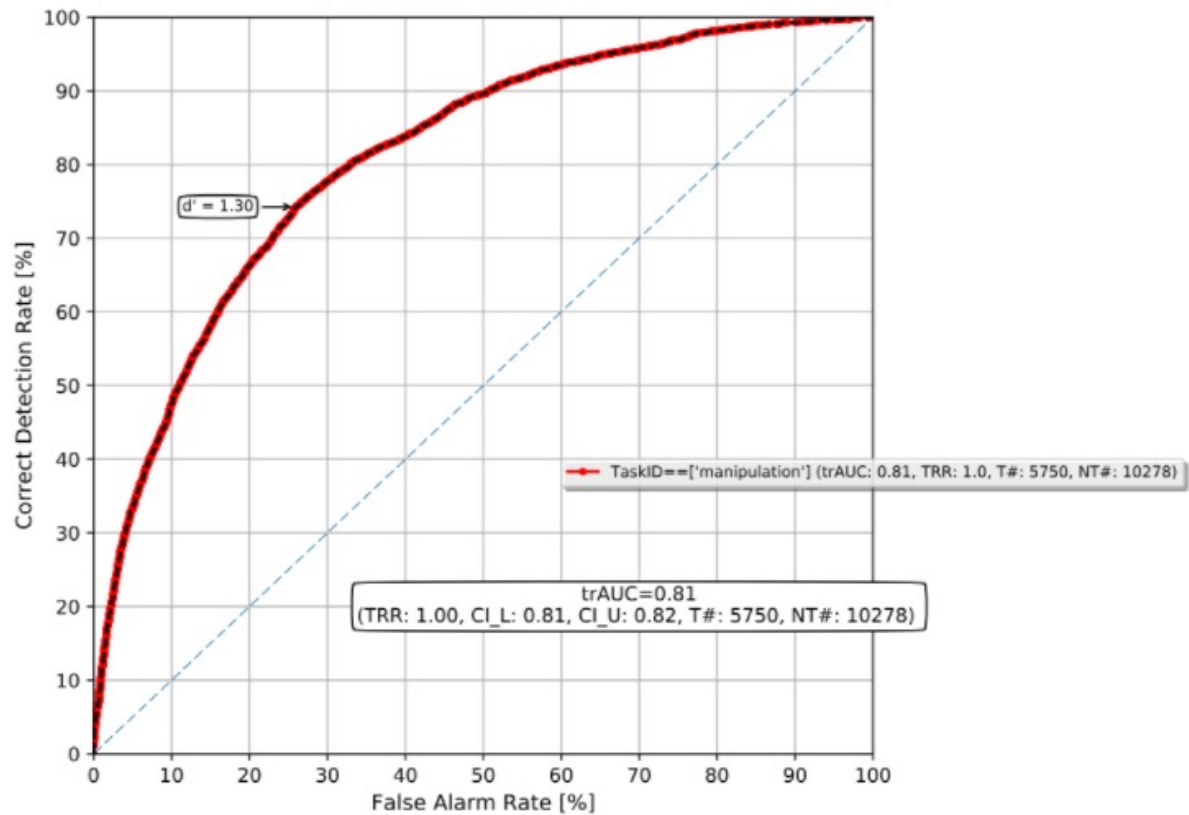


Figure: Team1 system performance without using the MFC19 dataset as training.

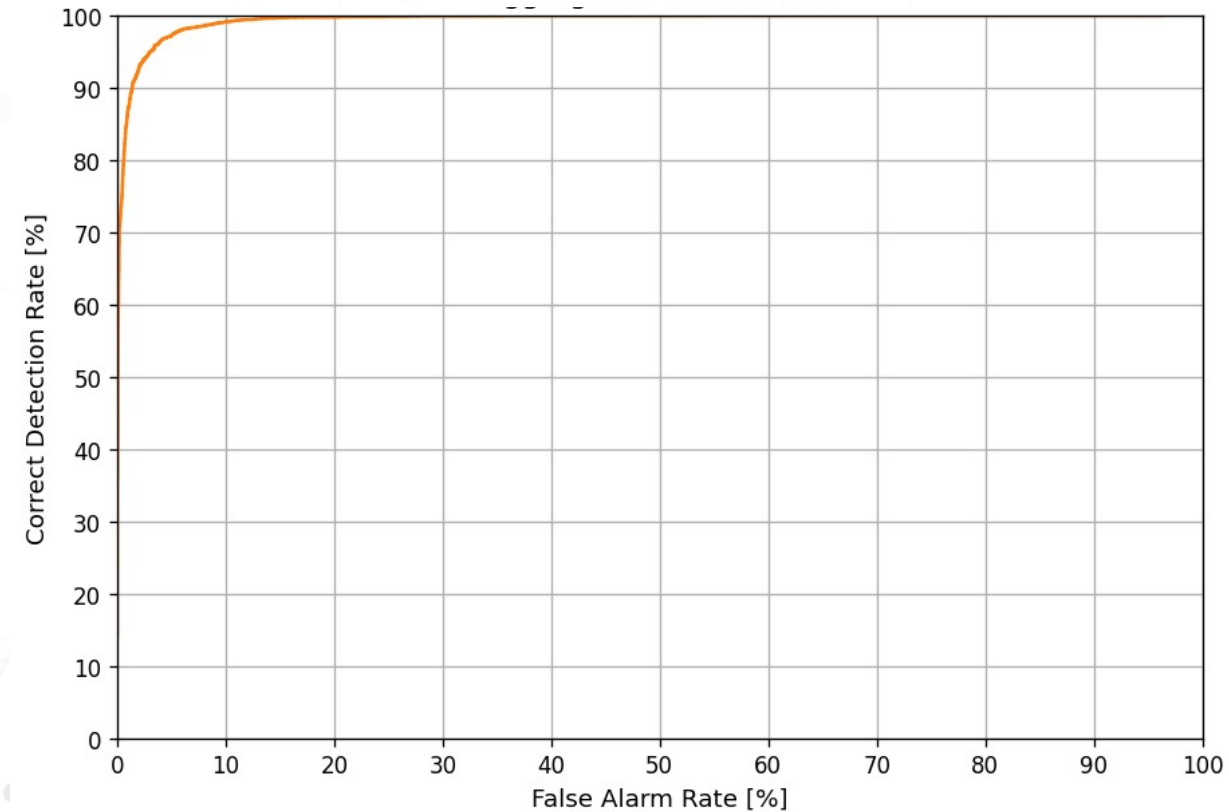


Figure : Team1 system performance using the MFC19 dataset as training.

OpenMFC Blackbox Evaluation Performance: Public Team

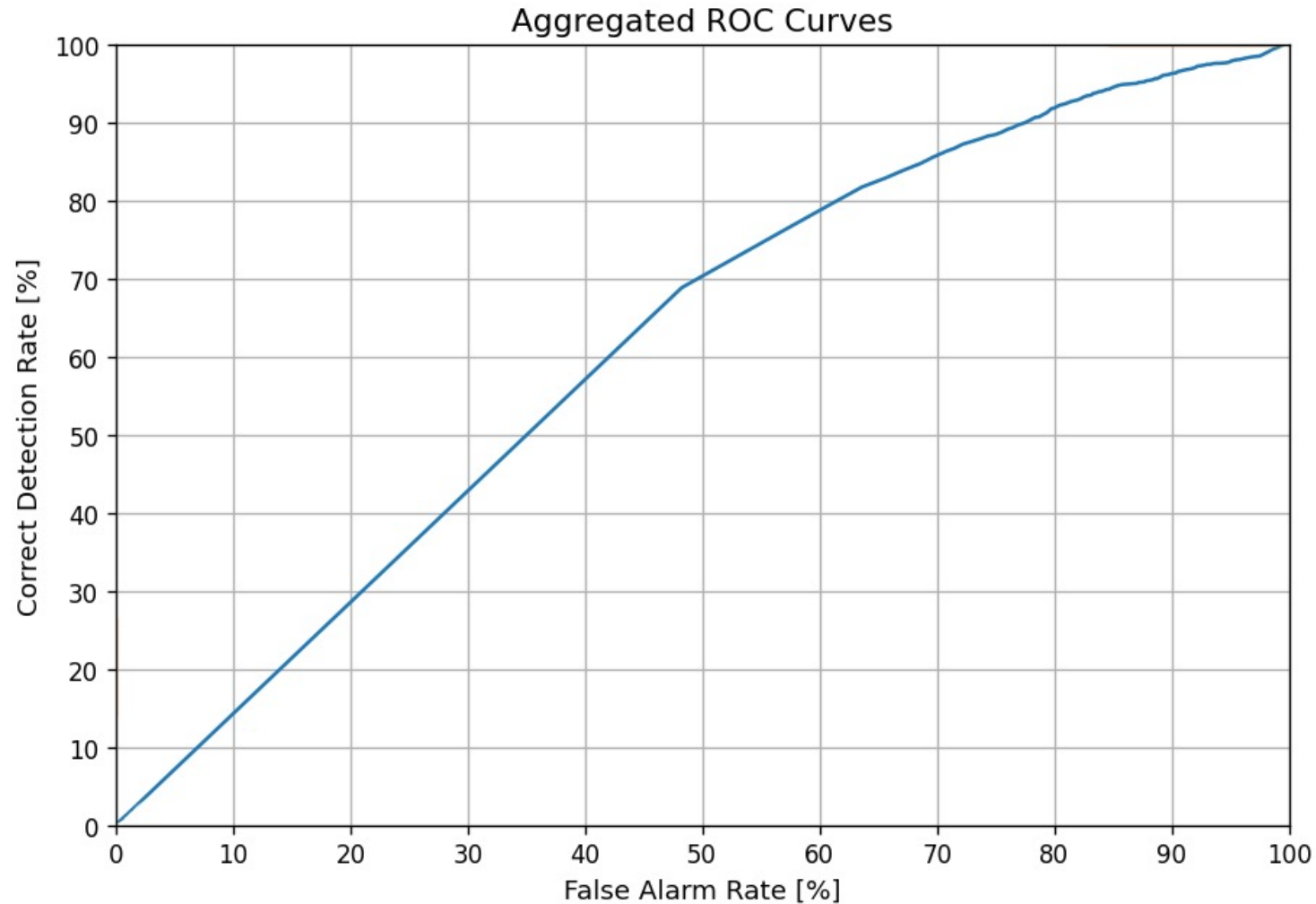


Figure: Team2 system performance without using MFC19 dataset as training.

Task:

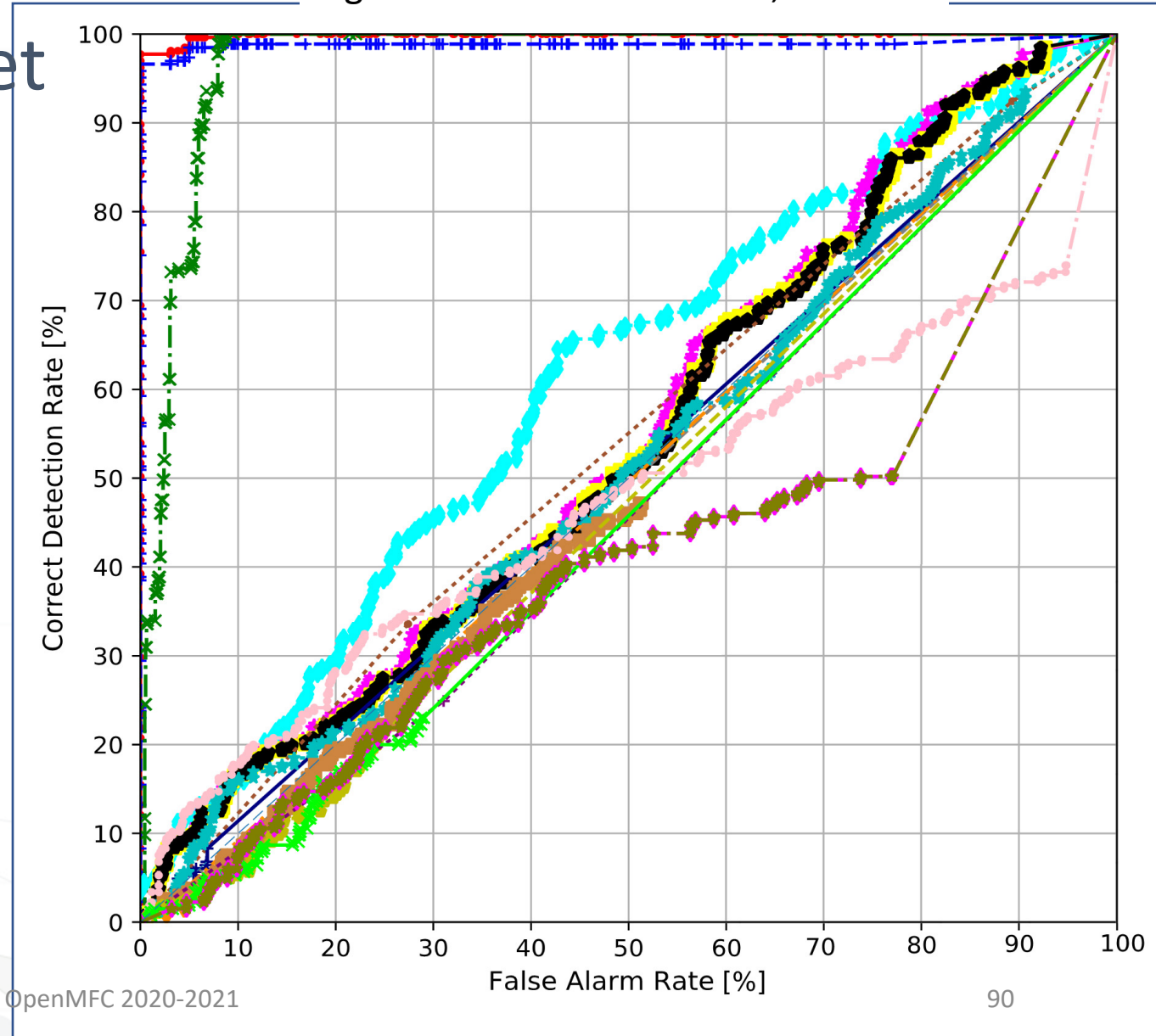
Video Manipulation Detection (VMD)

- VMD: to detect if the video has been manipulated
- Conditions:
 - Video Only (VO)
 - Video and Metadata (VM)
- Outline:
 - MediFor MFC results
 - OpenMFC Participant submission file
 - OpenMFC results

MediFor MFC19 Video Manipulation Detection Results

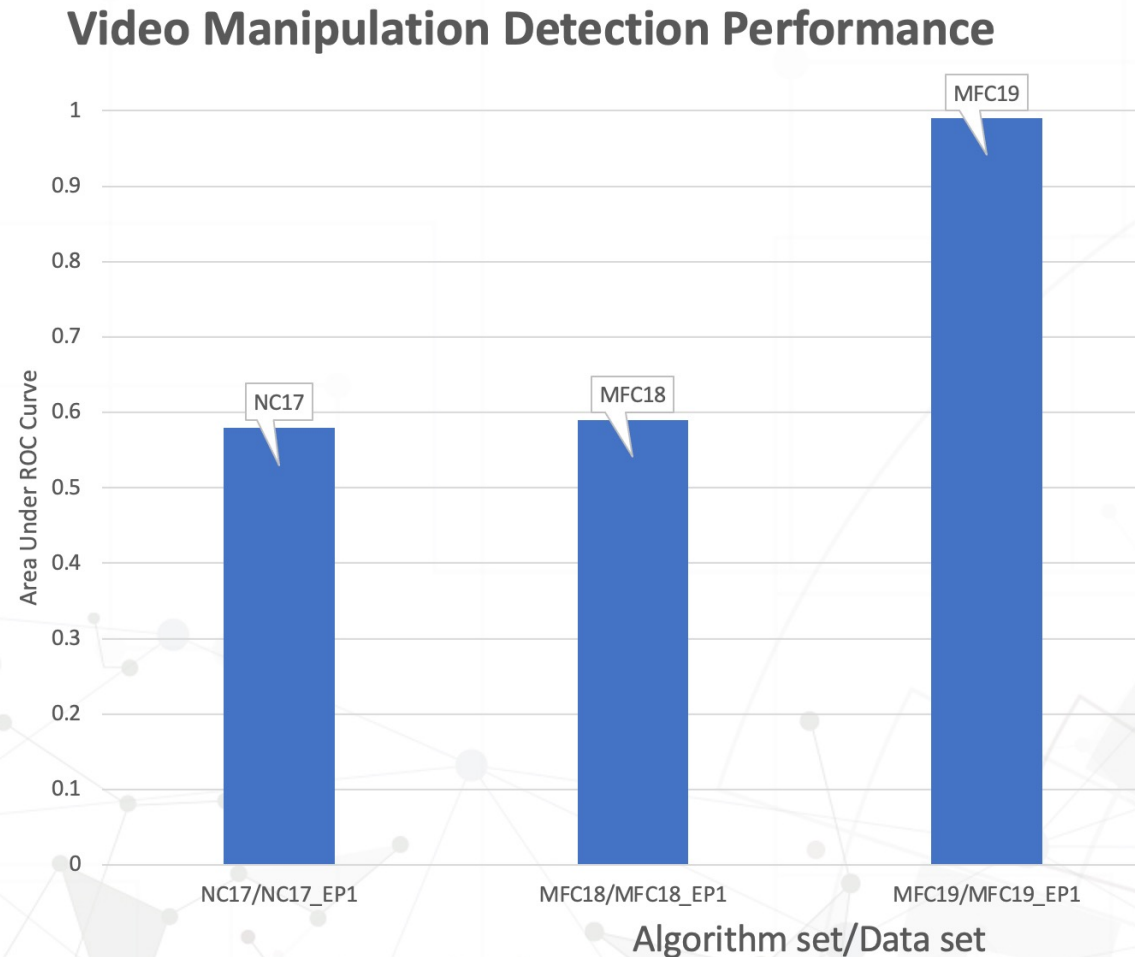
- OpenMFC 2020-2021 dataset
 - MFC19 EP1 dataset
 - 1.5K probe videos
- 19 analytic systems from 8 teams
 - Condition:
 - Video Only (VO)
 - Video + metadata (VM)
 - Highest AUC: 0.999;
CD@FAR =0.05: 0.997

Figure: VMD on MFC19 EP1, full data



MediFor MFC Year-to-year VMD Performance Comparison

- Not apples to apples comparison: dataset complexity also increases each year



OpenMFC Submission Files: VMD

- Task: VMD
- Submission file name : e.g., OpenMFC20_Video_SystemID_TeamID_Date.tar.gz
 - 1531 rows
 - Example: the first several rows of the submission file
 - ProbeFileID|ConfidenceScore|ProbeStatus
 - 001bd1016363c47079a6165535ae7145|0.136152726|Processed
 - 00a68871d70ada52a8ef2be33df178b9|0.633048435|Processed
 - 0109fcabf33f12ab5e947afb3950c602|0.535061206|Processed
 -

Task:

Image GAN Manipulation Detection (IGMD)

- IGMD: To detect GAN-manipulated images (e.g., created by a GAN model, locally/globally modified by a GAN filter/operation, etc.).
- Conditions:
 - Image Only (IO)
- Outline:
 - IGMD evaluation Dataset
 - MediFor MFC results
 - OpenMFC Participant submission file
 - OpenMFC results

IGMD Datasets

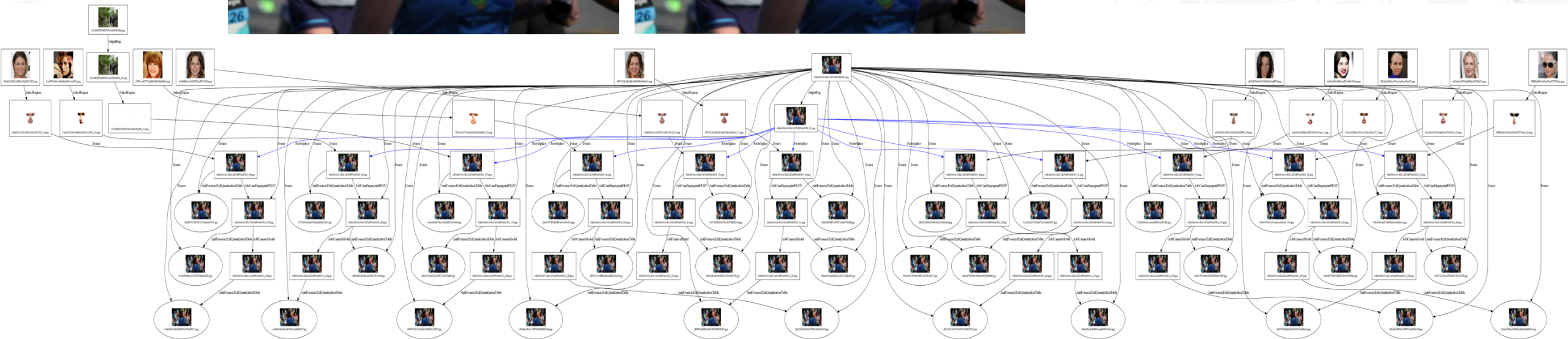
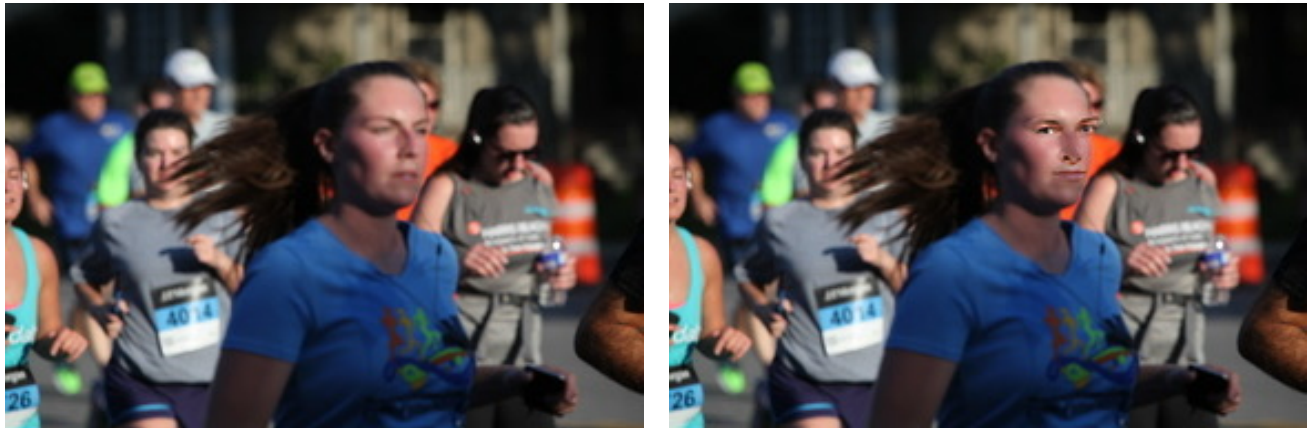
- Manipulations in the MFC18 Image GAN evaluation dataset:
 - Face swap – GAN vs. real face
 - GAN Fill, GAN Erase
 - CFA camera model
- Reference definition
 - Manipulated by GAN operation (with/without other operations) – IsTarget = ‘Y’
 - Manipulated by other operations (not GAN) – IsTarget = ‘Y’
 - NotManipulated – IsTarget = ‘N’

NIST Data Sets	Major operations	Journals	Probe	Date
MFC18 Eval GAN Image Full (total 1340 probes)	CFA Camera Model	89	191	06/30/2018
	ErasureByGAN	86	200	
	Face Swap	37	640	
	GAN Fill	55	110	
	World Image	-	199	

GAN Full Image Dataset: A face swap journal example

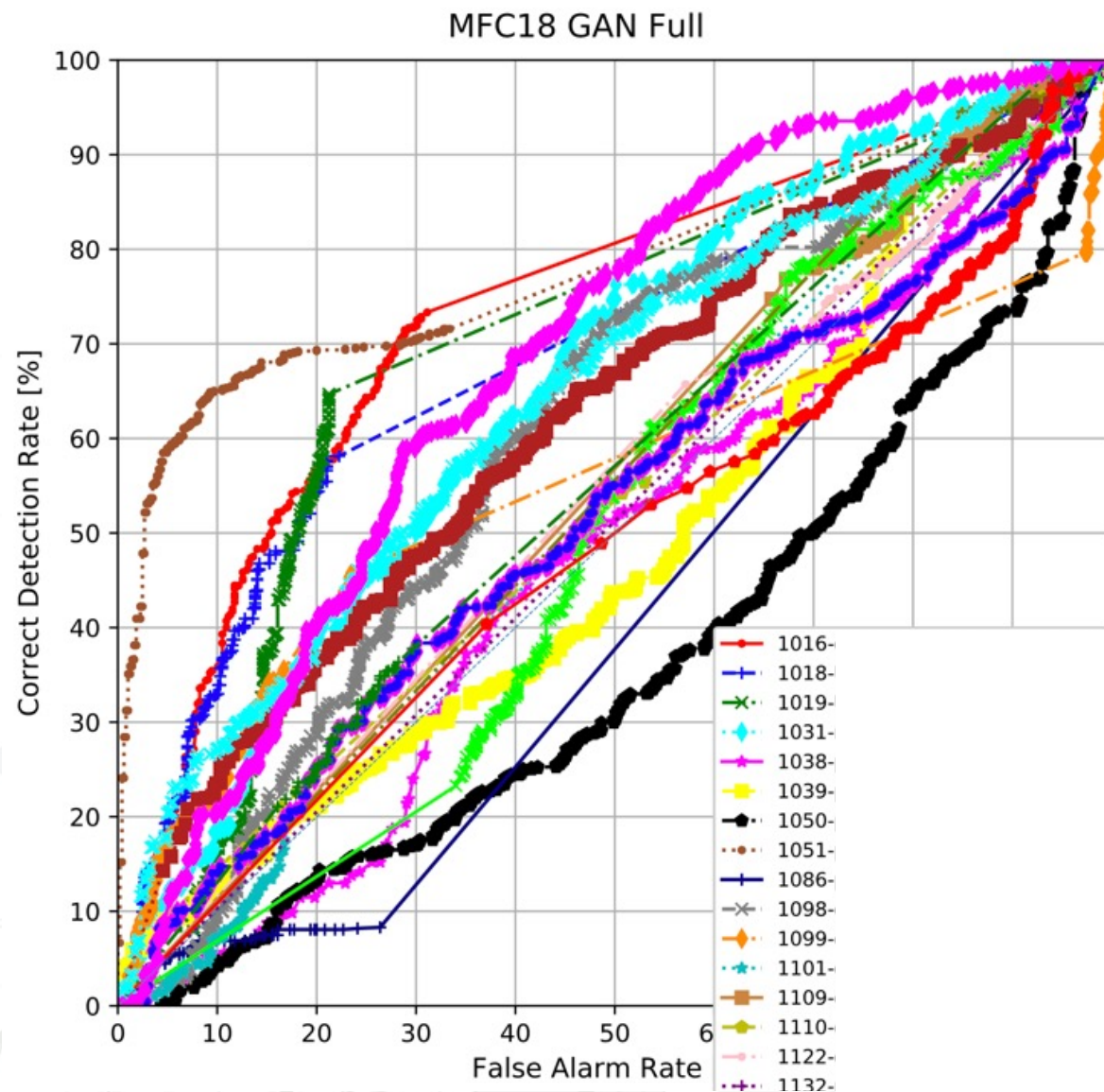
- The journal was created by PAR Government Systems Corporation

11 donor
33 manipulated



MediFor MFC18 IGMD-IO ROC

Full Scoring (25)



MediFor MFC18 IGMD-IO AUC

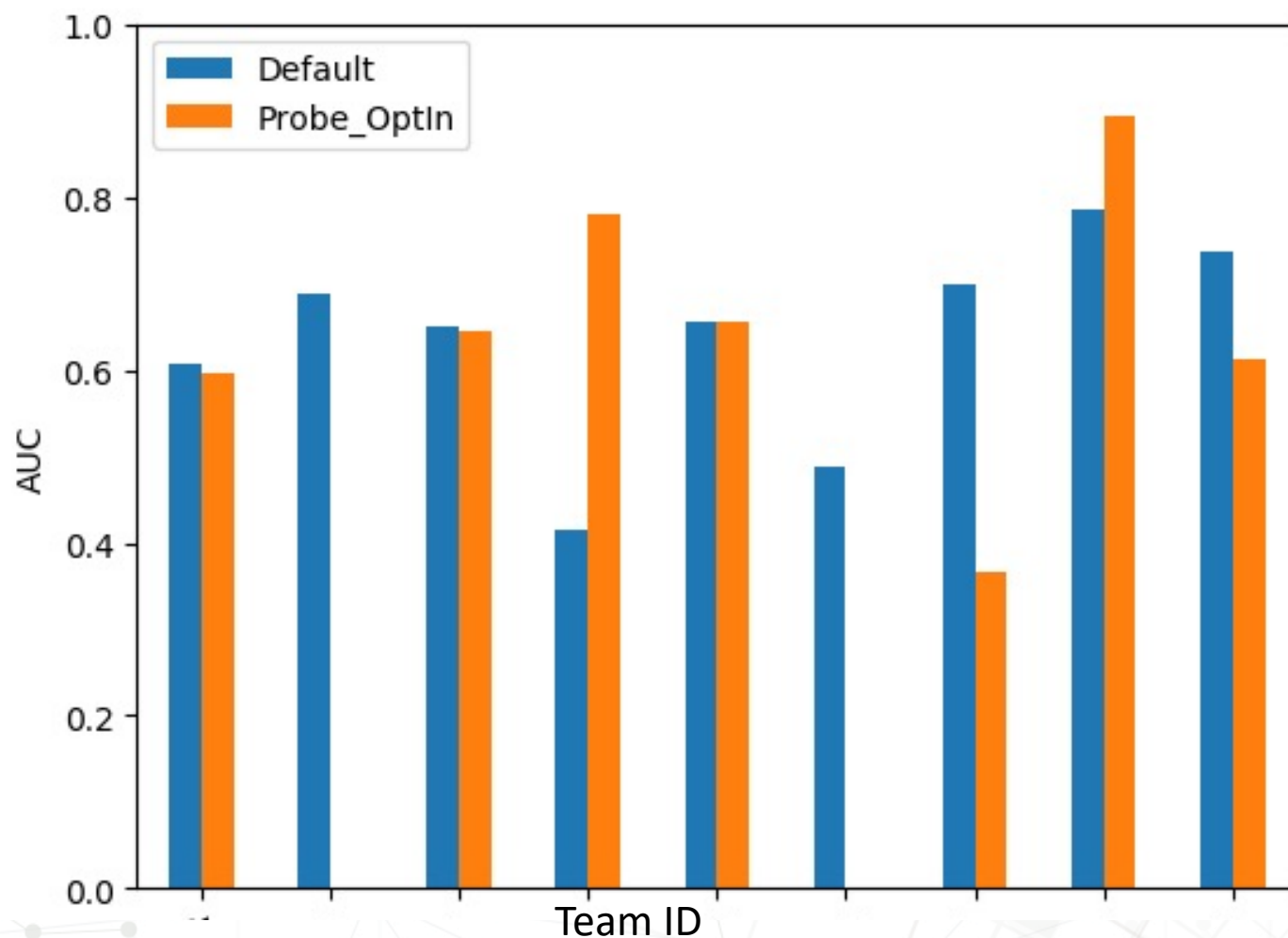


Figure: MediFor GAN Image AUC: Holistic vs. Opt-In

* Distinct bars may represent distinct systems.

MediFor MFC18 IGMD-IO CD@FAR=0.05

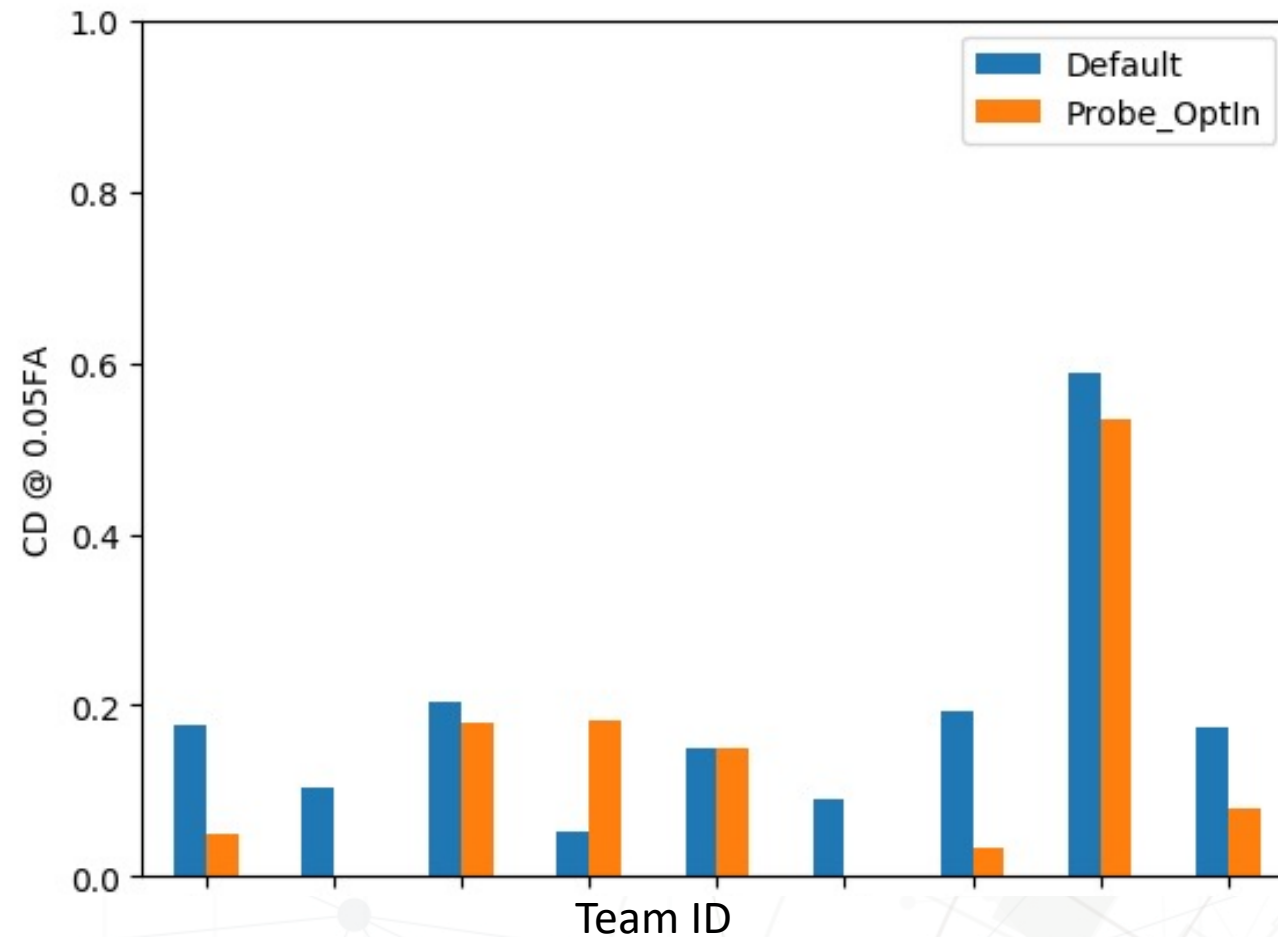


Figure: MediFor GAN Image CD@FAR=0.05: Holistic vs. Opt-In

- The blue bars are the evaluation results of the default systems on the full testing set.
- The orange bars are the evaluation results of the OptIn systems which selected a subset of testing samples for the evaluation only.

OpenMFC Submission Files: IGMD

- Task: IGMD
- Submission file name : e.g., OpenMFC20_ImageGAN_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
 -

OpenMFC Leaderboard: IGMD-IO

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

RANK	SUBMISSION ID	SUBMISSION DATE	TEAM NAME	SYSTEM NAME	AUC	CDR@0.05FAR
1	90	2021-07-10 09:56:14	UIIA	test	0.689716	0.207018
2	93	2021-07-30 18:13:11	UIIA	test	0.683956	0.187135
3	75	2021-06-14 04:12:30	UBMDFL_IGMD	dry-run	0.554261	0.012865
4	52	2021-06-01 15:41:20	UBMDFL_IGMD	dry-run	0.547125	0.009357
5	82	2021-06-23 09:21:26	UIIA	ptchatt	0.500033	0.051077

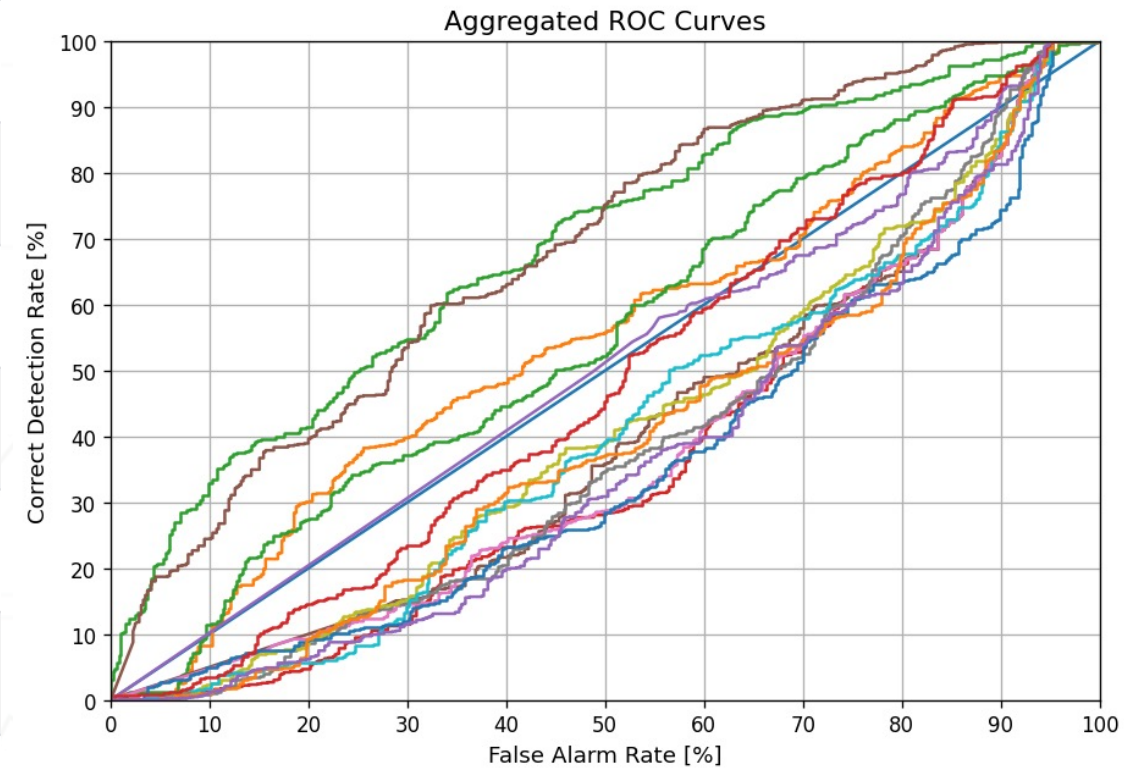


Figure : System performance on the MFC18 Image GAN dataset

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

Task:

Video GAN Manipulation Detection (VGMD)

- VGMD: To detect GAN/Deepfakes manipulated videos specifically while not detecting other forms of manipulations
- Conditions:
 - Video Only (VO)
- Outline:
 - VGMD dataset
 - Participant submission file
 - MediFor MFC results
 - OpenMFC results

VGMD Datasets

- Manipulations in the MFC18 video GAN evaluation dataset:
 - Deepfakes
 - Frame drop
 - GAN Inpainting
- Reference definition
 - Manipulated by GAN operation (with/without other operations) – IsTarget = ‘Y’
 - Manipulated by other operations (not GAN) – IsTarget = ‘Y’
 - NotManipulated – IsTarget = ‘N’

NIST Data Sets	Major operations	Journals	Probe	Date
MFC18 Eval GAN Video (total 118 probes)	PAR Deepfake	-	18	06/30/2018
	Dever Deepfake	-	40	
	Michigan GAN dropframe	-	20	
	PAR GAN Inpainting	-	20	
	World video	-	20	

GAN Video Example (1)

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



Manipulated

85edfcdfed8e2101a6a78936bc4be733.mp4

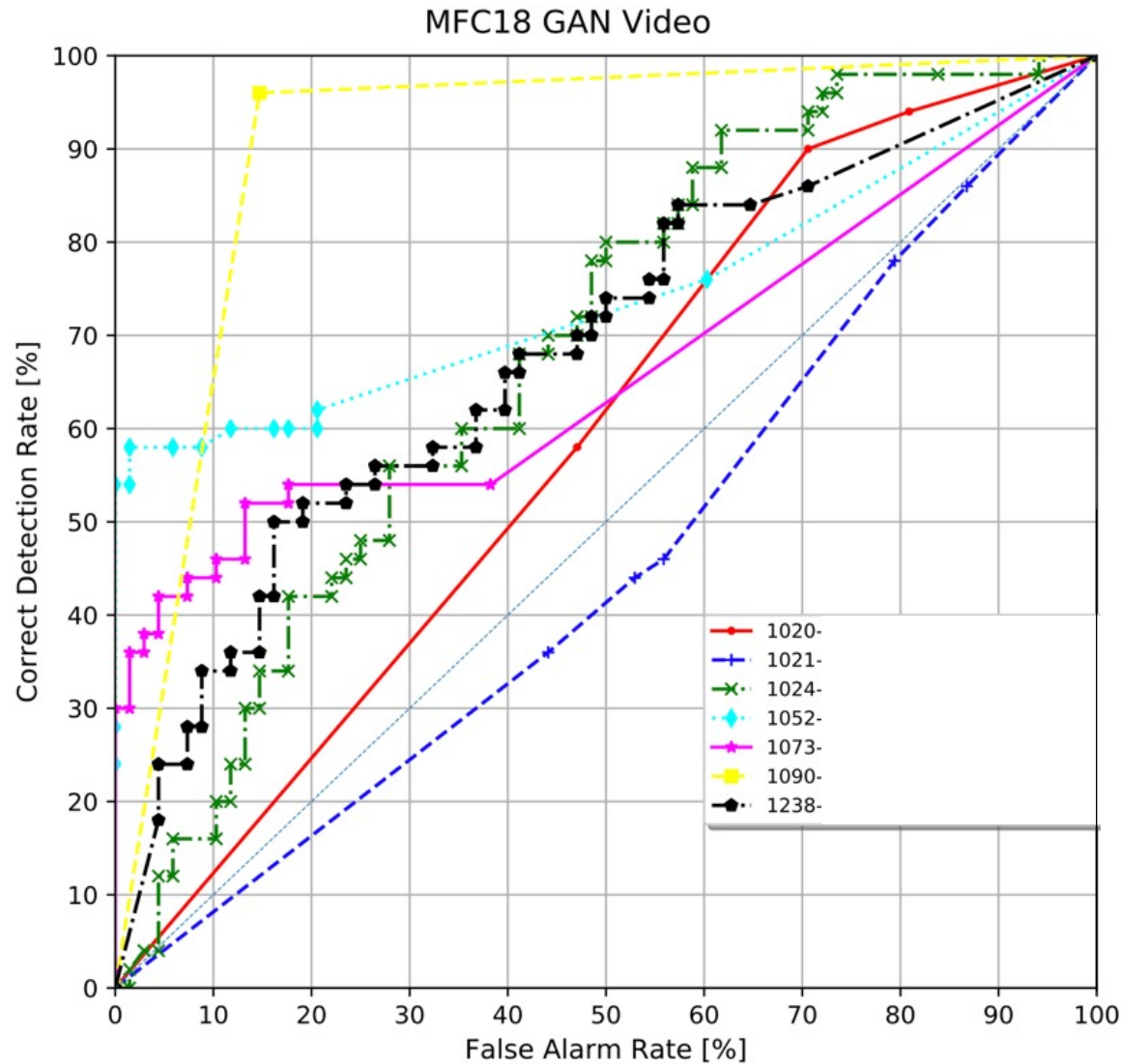
04572fb8d0156500fb82415ce1ed55aa.mp4

OpenMFC Submission Files: VGMD

- Task: VGMD
- Submission file name : e.g., OpenMFC20_VideoGAN_SystemID_TeamID_Date.tar.gz
 - 119 rows
 - Example: the first several rows of the submission file
 - ProbeFileID | ConfidenceScore | ProbeStatus
 - 00b85e9c14b8c6d02e57255fc71c84c0 | 0.752036187 | Processed
 - 01a61a91f9cdf22c62fa29c299d30778 | 0.287194415 | Processed
 - 0508c1016780c9aeb4b4aac1b1bac5e9 | 0.328568837 | Processed
 - 0744b28204e95f9a36dadeceb36a576c2 | 0.433211283 | Processed
 - 08a6f28deff8d5fa885075b852e9f08f | 0.23771142 | Processed
 -

MediFor MFC18 VGMD-IO ROC

Full Scoring (7)



MediFor MFC18 VGMD-IO AUC

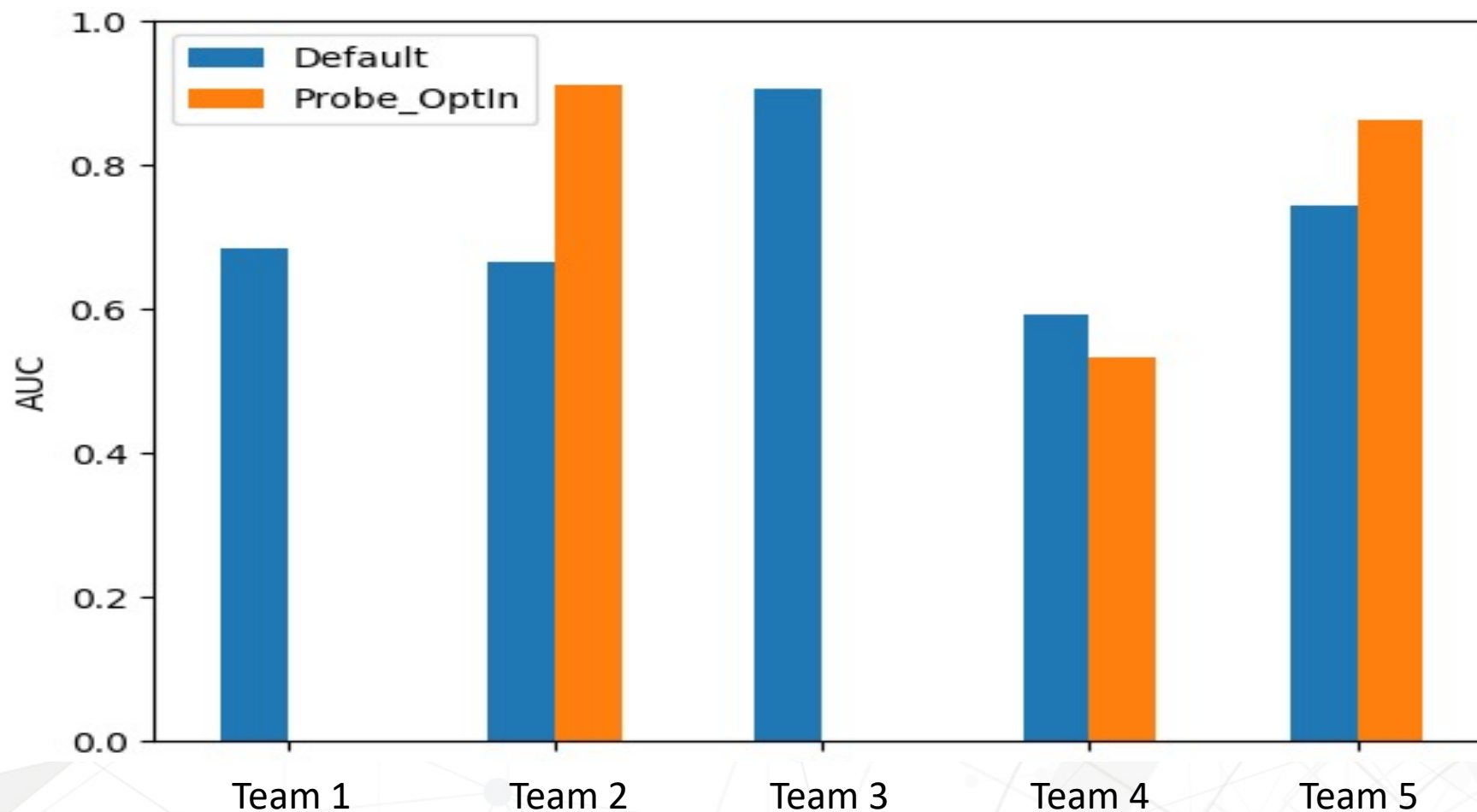


Figure: MediFor Video GAN AUC

5 teams: Team 1 and 3 only have default systems (blue bars).

MediFor MFC18 VGMD-IO CD@FAR=0.05

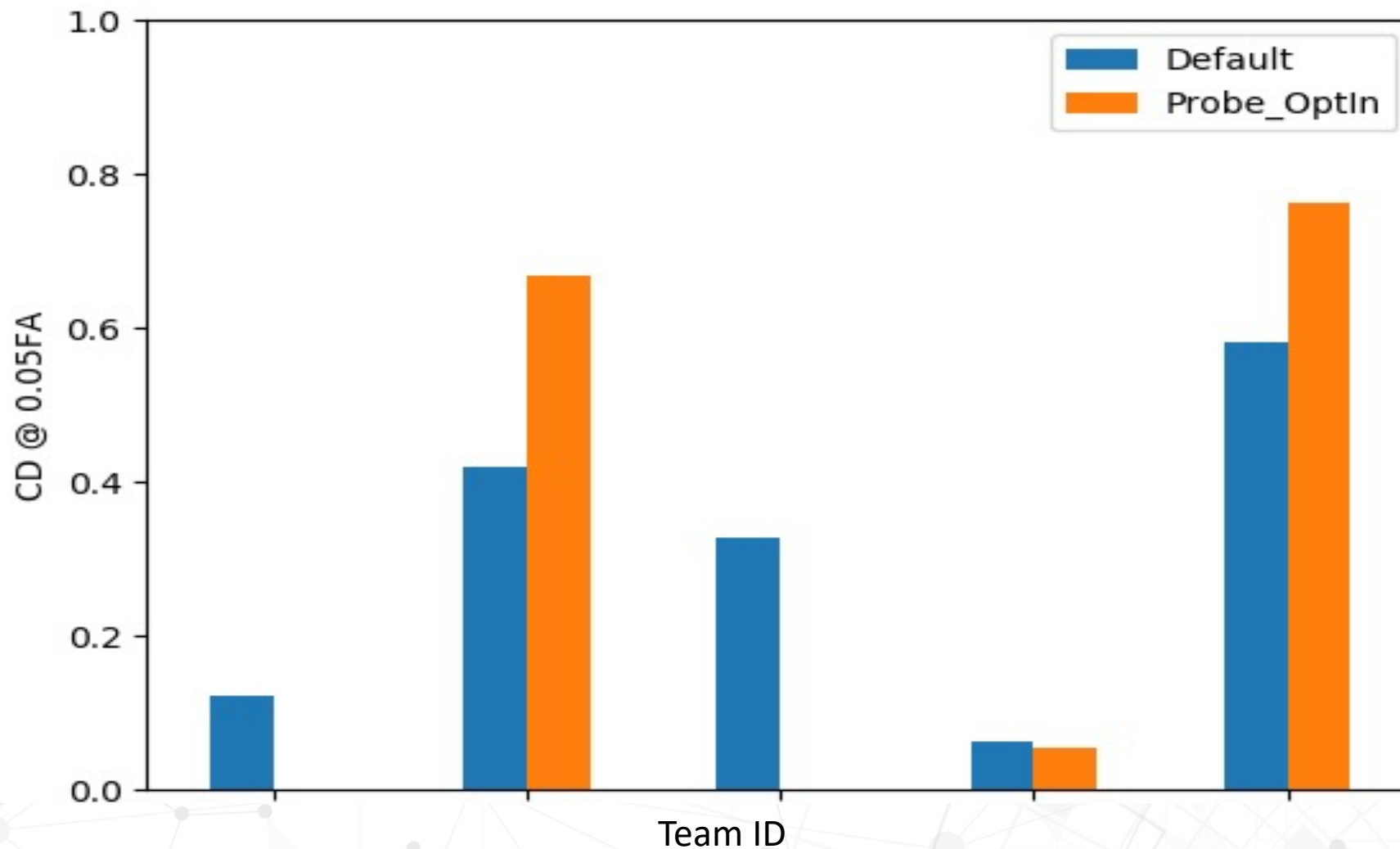


Figure: MediFor Video GAN CD@FAR=0.05

Takeaway

- MediFor MFC19/MFC18 results, cross-year comparisons
- OpenMFC results
- Media Forensics is still in the early stages of development
- **Join the OpenMFC program! <https://mfc.nist.gov>**

Questions?

OpenMFC team: mfc_poc@nist.gov

The background of the slide is a light gray with a complex, abstract pattern of thin lines and dots, resembling a network or circuit board. There are also larger, faint geometric shapes like rectangles and circles scattered throughout.

Thank You!

OpenMFC Evaluation Infrastructure

Lukas Diduch*

Haiying Guan⁺ and Yooyoung Lee⁺

*(CTR) Dakota Consulting, Inc. / MIG

⁺ NIST, Multimodal Information Group, Information Access Division

OpenMFC2021 Workshop : Day 2, Wednesday, Dec. 8, 2021

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.
- All images, graphs, and charts are original works created for DARPA MediFor as well as other NIST Programs.

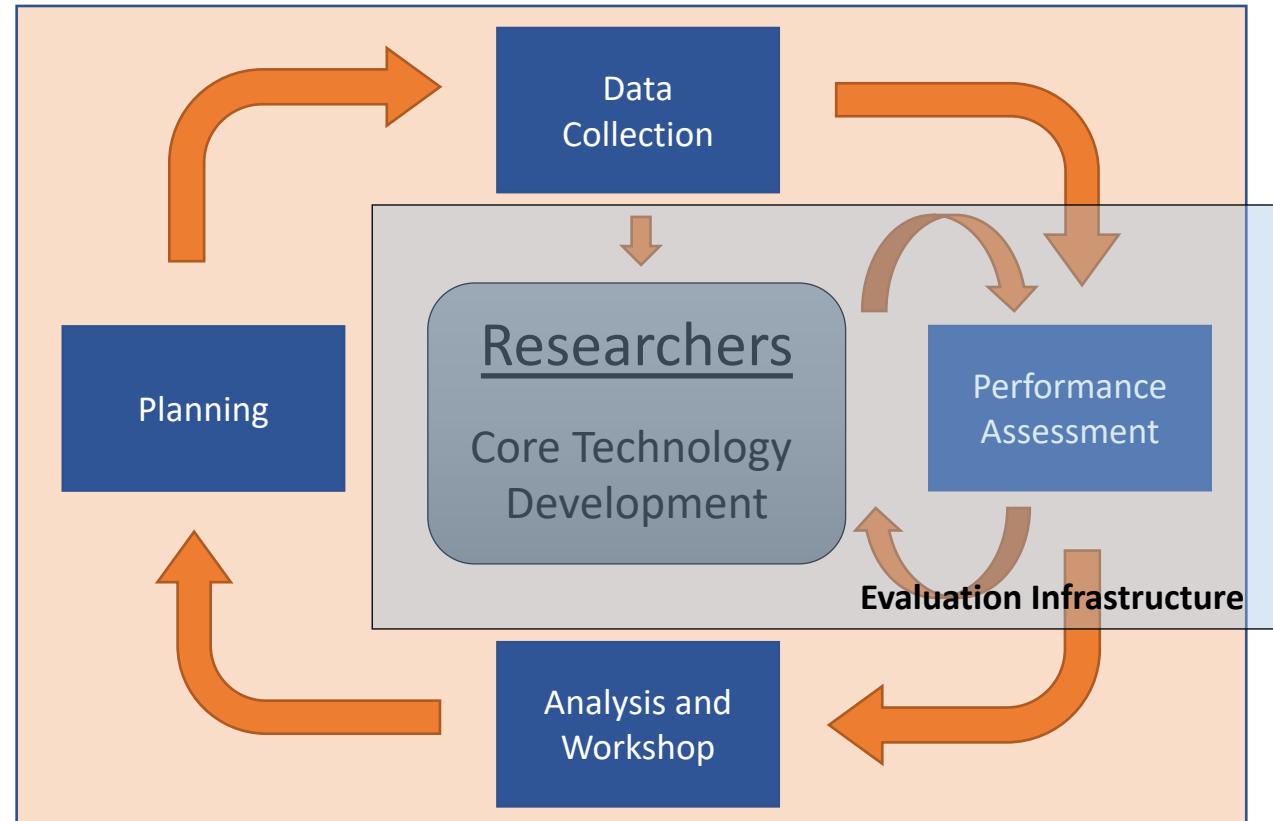
Acknowledgement

- NIST contributors
 - Peter Fontana
 - Timothee Kheyekhah
 - Jesse G. Zhang

Talk Overview

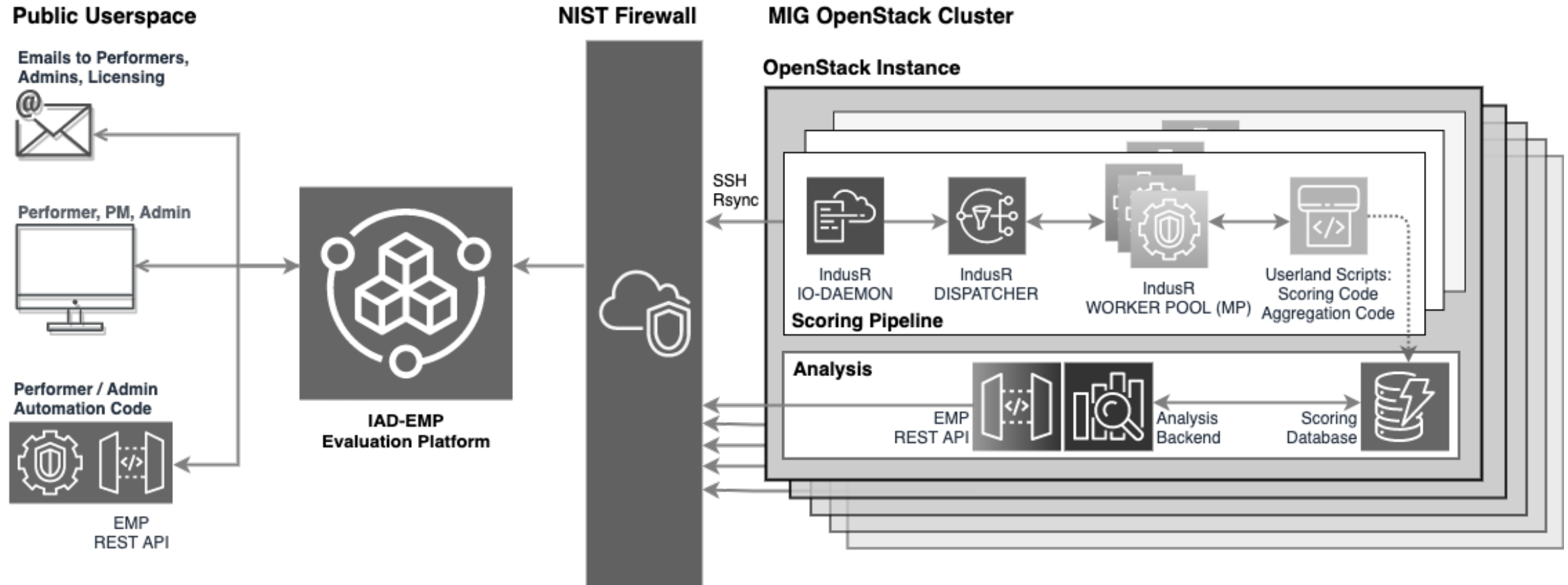
- Evaluation Driven Research Cycle
- OpenMFC Evaluation Infrastructure Components
 - Overview
 - Public-Facing Infrastructure
 - Website
 - Leaderboards
 - Internal Infrastructure
 - Indus Framework
 - Scoring Code

Evaluation-Driven Research



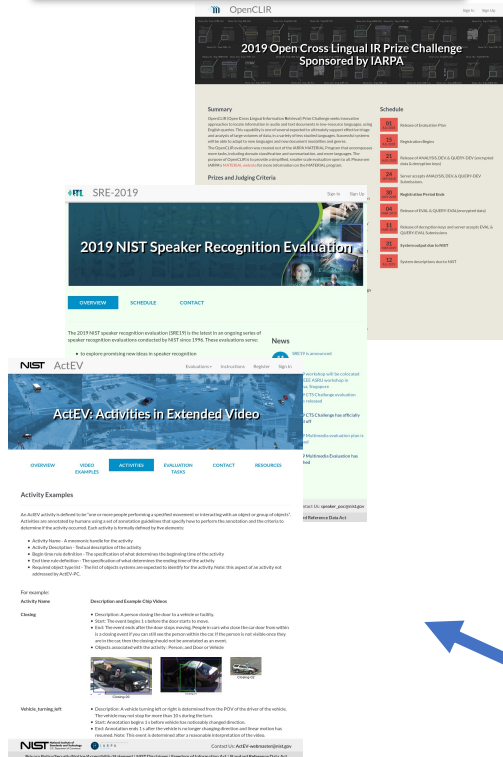
Evaluation Infrastructure Overview

Network Partitioning of the Evaluation system



Public-facing Infrastructure

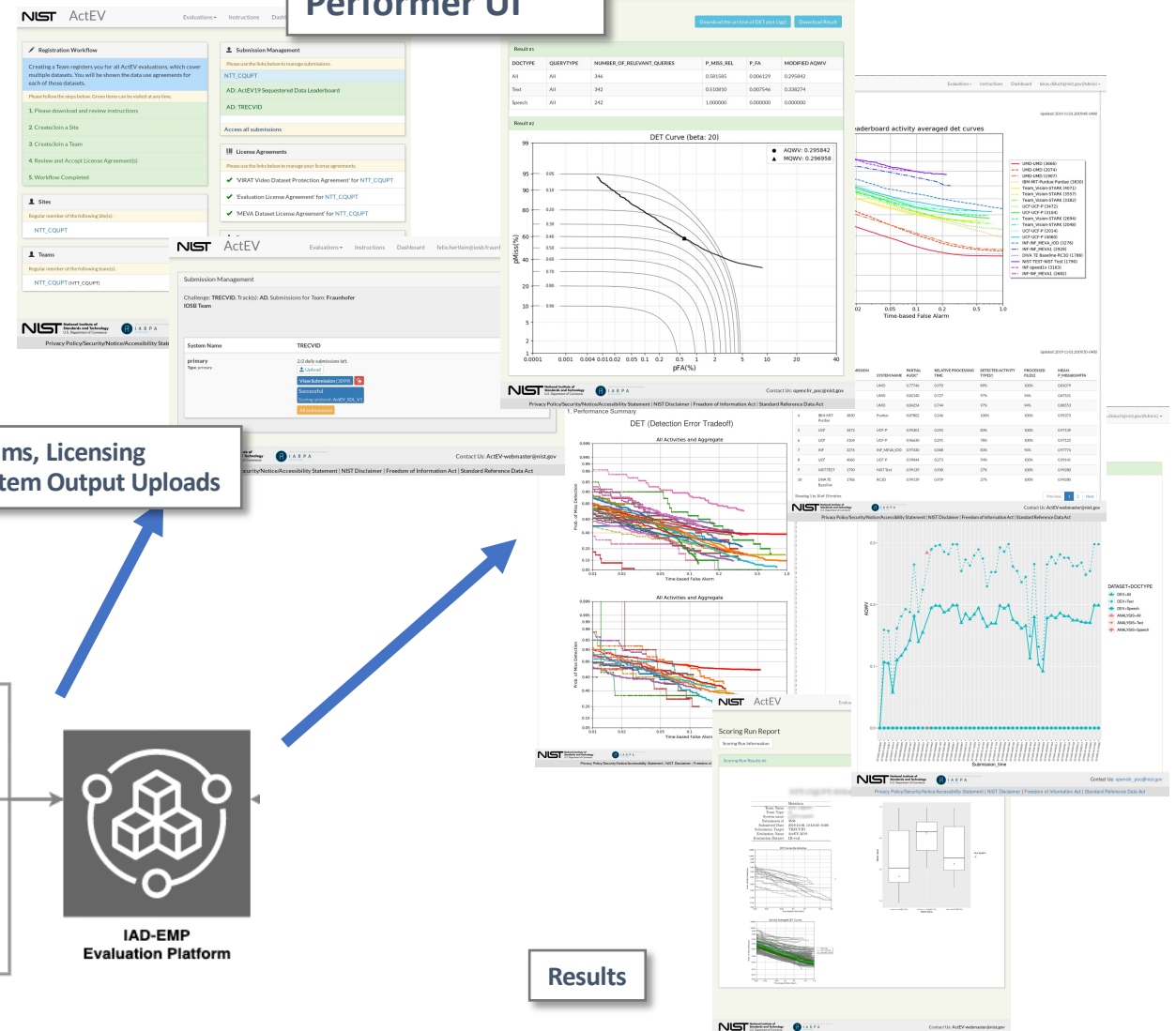
Content Management System (CMS)



Evaluation Admin UI



Performer UI



Teams, Licensing System Output Uploads

Public Userspace

Emails to Performers, Admins, Licensing

Performer, PM, Admin

Performer / Admin Automation Code



EMP
REST API



IAD-EMP
Evaluation Platform

Results

12/07-09/21

NIST OpenMFC 2020-2021

117

Content and Access Management

The collage features three NIST websites. The top left is the ActEV website, titled 'ActEV: Activities in Extended Video', showing a blue-tinted aerial view of a city. The top right is the OpenCLIR website, titled '2019 Open Cross Lingual IR Prize Challenge Sponsored by IARPA', displaying a grid of small images. The bottom left is the OpenMFC20 website, titled 'Open Media Forensics Challenge', showing a beach scene with people and umbrellas. A blue arrow points from the OpenMFC20 website towards the OpenMFC21 login form on the right.

The OpenMFC21 form is divided into two main sections. The top section is for login, with fields for 'Email' (containing 'new_user@mailor.org') and 'Password' (represented by dots). Below the password field is a 'Password confirmation' field (also with dots). A note specifies: '(12 characters minimum with at least one special character, capital letter and one digit)'. A teal 'Sign up' button is located below these fields. To the right of the login section is an icon of an envelope with a lock. The bottom section is for registration, with a large blue 'Log in' button. Below this button are links for 'Register new account', 'Forgot your password?', 'Didn't receive unlock instructions?', and 'Didn't receive confirmation instructions?'.

Performer Dashboard

General Participant Interface: Home Dashboard

1. Evaluation Workflow
2. Sites & Team Managements
3. Evaluation Tasks & Submission management
4. License agreements
5. Datasets Access

The screenshot displays the Performer Dashboard interface, which is organized into several sections. The top left section, titled "Registration Workflow", contains a list of six steps: "1. Create profile", "2. Create/Join a site", "3. Sign and upload license", "4. Create/Join a team", "5. Register to a track", and "6. Workflow completed!". The steps are numbered 1 through 6, with a blue circle containing the number 1 next to "3. Sign and upload license". Below this, the "Sites" section shows "Academia Sinica" with an "Owner" button. The "Teams" section shows "UIIA (Academia Sinica)" with an "Owner" button. The top right section, titled "Submission Management", contains a list of four tasks: "Image Manipulation Detection and Localization (MFC20)", "Video Manipulation Detection (MFC20)", "Video GAN Manipulation Detection (MFC20)", and "Image GAN Manipulation Detection and Localization (MFC20)". The tasks are numbered 1 through 4, with a blue circle containing the number 3 next to "Image Manipulation Detection and Localization (MFC20)". Below this, the "License Agreements" section contains two items: "'Open MFC2020 Evaluation Registration' for Academia Sinica" and "'Open MFC2020 Data Use Agreement' for Academia Sinica". The items are numbered 1 through 4, with a blue circle containing the number 4 next to the first item. The bottom right section, titled "Datasets", contains a list of two datasets: "OpenMFC Datasets" and "OpenMFC Datasets". The datasets are numbered 1 through 5, with a blue circle containing the number 5 next to "OpenMFC Datasets". The bottom of the dashboard features the NIST logo and the text "National Institute of Standards and Technology U.S. Department of Commerce". The contact information "Contact Us: mfc_poc@nist.gov" is located in the bottom right corner.

Registration Workflow

Please follow the steps below. Green items can be visited at any time.

1. Create profile
2. Create/Join a site
3. Sign and upload license
4. Create/Join a team
5. Register to a track
6. Workflow completed!

Sites

Academia Sinica **Owner**

Teams

UIIA (Academia Sinica) **Owner**

Submission Management

Please use the links below to manage submissions.

UIIA

- Image Manipulation Detection and Localization (MFC20)
- Video Manipulation Detection (MFC20)
- Video GAN Manipulation Detection (MFC20)
- Image GAN Manipulation Detection and Localization (MFC20)

License Agreements

Please use the links below to manage your license agreements.

- ✓ 'Open MFC2020 Evaluation Registration' for Academia Sinica
- ✓ 'Open MFC2020 Data Use Agreement' for Academia Sinica

Datasets

Please use the links below to view information about available datasets or to view download options.

[OpenMFC Datasets](#)

NIST National Institute of Standards and Technology
U.S. Department of Commerce

Contact Us: mfc_poc@nist.gov

Phases, Systems, Submission Management

- **Tracks** represent an evaluation task.
- **Phases** represent stages of a Track across a time-period.
- **Systems** represent different system instances/ implementations (e.g. training sets or system parameters etc.). Systems and Phases form a Matrix.
- **Submissions:** System Output to be scored against sequestered test dataset.

The screenshot shows the NIST ActEV Submission Management interface. At the top, there's a header with the NIST logo, 'ActEV', and navigation links for 'Evaluations', 'Instructions', 'Dashboard', and a user profile dropdown labeled 'Participant'. Below the header, the page title is 'Submission Management'. The main content area shows details for a challenge: 'Challenge: TRECVID, Track(s): AD, Submissions for Team: NIST-TEST'. There's a button 'Add new system' and a status '2 systems left.' Below this is a table with columns for 'System Name', 'TRECVID', 'TRECVID20', and 'TRECVID21'. The table lists two test systems, 'Test System I' and 'Test System II', both of type 'primary'. For each system, the table shows the status for each track (all are 'Phase closed, Submissions Disabled.') and provides links to view submissions, along with specific submission details and scoring protocols.

System Name	TRECVID	TRECVID20	TRECVID21
Test System I Type: primary	Phase closed, Submissions Disabled. View Submission (3806) FAIL-scoring Scoring-protocol: ActEV_SDL_V1 All Submissions	Phase closed, Submissions Disabled. View Submission (22841) Original FN: 42_zipbomb.zip FAIL-uncompress Scoring-protocol: ActEV19_AD All Submissions	Phase closed, Submissions Disabled. View Submission (23863) Original FN: baselineACT_1_AD.tgz FAIL-generate_report Scoring-protocol: ActEV19_AD All Submissions
Test System II Type: primary	Phase closed, Submissions Disabled. View Submission (1741) FAIL-scoring Scoring-protocol: ActEV19_AD FAIL-scoring Scoring-protocol: ActEV18_AD All Submissions	Phase closed, Submissions Disabled. View Submission (19749) Original FN: systemnist.zip FAIL-scoring Scoring-protocol: ActEV_SDL_V2 All Submissions	Phase closed, Submissions Disabled. All Submissions

Submissions

UI Generation : Submission Form example

- Customized evaluation parameters stored as Json-schemas
- Parsed & Validated in Ruby to generate the HTML & JavaScript interface

Form Features:

- Form elements (text input, drop-down list, multiple selections, etc.)
- Conditional forms (mutual exclusions between form inputs)

New Submission

Select submission parameters if applicable. Use the 'Browse' button to select a file from your computer and upload it to the server by clicking on 'Submit'.

System: test
Evaluation: OpenMFC 2020
Track: Image Manipulation Detection and Localization
Phase: MFC20
Allowed file formats: .url

Dataset
OpenMFC20 Image

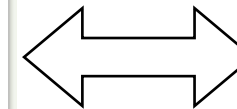
Input condition
Image only

Evaluation protocol
Detection
Localization

Specify Submission URL
Enter URL to fetch ressource from here.

Submit

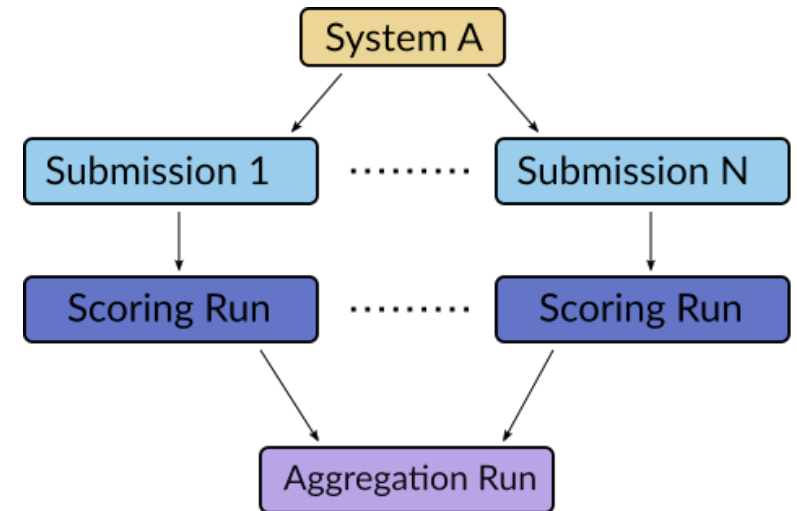
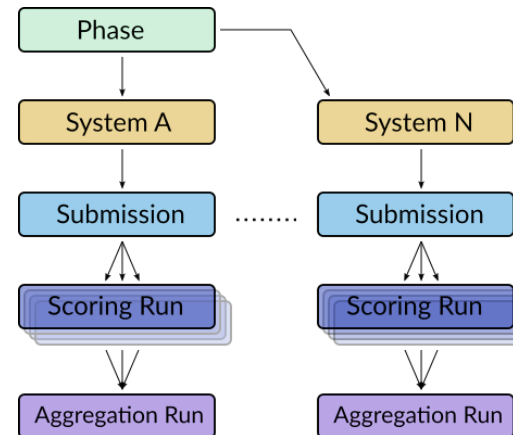
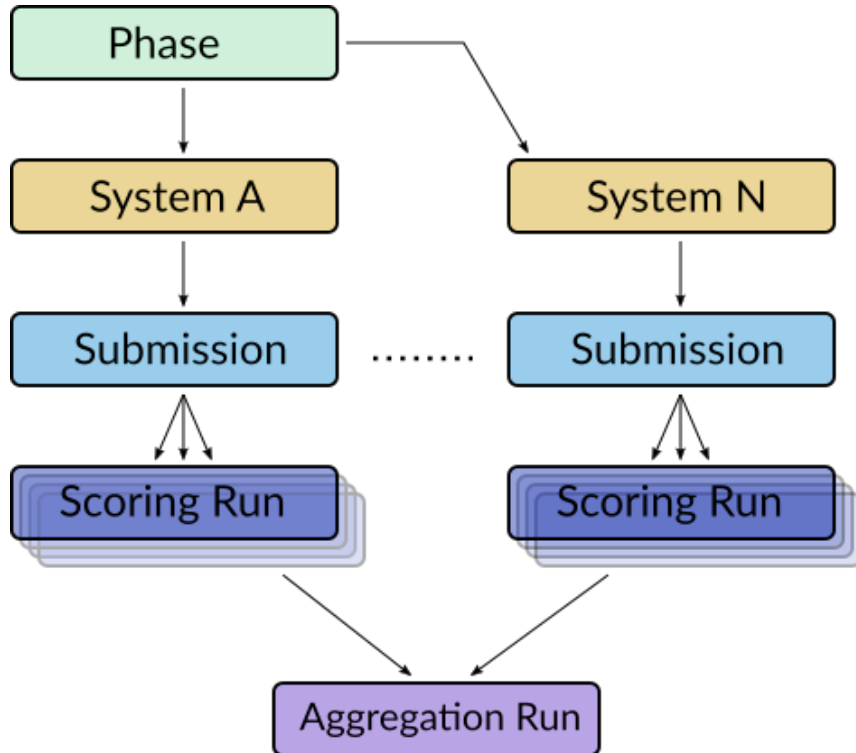
Back



```
{
  "type":"object",
  "required":[
    "Dataset",
    "Evaluation Protocol",
    "Input Condition"
  ],
  "properties":{
    "Dataset":{
      "type":"string",
      "enum":"nil"
    },
    "Evaluation Protocol":{
      "type":"array",
      "items":{
        "enum":[
          "Detection",
          "Localization"
        ]
      }
    },
    "Input Condition":{
      "type":"string",
      "enum":"nil"
    }
  }
}
```

Submissions, Scoring Runs, Aggregation Runs

- System Output is considered a **Submission**.
- Multiple **Scoring Runs** can be run against a submission.
- **Aggregation Runs** can be run against a set of submissions or scoring-runs, or aggregation-runs.



Scoring-Run Report

Scoring Run Report

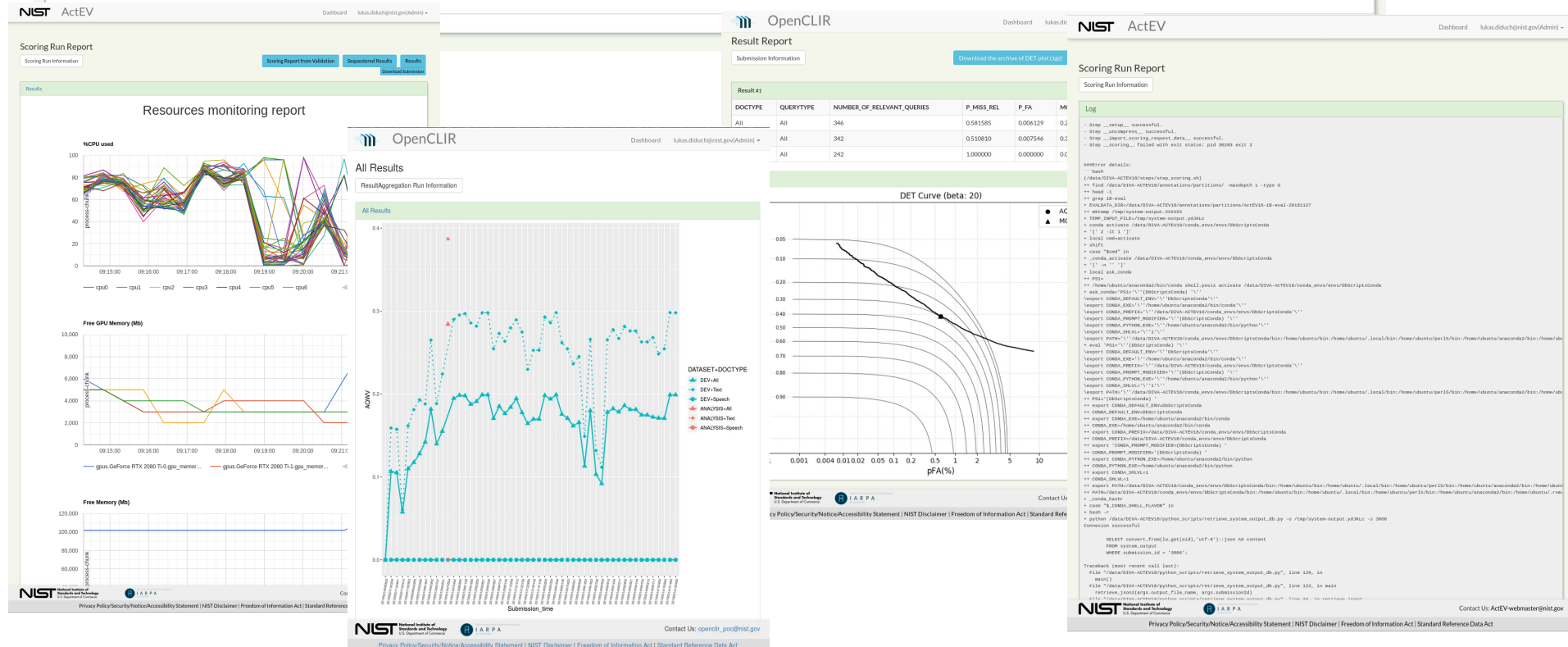
[Show Scoring Run Information](#)

Results

QUERY|TRR|SYS_RESPONSE|AUC|EER|FAR_STOP|AUC@FAR|CDR@FAR|CI_LEVEL|AUC_CI_LOWER|AUC_CI_UPPER|AUC_CI_LOWER@FAR|AUC_CI_UPPER@FAR|CDR_CI_LOWER@FAR|CDR_CI_UPPER@FAR

TaskID == ['manipulation'] | 1.0 | all | 0.616186 | 0.396983 | 0.05 | 0 | 0.071351 | 0.9 | 0.609714 | 0.622467 | 0 | 0 | 0.069902 | 0.072862

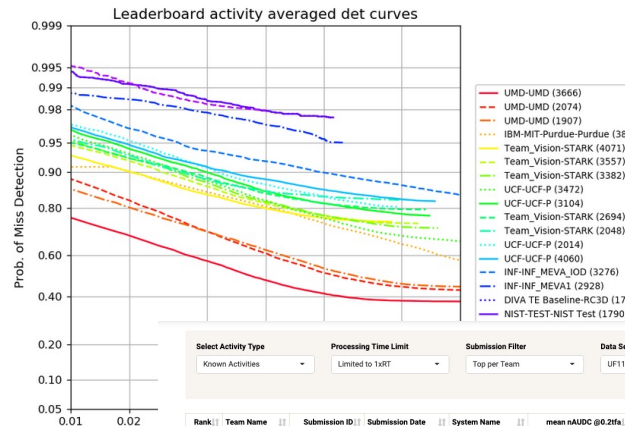
Within a few minutes after teams make a submission they can see scores / error-logs associated for their submission.



Leaderboards

SDL19-scoring-EO

Updated: 2019-11-01 20:09:48 -0400



SDL19-scoring-IR

Show 10 entries

RANK	TEAM NAME	SUBMISSION ID
1	UMD	3666
2	UMD	1907
3	UMD	2074
4	IBM-MIT-Purdue	3830
5	UCF	3472
6	UCF	3104
7	INF	3276
8	UCF	4060
9	NIST-TEST	1790

Select Activity Type: Known Activities Processing Time Limit: Limited to 1xRT Submission Filter: Top per Team Data Set: UFI11

Rank	Team Name	Submission ID	Submission Date	System Name	mean nAUC @0.2fa
1	CMU-DIVA	26095	2021-08-19	UF_EO	0.333
2	UCF	25908	2021-08-07	UCF-P	0.351
3	IBM-Purdue	26113	2021-08-21	IBM_EO	0.353
4	Vision Labs	26404	2021-09-22	vipy	0.376
5	UMD	26619	2021-10-17	UMD	0.389
6	UMD-Columbia	25031	2021-06-06	Columbia_System	0.402
7	UMCMU	25576	2021-07-12	UMCMU-T	0.492
8	Purdue	25782	2021-07-29	Purdue	0.494
9	IMNDS-IRU	24666	2021-05-15	Phu_atropgm	0.635

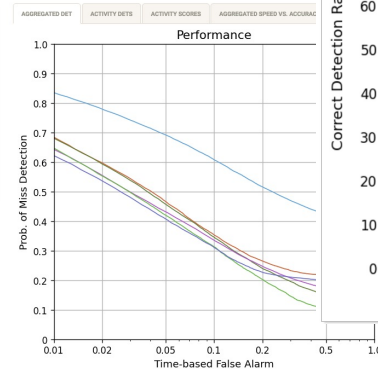
Scores Exploration

Submission Label:

Threshold Measure: nAUC@0.2fa

Graphics Settings: Fine

Download



IMDL-IO IMDL-IM IGMDL-IO IGMDL-IM VMD-VO VMD-VM VGMD-VO VGMD-VM

Image Detection & Localization (Image Only)

Updated: 2020-11-06 21:59:20 -0500

Show 10 entries

RANK	SUBMISSION ID	SUBMISSION DATE	TEAM NAME	SYSTEM NAME	AUC	CDR@0.05FAR	ROC CURVE	AVERAGE OPTIMAL MCC
1	10	2020-11-05 21:53:02.361371-05:00	UIIA	naive-efficient	0.616186	0.071351		
2	1	2020-08-25 16:01:19.003691-04:00	test_team	sys_test	0.494036	0.049217		

Showing 1 to 2 of 2 entries

Previous 1 Next

Download CSV

Updated:

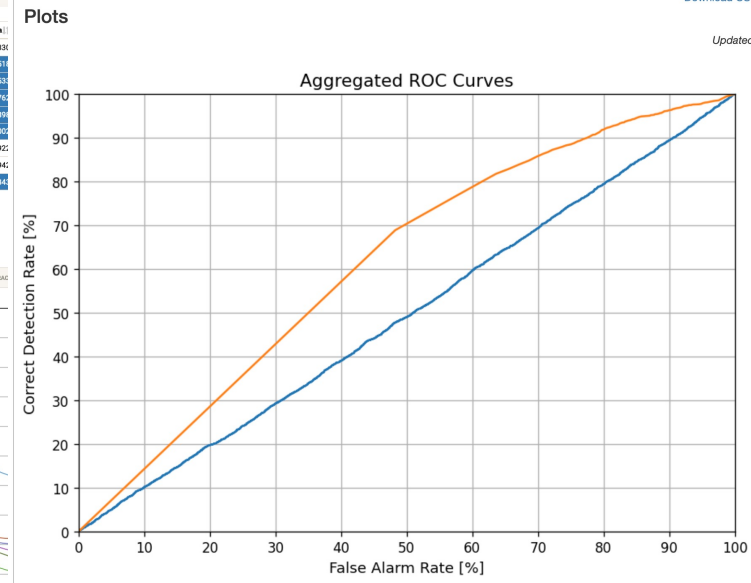


Image Manipulation Detection and Localization:

Image Only (IMDL-IO)

Video Manipulation Detection:

Video Only (VMD-VO)

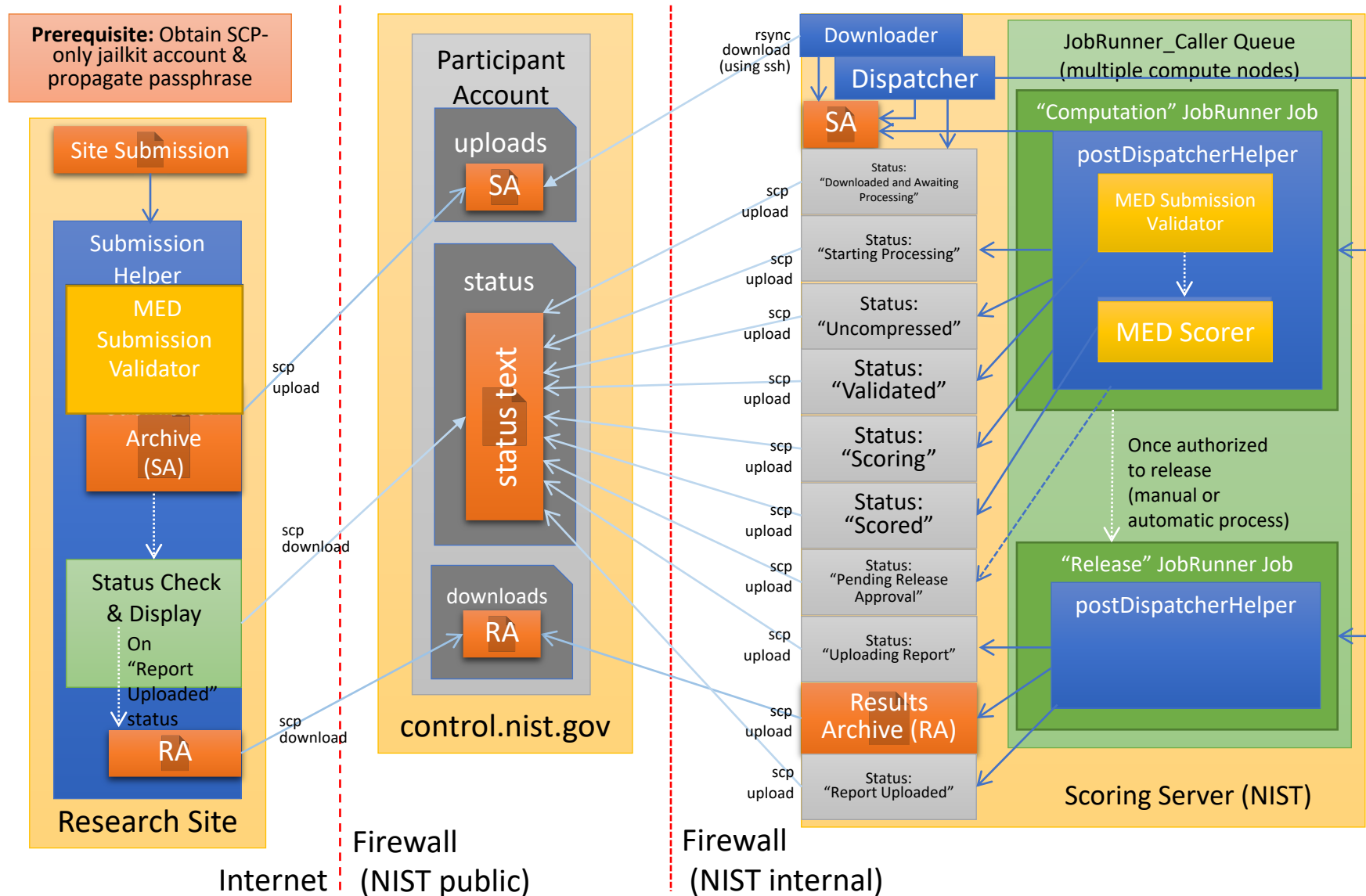
Video + Metadata (VMD-VM)

GAN Manipulation Detection:

Image Only (IGMD-IO)

Video Only (VGMD-VO)

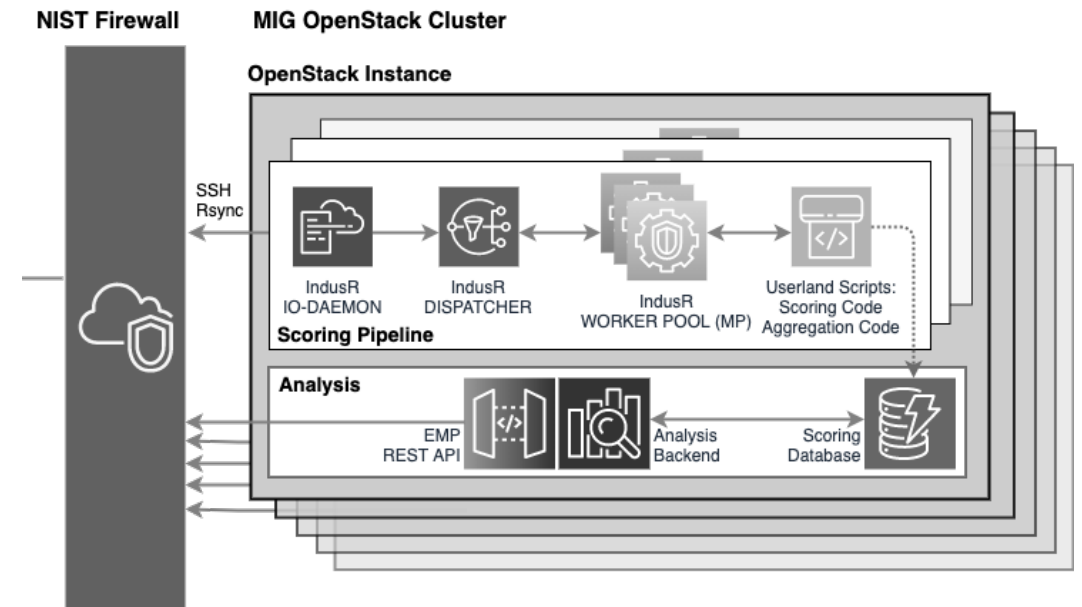
System-Output Submission Automation: Indus



IndusR Backend

Software overview

- IndusR is a Ruby *command line tool* for running pipeline jobs in concurrent and distributed environments.
- Each pipeline is defined by the user's **environment**, **config file** and associated set of **scripts**. The config file is used to fully describe the pipeline **parameters**, which includes server setup, sequence of steps with their respective scripts and associated hooks and hook parameters.



IndusR: Configuration Driven

Pipeline Config File :

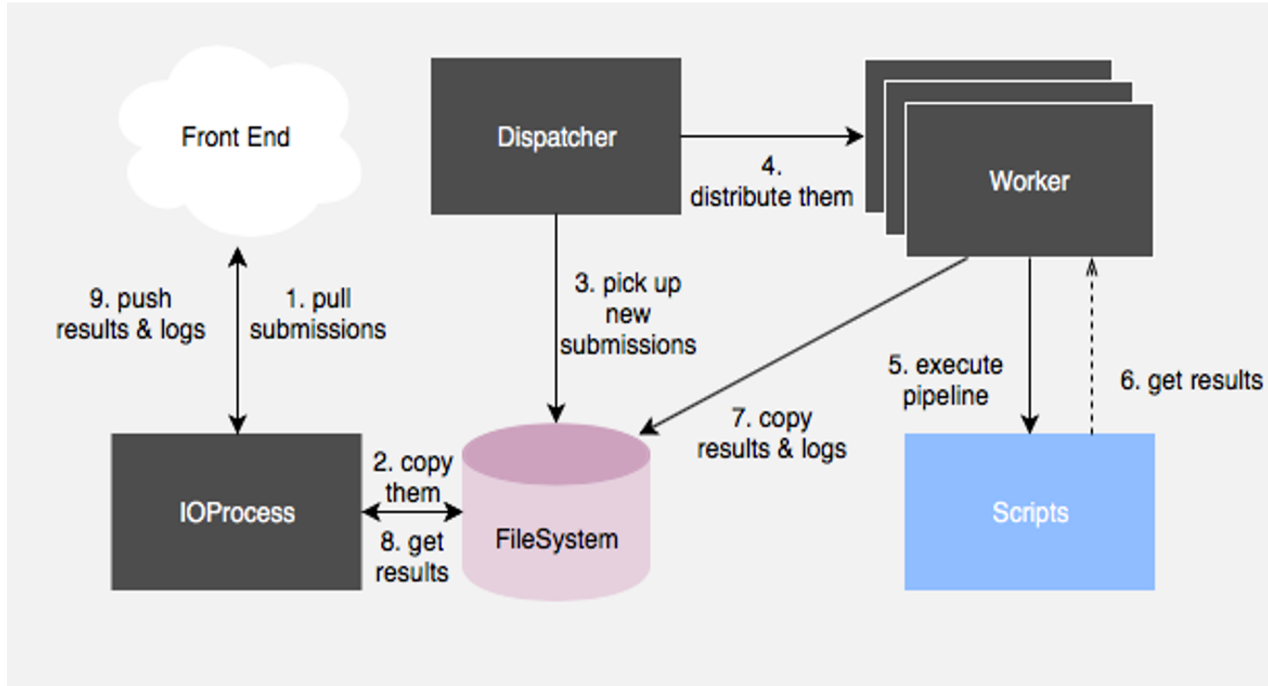
```
evalid: MFCBackend
redis: "redis://localhost:6379"
dispatch_type: scoring_run
eval_root_dir: !ENV INDUS_BE_EVAL_ROOT_DIR
io_interval: 45 # seconds
pipeline:
  setup: step_setup.sh
  download: step_download.sh
  uncompress: step_uncompress.sh
  validate: step_validate.sh
  detscore: step_detscore.sh
  locscore: step_locscore.sh
  updatedb: step_updatedb.sh

hooks: [...]
[paths...]
```

Additional Scoring Pipeline Hooks :

- IO Hooks
 - RSYNC/SSH: push, pull
 - REST API: push, pull, +complex query for pull
- POST step success/fail Hooks
 - bind a script to execute based on step condition.
- Configurable Log-scrubbing
 - Define scripts filtering out sensitive information and excessive detail.

IndusR Components



- **I/O process:** automatically retrieve and upload submissions/scoring- and aggregation-runs/status/results between WebUI and Scoring Cluster either via SSL or REST API.
- **Dispatcher process:** queue incoming runs
- **Worker processes:** Execute single step of a step-sequence expressed as a scoring- or aggregation pipeline running against the performers submission. Workers can be bound to individual steps.
- **Redis:** Key-value store to synchronize application state across all distributed/parallel processes.
- **Scripts:** Userland scripts written by evaluators.

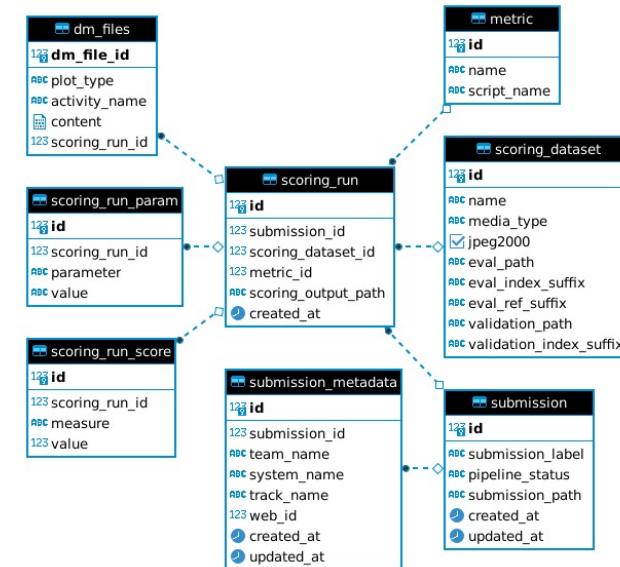
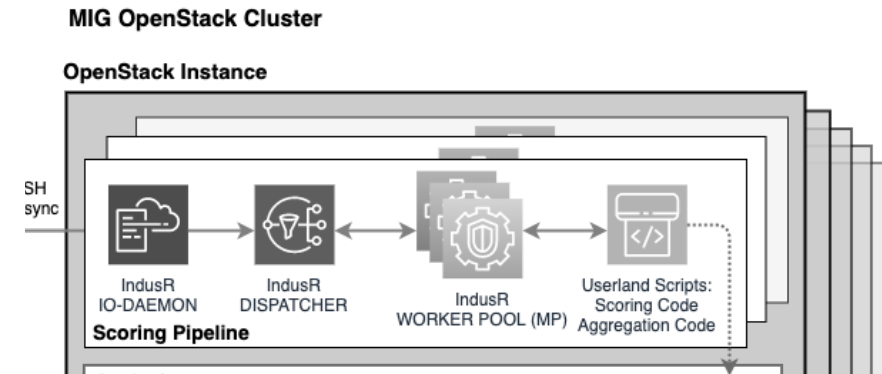
OpenMFC Implementation

Scoring Pipeline Implementation:

- One scoring pipeline handles all eval Tracks
- One aggregation pipeline handles all Leaderboards

Configuration Driven mainly through Database:

- Database stores
 - **Track configuration**
 - Track scores
- Leaderboards can use Database for aggregation or as a source for complex analysis online and offline.
- Advantages:
 - Adding new tasks and datasets quickly to the evaluation logic is straight-forward.
 - Database w/ Scores can be hosted anywhere.
- Disadvantages
 - Advanced setup.



Conclusion

As an **independent Scoring Entity** we are providing

- Fair **comparison of systems performance** across community peers
- **Standardized metrics** against **sequestered** Dataset(s)

Our Evaluation Infrastructure is

- Providing all **essential evaluation resources** online
- Facilitating and **managing scoring process** within eval-constraints
- Providing **tracking of progress** across different modalities
- Providing scoring computation resources (for open evaluations)

Leaderboard based **evaluation cycle** enables **rapid R&D**

Questions?

OpenMFC team: mfc_poc@nist.gov

Thank You!



Feedback and Discussion: OpenMFC Evaluation Leaderboard & System Analysis

Lukas Diduch, Haiying Guan
and Yooyoung Lee


Multimodal Information Group,
Information Access Division

OpenMFC2021 Workshop : Day 2, Wednesday, Dec. 8, 2021



Infrastructure Discussion Outline

- Leaderboard upgrade
 - Anonymous submission to the leaderboard
 - Feedback collection for the leaderboard/submission
- Online system analysis tool
 - Holistic Evaluation vs. Opt-In
 - Selective Scoring

The background of the slide is a light gray with a complex network of thin lines and dots, resembling a molecular or digital structure. There are also larger, faint geometric shapes like rectangles and circles scattered across the background.

Discussion: Holistic vs. Opt-In

Holistic vs. Opt-In Technologies

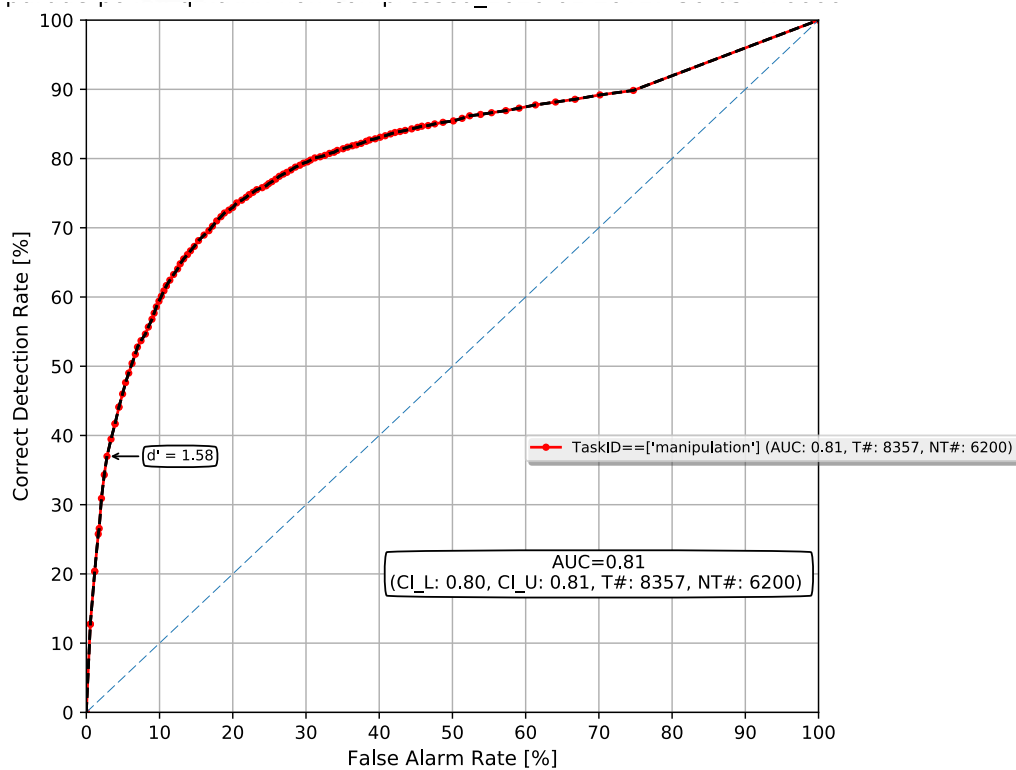
- Evaluation challenge:
 - Some media forensic systems have the option to return no response
 - E.g., face illumination consistency systems should not respond if no face was found in the image

Probe Status	Description
Processed	probe was fully processed
NonProcessed	probe was not processed due to a system failure of some kind occurring
FailedValidation	probe failed the MediScore Validator tool and will be given a score of 0

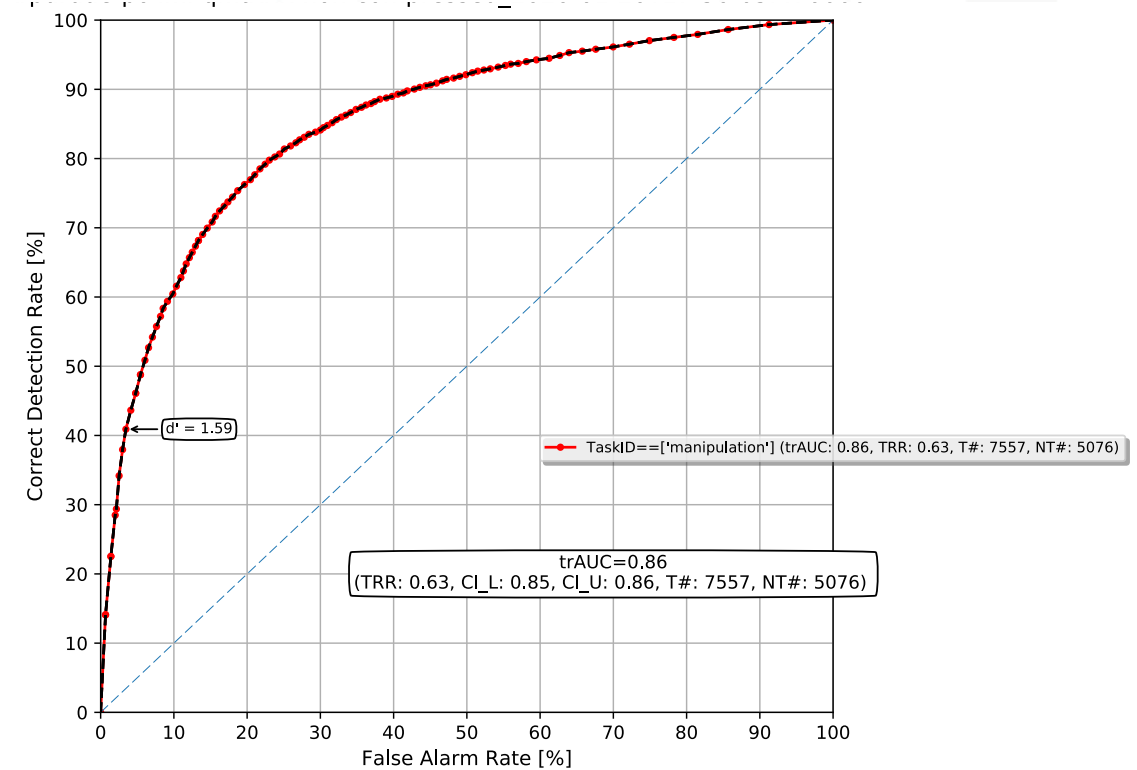
- NIST reports:
 - Holistic performance measures: score all trials
 - Opt-In performance measures:
 - Trial Response Rate (TRR) – Percent of Processed, NonProcessed, and FailedValidation images
 - Performance measures excluding opt'd out probes

Holistic vs. Opt-In Technologies

- Some media forensic systems respond only to certain media types
 - e.g., jpeg compression systems should not respond if input is not in jpeg format



(a) Holistic



(b) Opt-In

MediFor IMD: Opt-In

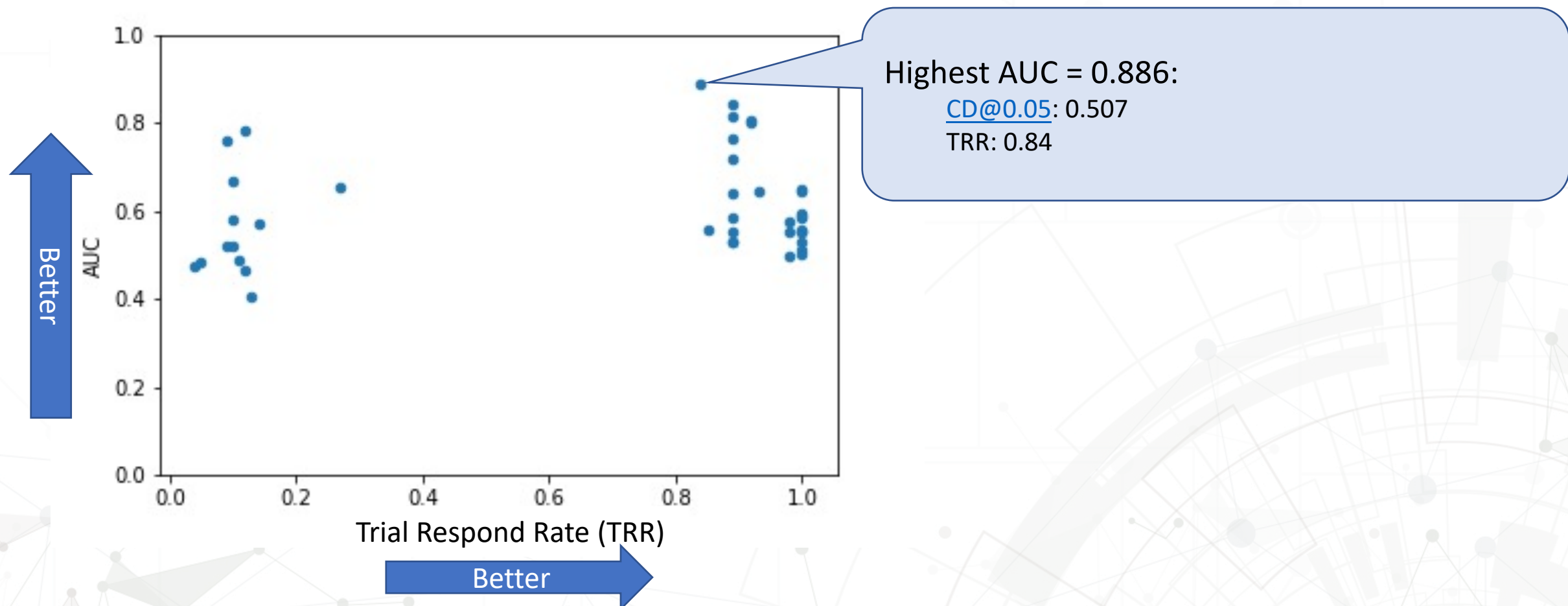


Figure: Image detection Opt-In system Area Under the Curve (AUC) vs. Trial Response Rate (TRR) on MFC19 EP1 Image dataset (each point is an analytic system)

System Performance Analysis: Selective Scoring

- Concept definition
- MediFor MFC19 selective scoring
- MediFor MFC19 results

Selective Scoring - Concept

- Question addressed
 - Some analytic systems only applied to a certain manipulation. Those systems should not be penalized on the test samples that they are not applicable to (e.g., double JPEG detector does not work on PNG images).
- MFC supports selective scoring evaluation infrastructure
 - Allow researchers to solve subdomain problems
 - Allow evaluation program to analyze the system performance based on
 - Different manipulation operations;
 - Different manipulation semantics (queried by the combination of manipulation operations and other metadata)
 - Deep understanding different systems' performance for overall system integration.
 - Full usage of data: data collection and annotation cost is very high. With selective scoring, the single journal provides multiple test cases for different subtask evaluations.

Selective Scoring – Infrastructure

- Performers declare the limits/presumptions of their system
 - A system description
 - Suggest suitable selective scoring condition.
- Structured Manipulation Data
 - Structured metadata collection and annotation: the manipulation Journaling Tool (JT) records manipulation operations with a graph step by step.
- Query-able reference ground-truth data
 - Evaluation reference data generation: the evaluation data generation tool, TestMaker extracts the metadata and put in reference ground-truth.
- Evaluation Scoring Software
 - NIST evaluation scoring software: selective (query-based) evaluation: selected trials (based on the query) will be scored, while unselected trials will NOT be scored.

Image Selective Scoring

- Query definition
- MFC19 image detection selective scoring result summary
- MFC19 image localization selective scoring result summary

Selective Scoring – MFC19 Image Selective Scoring Queries

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 techniques 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 NIST OpenMFC 2020-2021	5750

Selective Scoring for System Performance Analysis

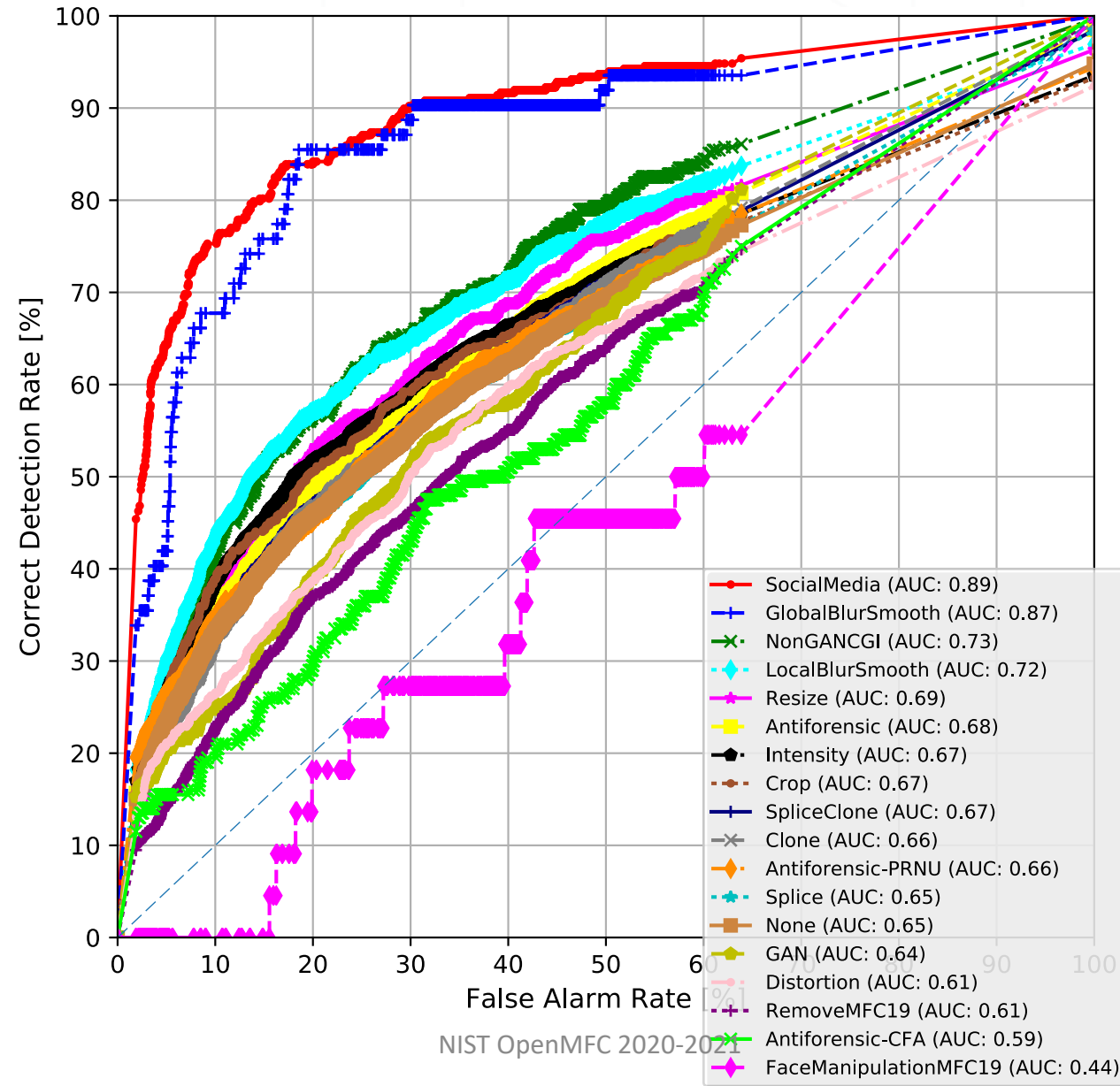


Image Detection Selective Scoring – Top 10 system AUC

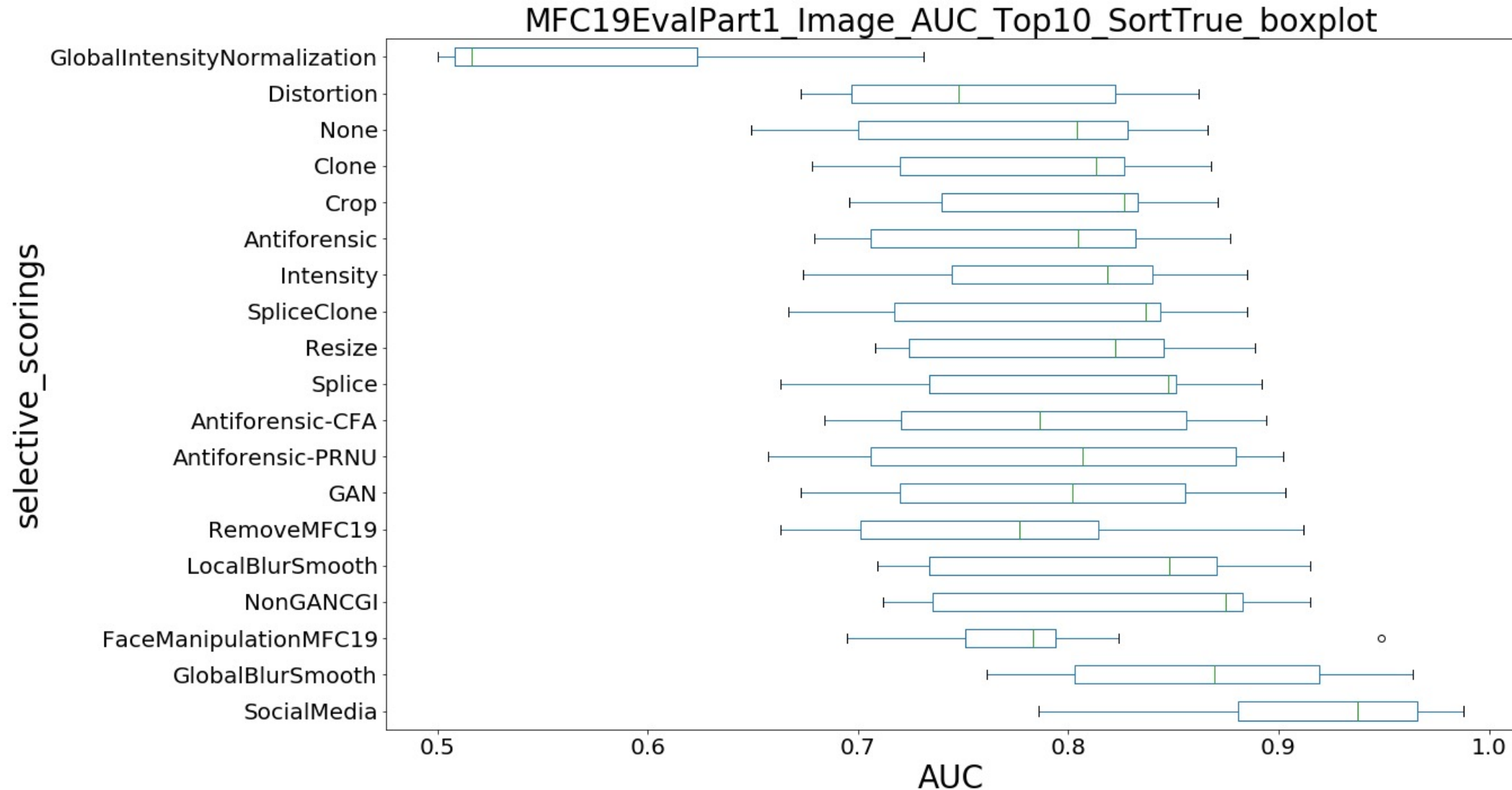
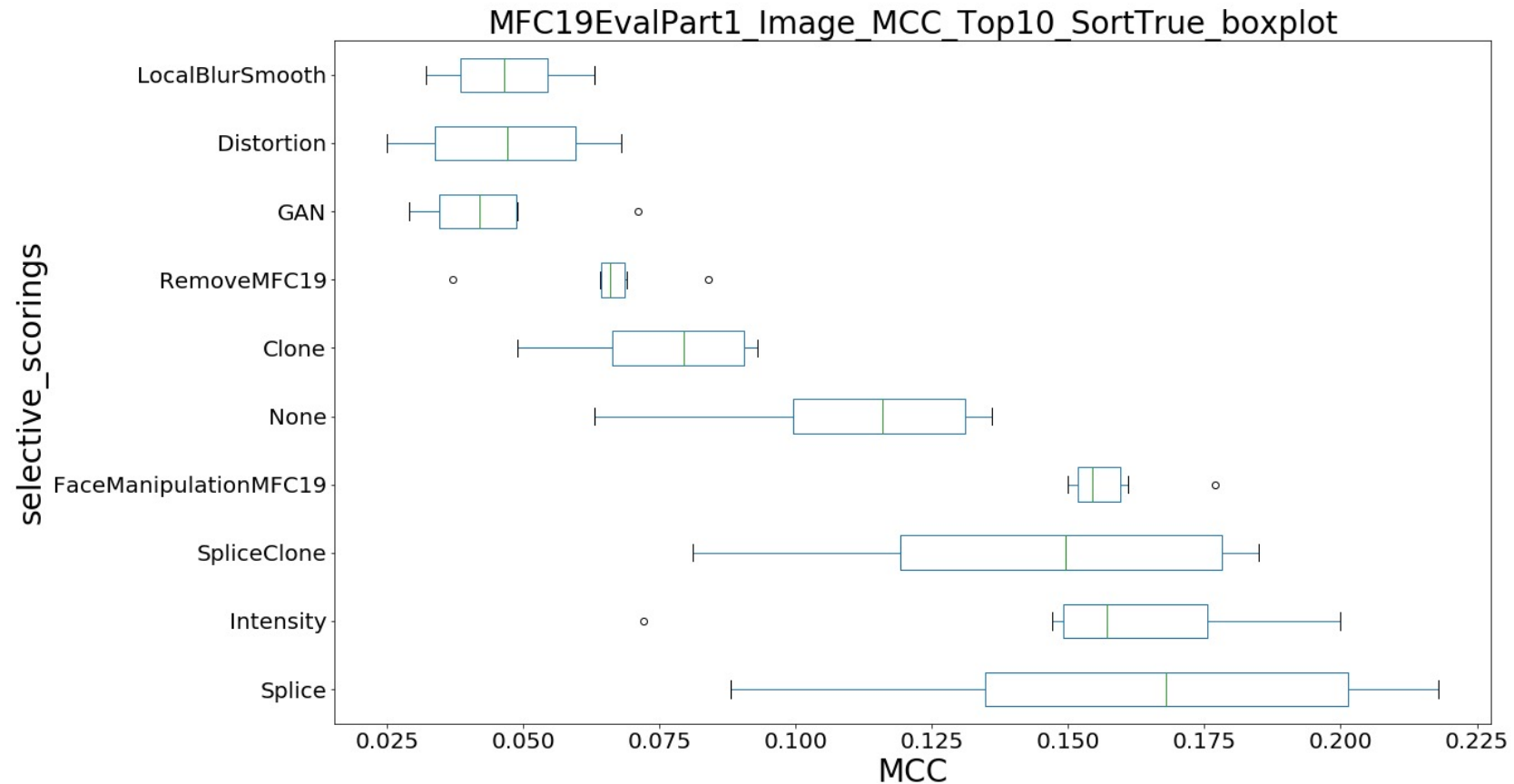


Image Localization Selective Scoring – Top 10 system MCC



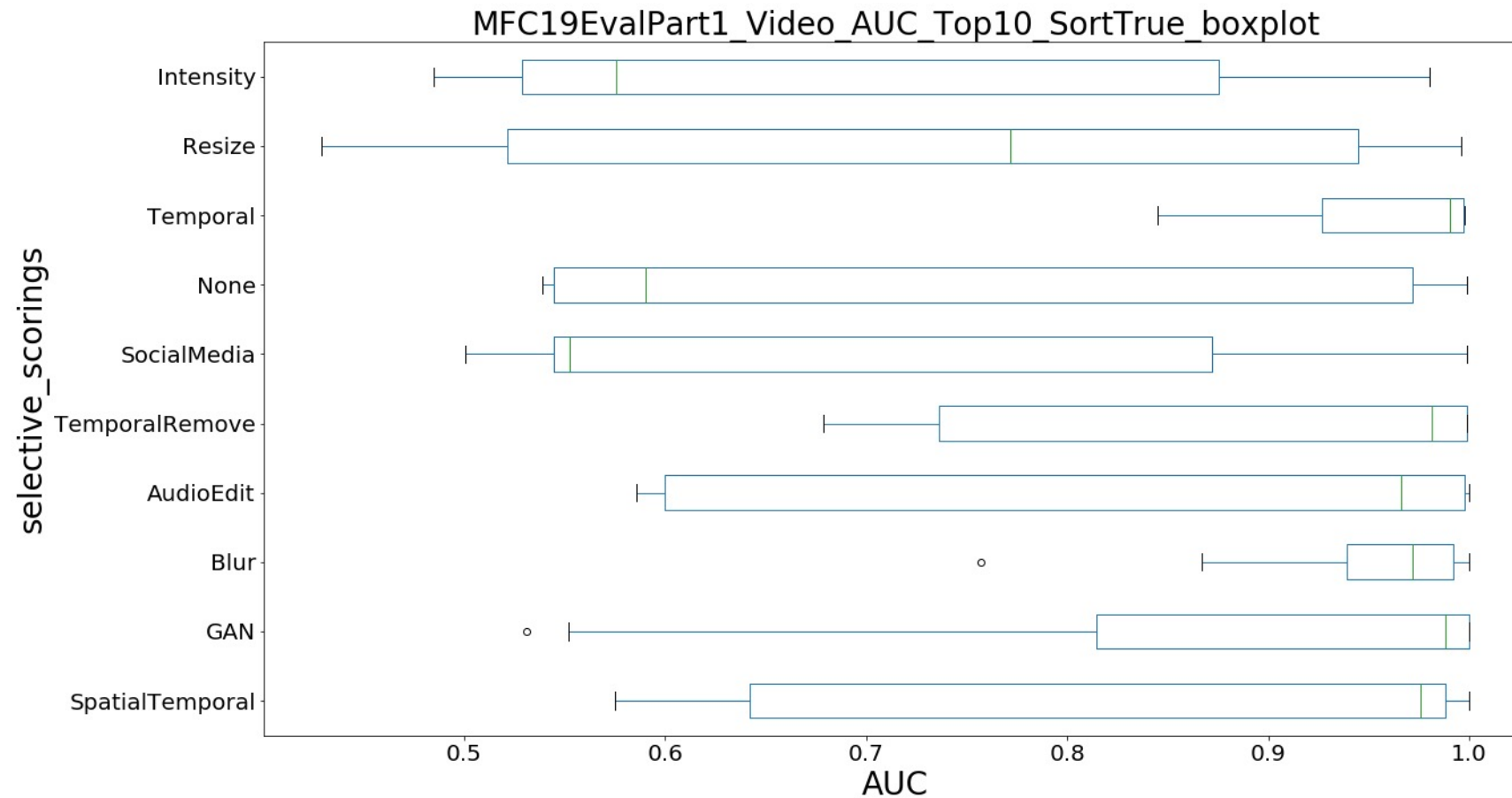
Video Selective Scoring

- Query definition
- MFC19 video detection selective scoring result summary
- MFC19 video localization selective scoring result summary

Selective Scoring – MFC19 Video Selective Scoring Queries

Name	Definition	Count
Temporal	Frames are edited (with Add,Splice,Clone,Move) at different time in video beside Remove	28
TemporalRemove	Frames are removed at different time in video	40
SpatioTemporal	Pixels on a frame are edited (with Splice,Clone,Remove,Overlay, CGI) across frames in video	212
Intensity	A range of intensity pixel values is changed across frames in video	91
SocialMedia	Any social media (e.g., Youtube)	234
AudioEdit	Any operations that edit audio of video	195
Resize	Video dimensions are changed	9
Blur	Any techniques that use a low-pass filter (locally/globally) to remove outlier pixels (e.g., noise) across frames in video	47
GAN	Any operations that use GAN-based techniques in video	11
All	All without selective scoring	369

Video Detection Selective Scoring – Top 10 system AUC



Questions?

OpenMFC team: mfc_poc@nist.gov

Thank You!

NIST Open Media Forensics Challenge



OpenMFC 2021 -2022 Work Plan

Haiying Guan

Yooyoung Lee, Lukas Diduch, and Ilia Ghorbanian

Multimodal Information Group,
Information Access Division , ITL, NIST

OpenMFC2021 Workshop : Day3, Thursday, Dec. 9, 2021

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.
- All images, graphs, and charts are original works created for DARPA MediFor as well as other NIST Programs.

Acknowledgement

- External collaborators:
 - Prof. Siwei Lyu in University at Buffalo
 - Prof. Jennifer Newman, Prof. Yong Guan, Li Lin from Iowa State University, and Prof. Roy A. Maxion from Carnegie Mellon University
- NIST contributors
 - Jonathan Fiscus
 - Timothee Kheyrkhah
 - Peter Fontana
 - Jesse G. Zhang

Outline

- OpenMFC 2021-2022 Evaluation Plan
 - OpenMFC 2021-2022 evaluation task
 - OpenMFC 2021-2022 datasets
 - Leaderboard upgrade
- Workshop publication
- Collaboration

OpenMFC 2021-2022 Evaluation Tasks and Datasets

- Image Manipulation Detection and Localization (IMDL)
 - Task: to detect if the image has been manipulated, and then to localize the manipulated region
 - Dataset: OpenMFC2020-2021 data (MFC19 EP1) – 16K
- Video Manipulation Detection (VMD)
 - Task: to detect if the video has been manipulated
 - Dataset: OpenMFC2020-2021 data (MFC19 EP1) – 1.5K
- Image GAN/Deepfakes Manipulation Detection (IGMD)
 - Task: to detect GAN-manipulated images
 - Dataset 1: OpenMFC2020-2021 data (MFC18 image GAN challenge) – 1.3K
 - Dataset 2: New data coming: OpenMFC 2020 Image GAN evaluation dataset
- Video GAN/Deepfakes Manipulation Detection (VGMD)
 - Task: to detect GAN/Deepfakes manipulated videos
 - Dataset 1: OpenMFC2020-2021 data (MFC18 video GAN challenge) – 118
 - Dataset 2: New data coming: OpenMFC 2022 video Deepfakes evaluation dataset

OpenMFC 2021-2022 New Tasks

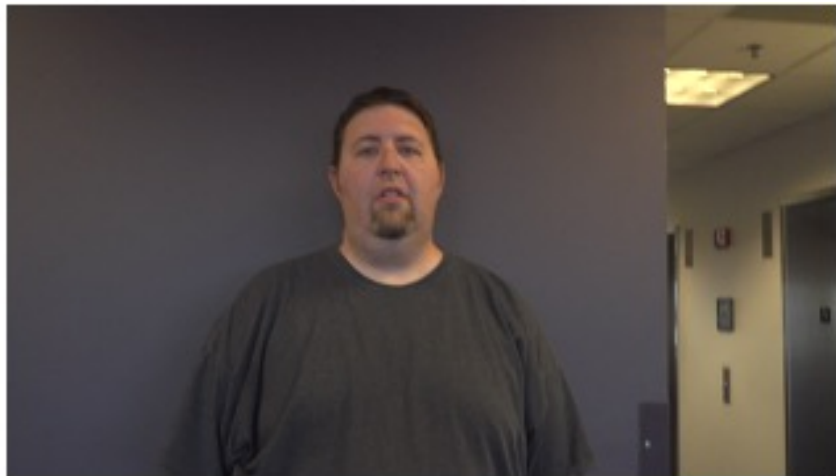
- Image Splice Manipulation Detection (ISD)
 - Purpose: reduce the task complexity for easy entry to the program
 - Good entry point for the research: college student, etc.
 - Approach: constrain on the scope with well-defined operation – Splice
 - Task: to detect if the image has been spliced
 - Condition: Image Only
 - Data: extracted 2K splice/original images from OpenMFC2020-2021 data
- Steganography Image Detection (StegD)
 - Collaborated with Jennifer's team for this task
 - To detect if the image contains stego
 - Condition: Image Only
 - Data: 480 stego + original images

OpenMFC Evaluation Protocol

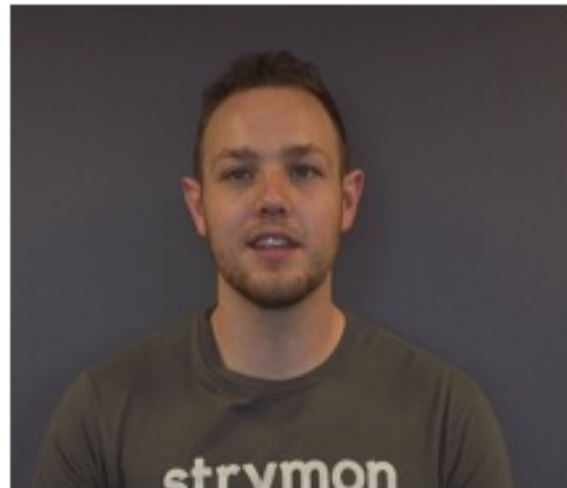
- Development data or previous evaluation data released by NIST is not recommended for any type of training
- Any machine learning or statistical analysis algorithm should complete training, model selection, and tuning prior to performing the detection task
- Trial Independence: Each trial in a task index file should be independent from each other.

Video Deepfakes Detection Test Example (1)

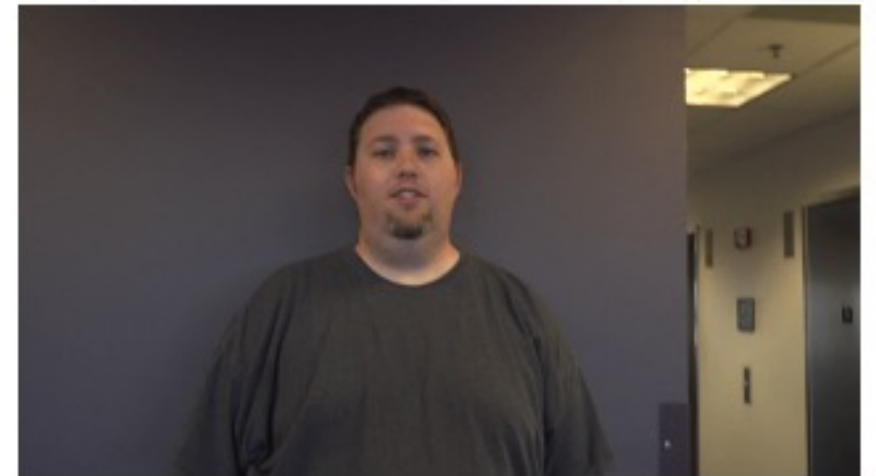
- NIST Deepfakes example
 - Original and donor videos are from MFC18 dataset, collected by UC Denver.
 - The deepfakes tool: DeepFaceLab
 - The deepfaked video was generated by NIST team (Ilia Ghorbanian's work)



(a) Original video



(b) Donor video

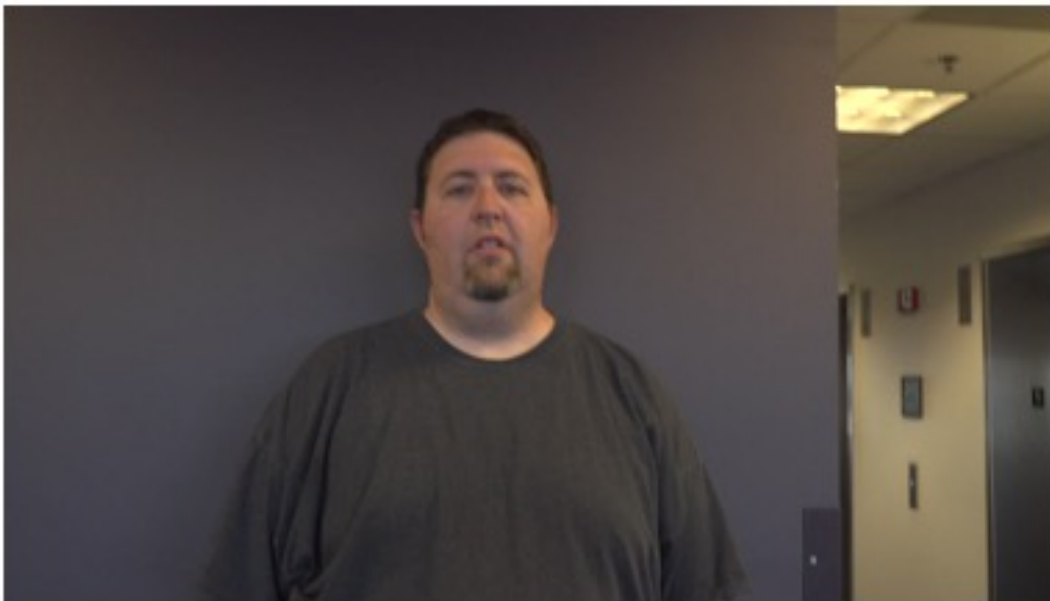


(c) Deepfaked video

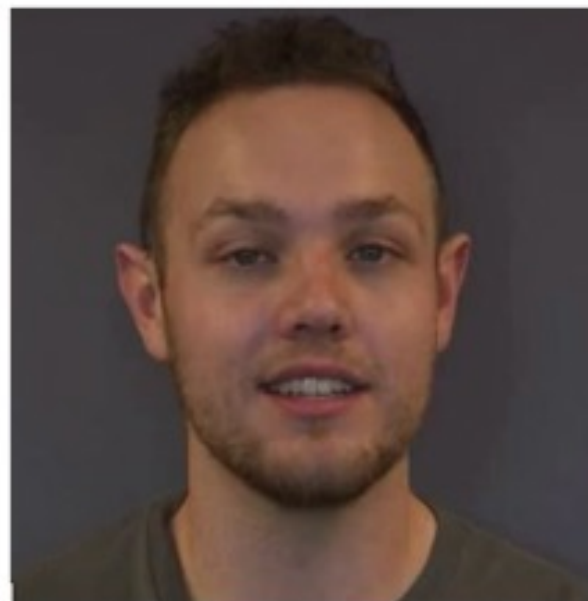
Figure: A Deepfaked video example using the MFC videos and the DeepFaceLab tool.

Video Deepfakes Detection Test Example (2)

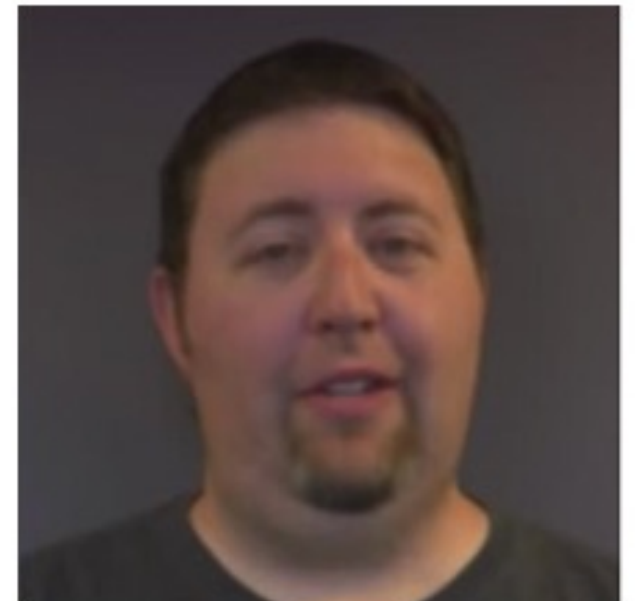
- Collaborated with Prof. Lyu's team in University at Buffalo.
- Tool: celeb-DF's face reenact tool
- Original and donor video are from MFC18 dataset, collected by UC Denver's team



(a) Original (target) video



(b) Source (facial expression donor) video



(c) Face reenact video of the similar facial expression

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

Stego Image Detection (StegD)

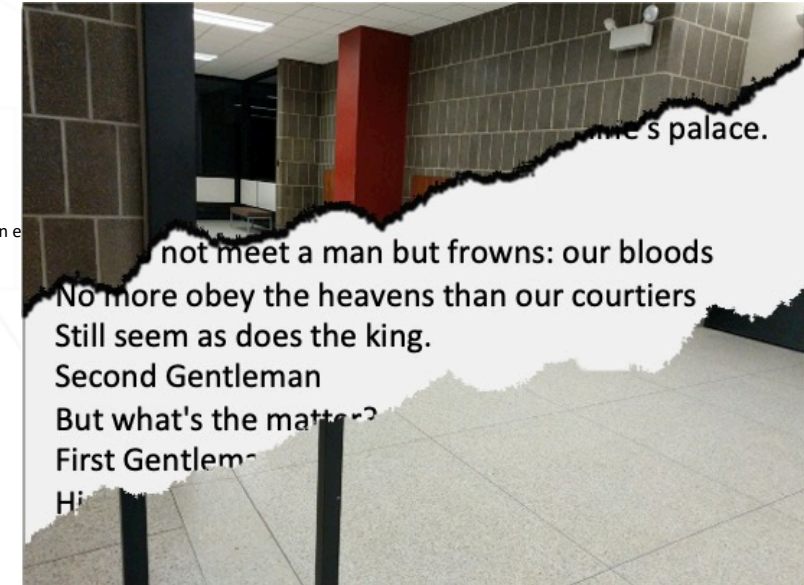
- Collaborated with Prof. Newman's team in ISU



Cover image

Cymbeline
ACT I
SCENE I. Britain. The garden of Cymbeline's palace.
Enter two Gentlemen
First Gentleman
You do not meet a man but frowns: our bloods
No more obey the heavens than our courtiers
Still seem as does the king.
Second Gentleman
But what's the matter?
First Gentleman
His daughter, and the heir of's kingdom, whom
He purposed to his wife's sole son--a widow

Payload



Steganographic image

“Steganography is the technique of hiding secret data within an ordinary, non-secret, file or message in order to avoid detection; the secret data is then extracted at its destination. The use of steganography can be combined with encryption as an extra step for hiding or protecting data.” Steganography hides data (payload) in an innocent file (cover), producing a steganographic image.

OpenMFC 2021-2022 Task Summary

Table: Tasks and Conditions

Task	Task Synopsis	Evaluation Condition		Dataset Size
		Media Only	Media + Metadata	
IMDL	Detect all image manipulations and localize non-global manipulations	Y	Y	16K
VMD	Detect all video manipulations	Y	Y	1.5k
IGMDL	Detect image GAN manipulations and localize non-global manipulations	Y	-	1.3k/TBD
VGMD	Detect all video GAN manipulations	Y	-	118/TBD
ISD	Detect image splice manipulation	Y	-	2k
StegD	Detect image steganography	Y	-	480

Nine OpenMFC 2021-2022 Leaderboards

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

- **Deepfakes/GAN Manipulation Detection**
 - IGMD-IO
 - 1.3k (OpenMFC 2020-2021/MFC18 data)
 - **New Dataset**
 - VGMD-VO
 - 118 (OpenMFC 2020-2021/MFC18 data)
 - **New Dataset**
- **Image Manipulation Detection and Localization:**
 - IMDL-IO: 16K (OpenMFC 2020-2021 MFC19)
 - Removed: IMDL-IM
 - **ISD-IO: 2k**
- **Video Manipulation Detection:**
 - Video Only (VMD-VO): 1.5K (OpenMFC 2020-2021/MFC19 data)
 - Video + Metadata (VMD-VM): 1.5K (OpenMFC 2020-2021/MFC19 data)
- **Steganography Manipulation Detection**
 - **StegD-IO: 480**

Takeaway

- OpenMFC 2021-2022 evaluation task
 - Keep old tasks
 - Proposed two new tasks
- Leaderboard upgrade
 - New leaderboards
 - Image Splice Detection
 - Steg Detection
- OpenMFC 2021-2022 datasets
 - New: Steg detection dataset
 - Ongoing:
 - GAN image detection dataset
 - Deepfakes video dataset

What's next?

- Workshop publication
 - Slides/Papers/Reports etc.
- Community engagement and collaboration
- NIST
 - OpenMFC2021-2022 Eval. Plan
 - ISD dataset release
 - Deepfakes dataset generation
 - Leaderboard implementation and submission pipeline testing
- Researchers:
 - **Join the OpenMFC program!**

Questions?

OpenMFC team: mfc_poc@nist.gov

The background of the slide is a light gray with a complex, abstract pattern of thin lines and dots, resembling a network or circuit board. There are also larger, faint geometric shapes like rectangles and circles scattered across the background.

Thank You!

Feedback and Discussion: OpenMFC 2021-2022 Evaluation

Haiying Guan

Ilia Ghorbanian, Lukas Diduch, and Yooyoung Lee

Multimodal Information Group,
Information Access Division

OpenMFC2021 Workshop : Day 3, Thursday, Dec. 9, 2021

OpenMFC 2020-2021

Video Deepfakes Detection Evaluation Dataset

- NIST team:
 - Ilia Ghorbanian, Haiying Guan, Lukas Diduch, and Yooyoung Lee
- External collaborators :
 - Prof. Siwei Lyu, Shan Jia, and Yan Ju in University at Buffalo

Deepfakes Evaluation Dataset: Objective

- Benchmark dataset for evaluation
 - Lab algorithm evaluation vs. in-the-field evaluation (real-world application)
 - Avoid the potential pitfall
 - Reduce the systematic transition difficulties
 - Bridge the gap
 - Continuous year-to-year report
- Adapt to the dynamic evaluation updates:
 - emerging software and tools (GAN, Deepfakes, CGI, etc.)

OpenMFC Deepfakes Discussion Topics

1. Are there any public releasable Deepfaked data?
2. Are there any public releasable real face image/video datasets that you're aware of?
3. What Deepfakes tools/algorithm you know? Which one are you working on?
4. What is the key factors you can think for generating high quality Deepfaked videos?
5. What are the media forensic technologies for deepfake detection as you known?
6. What is the key factors do you think that can affect the media forensic detection technologies significantly?
7. What are the most common artifacts that Deepfake tool can leave behind?
8. What is the minimum amount of videos is expected for deepfake detection training/testing?

Available Deepfakes Tools

Name	Reference	Copyright	Initial Release	Input/output	Required Hardware
First Order Motion Model for Image Animation	arxiv.org/abs/2003.00196	CC BY-NC 4.0	Dec 2019	GIF- video+image/GIF- video	Nvidia GPU - CPU
DeepFaceLab	arxiv.org/abs/2005.05535	GPL-3.0	Jun 2018	Video + video/ video	Nvidia GPU - CPU
FakeApp	malavida.com/en/soft- /fakeapp/	“free”	Jan 2018	Video + Video/Video	Nvidia GPU
Deepfakes Faceswap	faceswap.dev	GPL-3.0	Dec 2017	Video + Video – Image/Video	Rec: Nvidia GPU Min: CPU
Celeb-DF	arxiv.org/abs/1909.12962	Not available for public use	Sep 2019	Video + Video/Video	Nvidia GPU
Reface	Reface.app	reface.app/ terms/	Feb 2020	Image + Choose a video/video	Android: 5.1 and up IOS: 13 or later
FaceApp	Faceapp.com	www.faceapp.co m/terms-en.html	Feb 2017	Image + Image/Image	Android: 5.3 and up IOS: 13 or later

Available GAN Models (1) *

Name	Reference	Copyright	Initial Release	Input/Output	Required Hardware
StyleGan	github.com/NVlabs/stylegan	Nvidia Source Code License	Feb 2019	Image-Image/Image	12GB Nvidia GPU
StyleGan2	github.com/NVlabs/stylegan2	Nvidia Source Code License	Feb 2020	Image-Image/Image	12GB Nvidia GPU
StyleGan3	github.com/NVlabs/stylegan3	Nvidia Source Code License	Oct 2021	Image-Image/Image	12GB Nvidia GPU
Pix2Pix	arxiv.org/abs/1611.07004	github.com/phillipi/pix2pix/blob/master/LICENSE	2016	Image-Image/image Video-image/video	Nvidia GPU
SN-GAN	arxiv.org/abs/1802.05957	IT License	Feb 2018	Image-Image/Image	Nvidia GPU

Available GAN Models (2) *

Name	Reference	Copyright	Initial Release	Input/Output	Required Hardware
MMD-GAN	arxiv.org/abs/1705.08584	BSD-3-Clause	Jan 2018	Image-Image/Image	Nvidia GPU
Glow	arxiv.org/abs/1807.03039	MIT License	Jun 2018	Image-Image/Image	Nvidia GPU
ProGAN	arxiv.org/abs/1710.10196	Attribution-NonCommercial 4.0 International	Oct 2017	Image-Image/Image	Nvidia GPU
PixelCNN	arxiv.org/abs/1606.05328		Nov 2016	Image-Image/Image	Nvidia GPU

Available Datasets (1) *

Name	Reference	Copyright	Initial Release	Size	Description
FFHQ	github.com/NVlabs/ffhq-dataset	CC BY-NC-SA 4.0	Jun 2019	2.56 TB	Human face dataset with over 70,000 high-quality PNG images.
CelebA-HQ	github.com/taiki-as/progressive_growing_of_gans	CC BY-NC 4.0	Oct 2017	89 GB	The CelebA-HQ dataset is a high-quality version of CelebA that consists of 30,000 human face images
LSUN	www.yf.io/p/lsun		Jun 2015	21.93 GB	Large-scale Scene Understanding Challenge dataset consists of 10 scene categories and 20 object categories.
ImageNet	image-net.org		2010	150 GB	This dataset spans 1000 object classes and contains 1,281,167 training images, 50,000 validation images and 100,000 test images
CityScapes	cityscapes-dataset.com	www.cityscapes-dataset.com/license/	2016	10.86 GB	large-scale dataset that contains a diverse set of stereo video sequences recorded in street scenes from 50 different cities, with high quality pixel-level annotations of 5 000 frames in addition to a larger set of 20 000 weakly annotated frames.

Available Datasets (2)*

Name	Reference	Copyright	Initial Release	Size	Description
COCO-stuff	github.com/nghyhtrome/cocostuff	COCO images: Flickr Terms of use COCO annotations: Creative Commons Attribution 4.0 License COCO-Stuff annotations & code: Creative Commons Attribution 4.0 License	Mar 2018	18 GB	COCO-Stuff augments all 164K images of the popular COCO dataset with pixel-level stuff annotations
ADK20k	sceneparsing.csail.mit.edu/	BSD 3-Clause License	2016	1.14 GB	Images of daily scenes.
AFHQ v2	github.com/clovaai/stargan-v2	Attribution-NonCommercial 4.0 International	2021	6.48 GB	High resolution images of animal faces
VGGFace2	github.com/oxvgg/vgg_face2		2018	40.25 GB	dataset contains 3.31 million images of 9131 subjects, with an average of 362.6 images for each subject

FaceForensics++¹

Deepfake video dataset containing 1,000 videos with sources and target ground truth

Videos created using:

- Face2Face and NeuralTextures for facial reenactment
- Deepfakes and FaceSwap for the face swap process

DeepFake Detection Challenge Dataset¹

Facebook's commissioned dataset containing over 100,000 videos with 3,426 subject, resulting in over 25 TB of raw data.

- Average of 14.4 videos per person with most videos being shot in 1080p
- Realistic and nonrealistic examples included to represent most possibilities
- Videos generated using different methods for more variation

Methods used:

- NeuralTalkingHeads
- DFAE model
- FaceSwap
- FSGAN
- StyleGAN
- Manual Retouching

Celeb-DF¹

A Deepfake dataset with the goal of having higher quality examples of Deepfaked media, containing 5,369 examples of Deepfaked videos.

- 5,369 videos corresponding to over 2 million frames
- Sourced from publicly available Youtube videos of interviews with celebrities.
- Proven to be a more challenging dataset compared to FaceForensics++
- Higher resolution on the Deepfaked faces using encoder and decoders with more layers and increased dimensions

Deepfakes Tool Study

Media manipulation tools that are being tested or are planned to be tested:

DeepFaceLab

FaceApp (mobile application)

Reface (mobile application)

First Order Motion Model for Image Animation

Deepfakes FaceSwap

DeepSwap

FakeApp

DeepFaceLab¹

DeepFaceLab is one of the first publicly available and Open Source Deepfake tools and ever since its release on Github on 2018, it has been seeing constant updates and tweaks to its code and algorithm.

- Available on MacOS, Windows, and Linux
- Relatively easy to use software for cropping, aligning, and swapping the faces automatically.
- The color correction is done manually by the user, can be done either by hand in a video editing software or using the provided tools.
- Input is two videos, one source and the other destination, with the output being one video.
- The first video created by a fresh model took about 16 hours to look natural

¹<https://github.com/iperov/DeepFaceLab>, <https://arxiv.org/abs/2005.05535>

DeepFaceLab Example



- Video created on a model that is trained for about 24 hours, and took about 4 hours to get to this place.

Deepfake FaceSwap¹

FaceSwap is also a publicly available and Open Source Deepfake tool with a relatively large community forum and number of tutorials.

- GUI application for ease of use
- The code can also be downloaded and used because of the Open Source nature of the project
- Input can be an image + video or video + video, with the image being source and video the destination.
- Everything is done automatically by the software, no manual retouching needed
- Needs the most amount of time out of all the other tools, with the recommended minimum time being 48 hours of training and the recommended time being one week of training

¹ <https://www.faceswap.dev>

Reface¹

Similar to FaceApp, Reface is also a mobile application available on Android and IOS which allows for simple Deepfake videos with specific clips of movies or music videos.

- Available on both Android on IOS free of charge with watermark, and a subscription to remove the watermark
 - Input is only one picture
 - User can choose one of the specific clips available in the app for the faceswap
- Generates the results within minutes, if not seconds. The generated videos are usually fairly impressive

¹ <https://reface.app/>

Questions?

OpenMFC team: mfc_poc@nist.gov

The background of the slide is a light gray with a complex, abstract pattern of thin lines and dots, resembling a network or circuit board. There are also larger, faint geometric shapes like rectangles and circles scattered across the background.

Thank You!



Open Media Forensics Challenge (OpenMFC) 2021 Workshop

Closing Remarks

Jim Horan

Multimodal Information Group,
Information Access Division

Dec. 7-9, 2021

