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

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

INFORMATION TECHNOLOGY LABORATORY

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 Fraud[1]

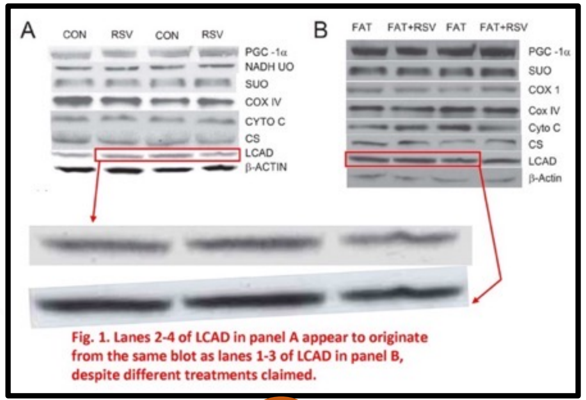
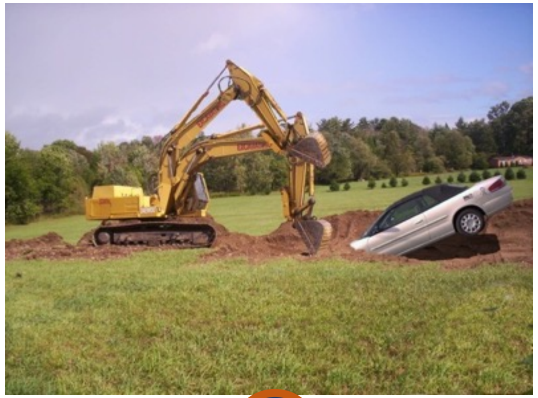


Fig. 1. Lanes 2-4 of LCAD in panel A appear to originate from the same blot as lanes 1-3 of LCAD in panel B, despite different treatments claimed.

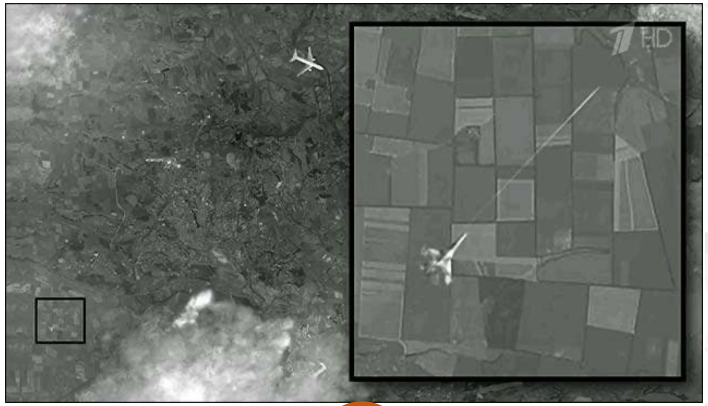
Insurance Fraud[2]



Social Media[3]

[redacted]

News/Magazines[4]



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

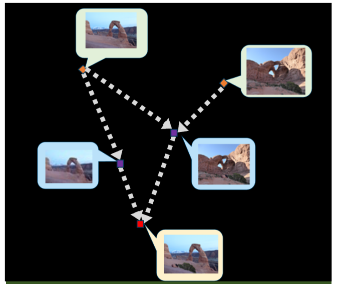
NIST role: define tasks and metrics, and manage technical evaluations of media forensic technologies

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

Deepfake

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

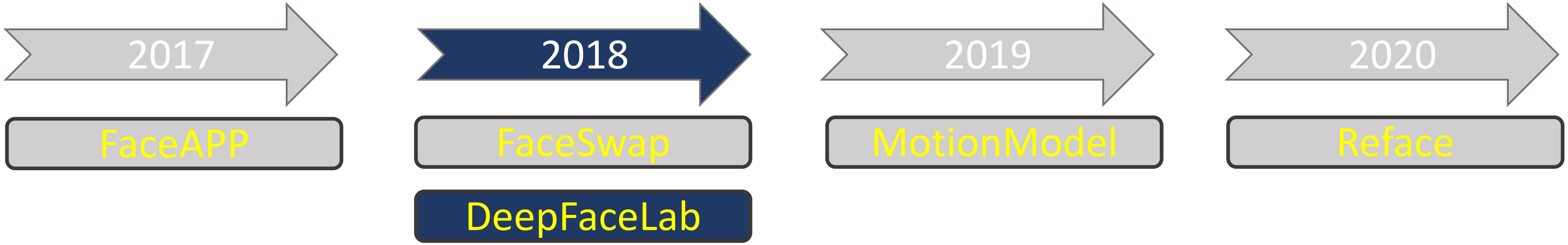
Completed Test

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

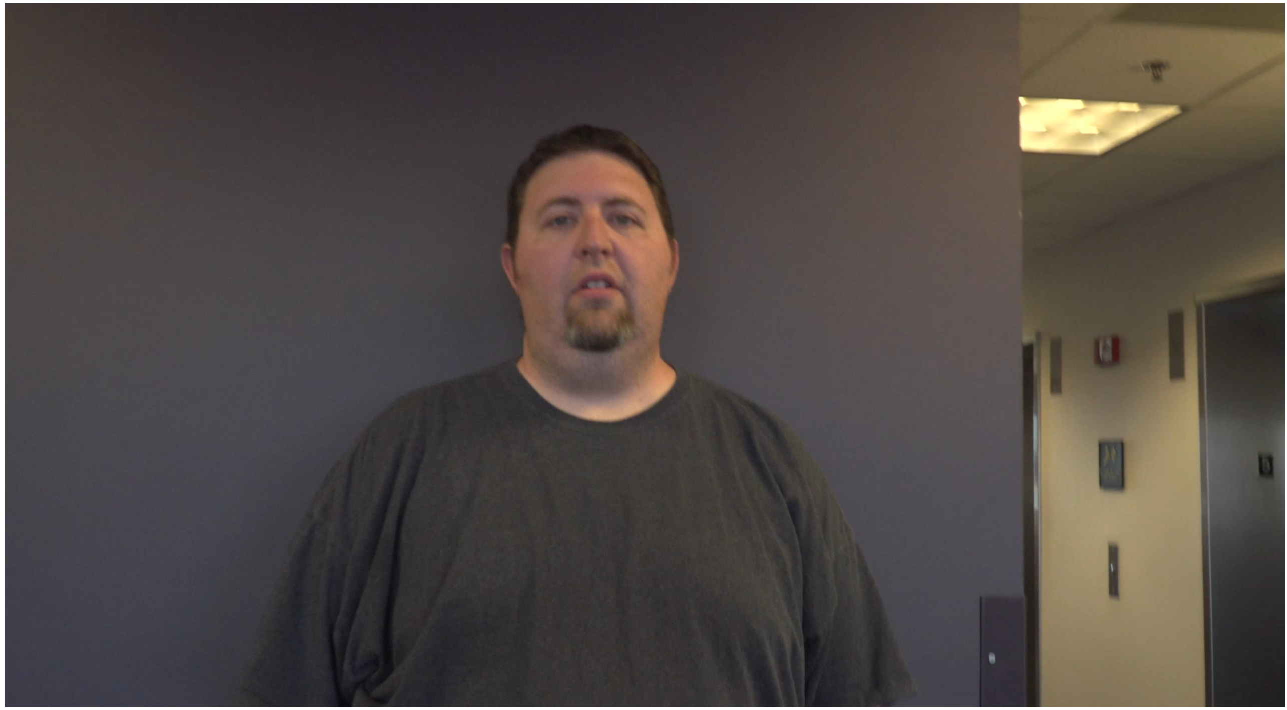
GAN Datasets	Release
CityScapes, ADK20k	2016
CelebA-HQ,	2017
COCO-stuff, VGGFace2	2018
FFHQ	2019
AFHQ v2	2021



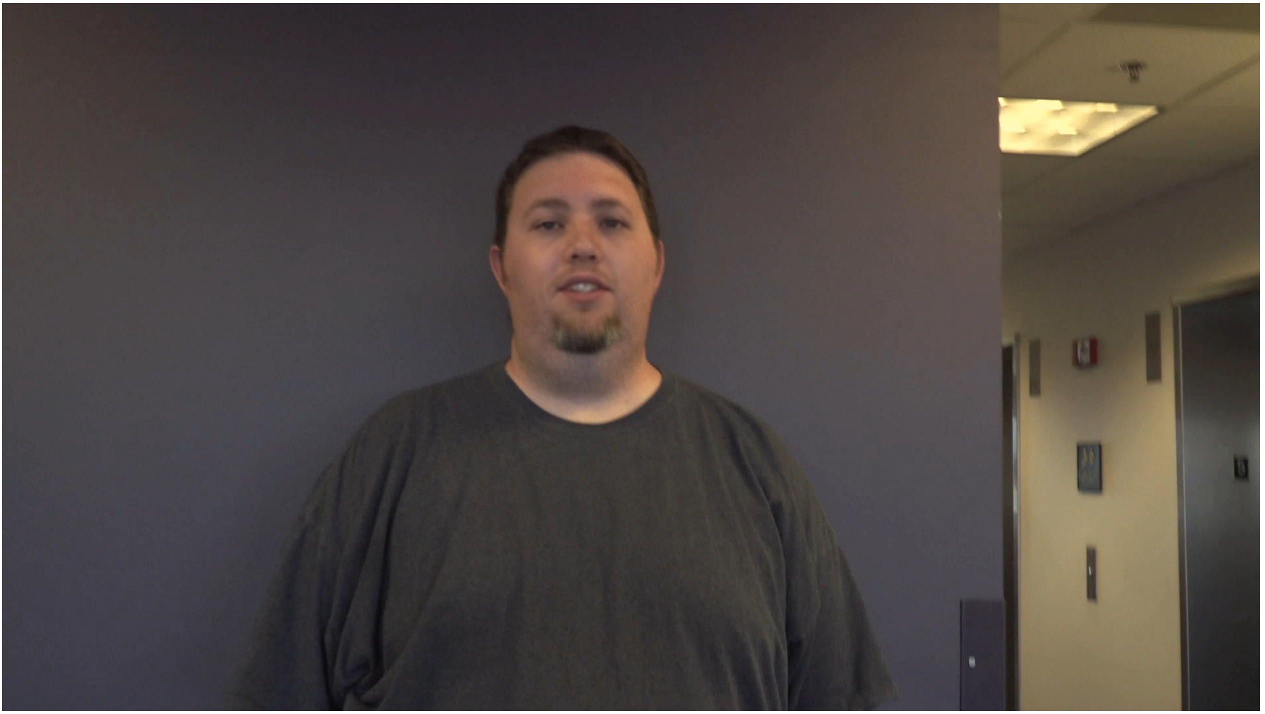
2017 | 2018 | 2019 | 2020

FaceAPP | FaceSwap | MotionModel | Reface

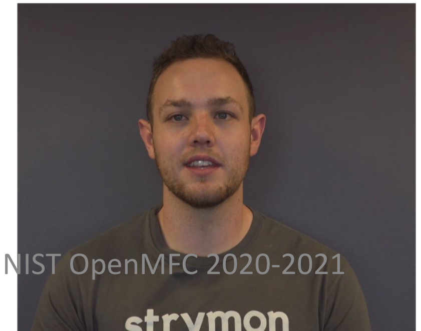
DeepFaceLab



Base



Deepfaked



Donor

2017

FaceAPP

2018

FaceSwap

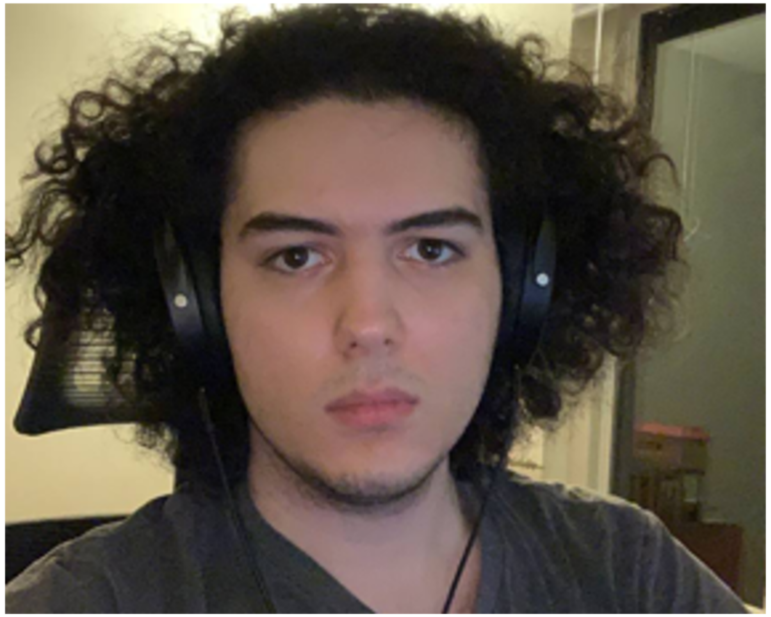
DeepFaceLab

2019

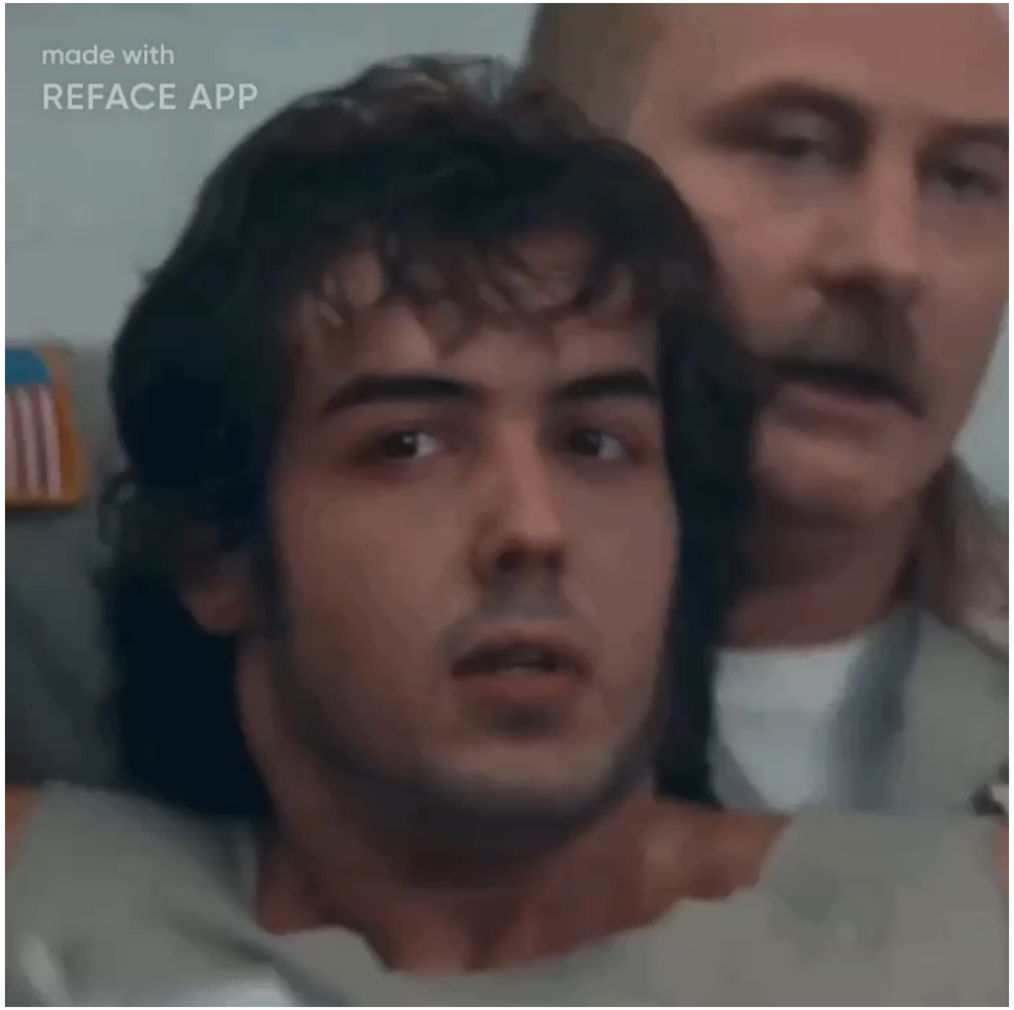
MotionModel

2020

Reface



Donor



Deepfaked

Web-based Leaderboard

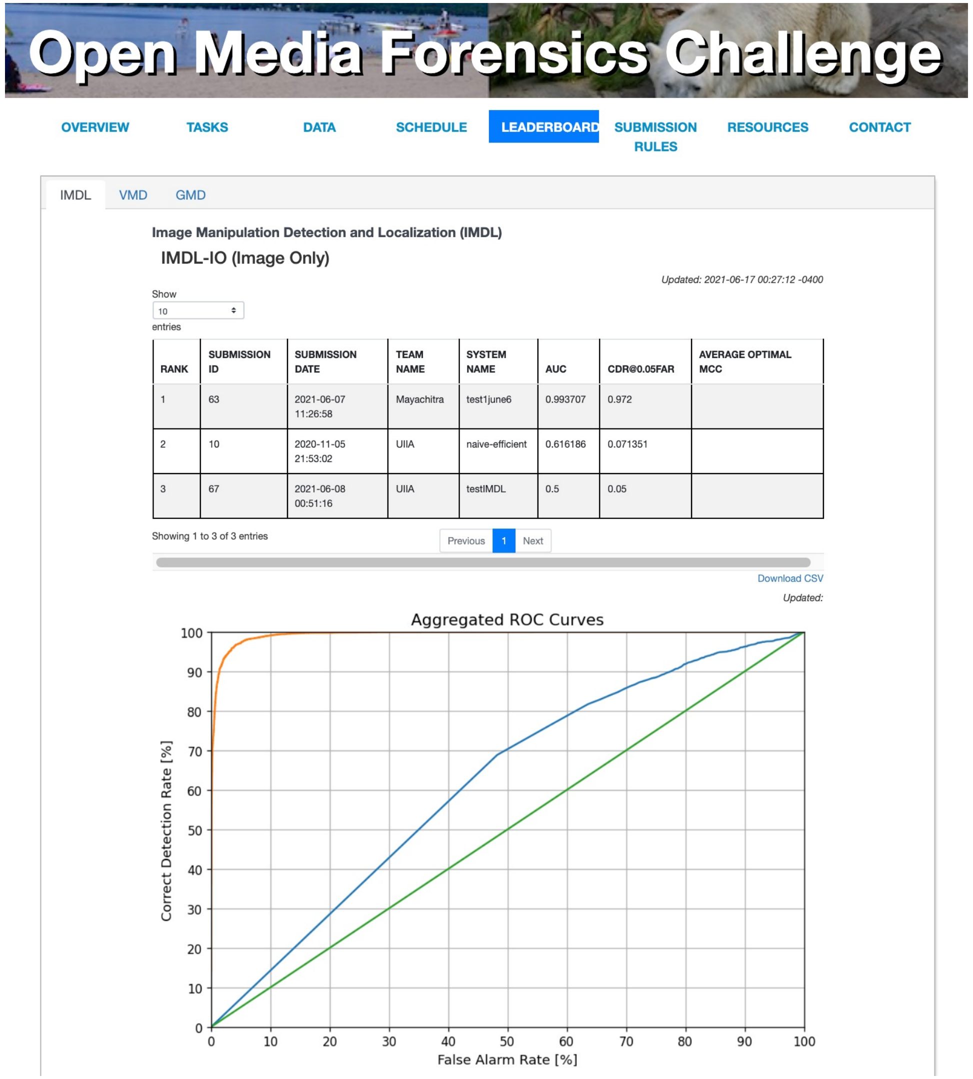


Image Manipulation Detection and Localization (IMDL)

IMDL-IO (Image Only)

Updated: 2021-06-17 00:27:12 -0400

Show 10 entries

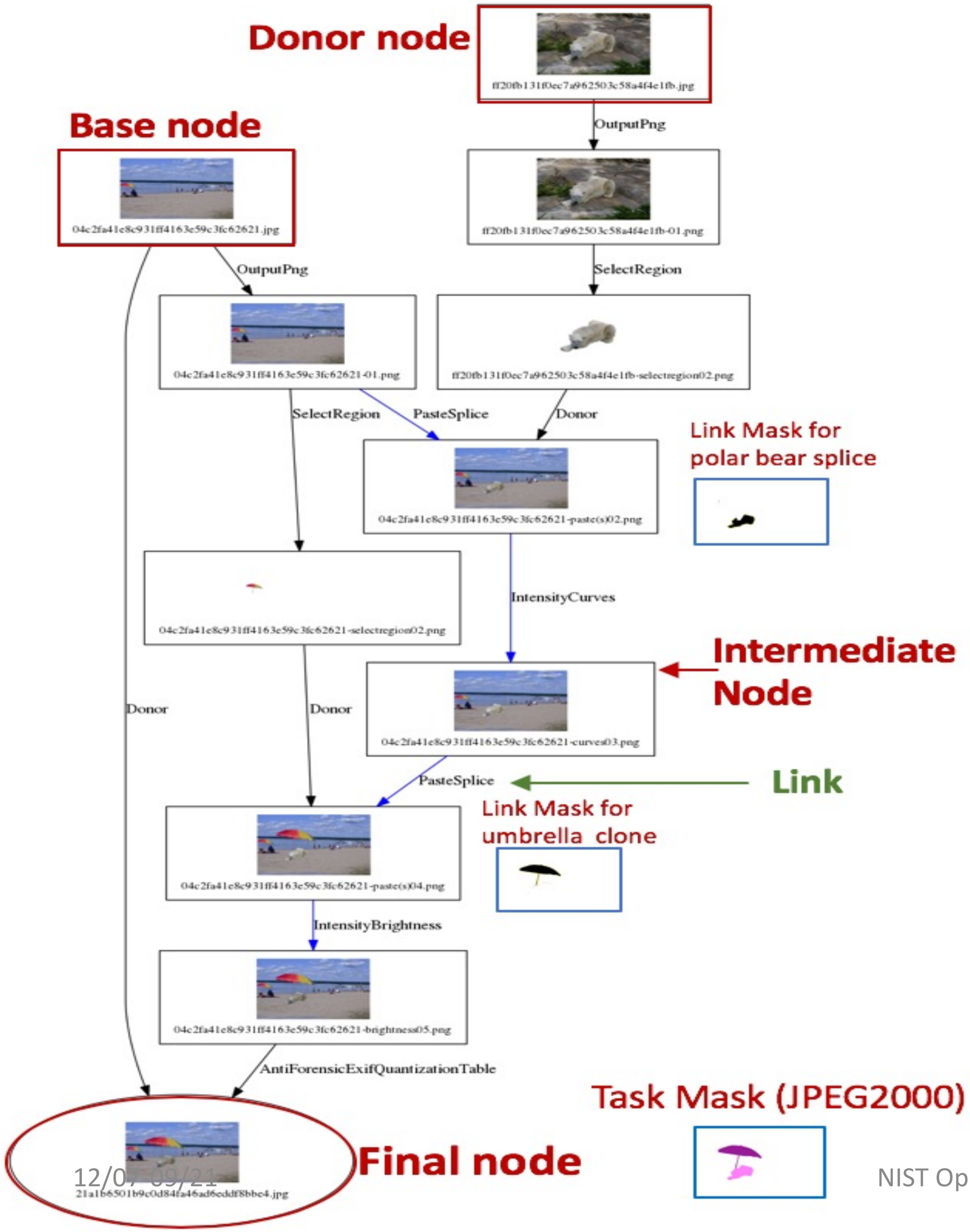
RANK	SUBMISSION ID	SUBMISSION DATE	TEAM NAME	SYSTEM NAME	AUC	CDR@0.05FAR	AVERAGE OPTIMAL MCC
1	63	2021-06-07 11:26:58	Mayachitra	test1june6	0.993707	0.972	
2	10	2020-11-05 21:53:02	UIIA	naive-efficient	0.616186	0.071351	
3	67	2021-06-08 00:51:16	UIIA	testIMDL	0.5	0.05	

Showing 1 to 3 of 3 entries

Updated:

Quick Turnaround Leaderboard Evaluation

https://mfc.nist.gov


Donor node
ff20fb131f0ec7a962503c58a4f4e1fb.jpg
Base node
04c2fa41e8c931ff4163e59c3fc62621.jpg
OutputPng
ff20fb131f0ec7a962503c58a4f4e1fb-01.png
OutputPng
SelectRegion
04c2fa41e8c931ff4163e59c3fc62621-01.png
ff20fb131f0ec7a962503c58a4f4e1fb-selectregion02.png
SelectRegion
PasteSplice
Donor
Link Mask for polar bear splice
04c2fa41e8c931ff4163e59c3fc62621-paste(s)02.png
04c2fa41e8c931ff4163e59c3fc62621-selectregion02.png
IntensityCurves
Intermediate Node
Donor
Donor
04c2fa41e8c931ff4163e59c3fc62621-curves03.png
PasteSplice
Link
Link Mask for umbrella clone
04c2fa41e8c931ff4163e59c3fc62621-paste(s)04.png
IntensityBrightness
04c2fa41e8c931ff4163e59c3fc62621-brightness05.png
AntiForensicExifQuantizationTable
Task Mask (JPEG2000)
Final node
21a1b6501b9c0d84fa46ad6eddf8bbe4.jpg

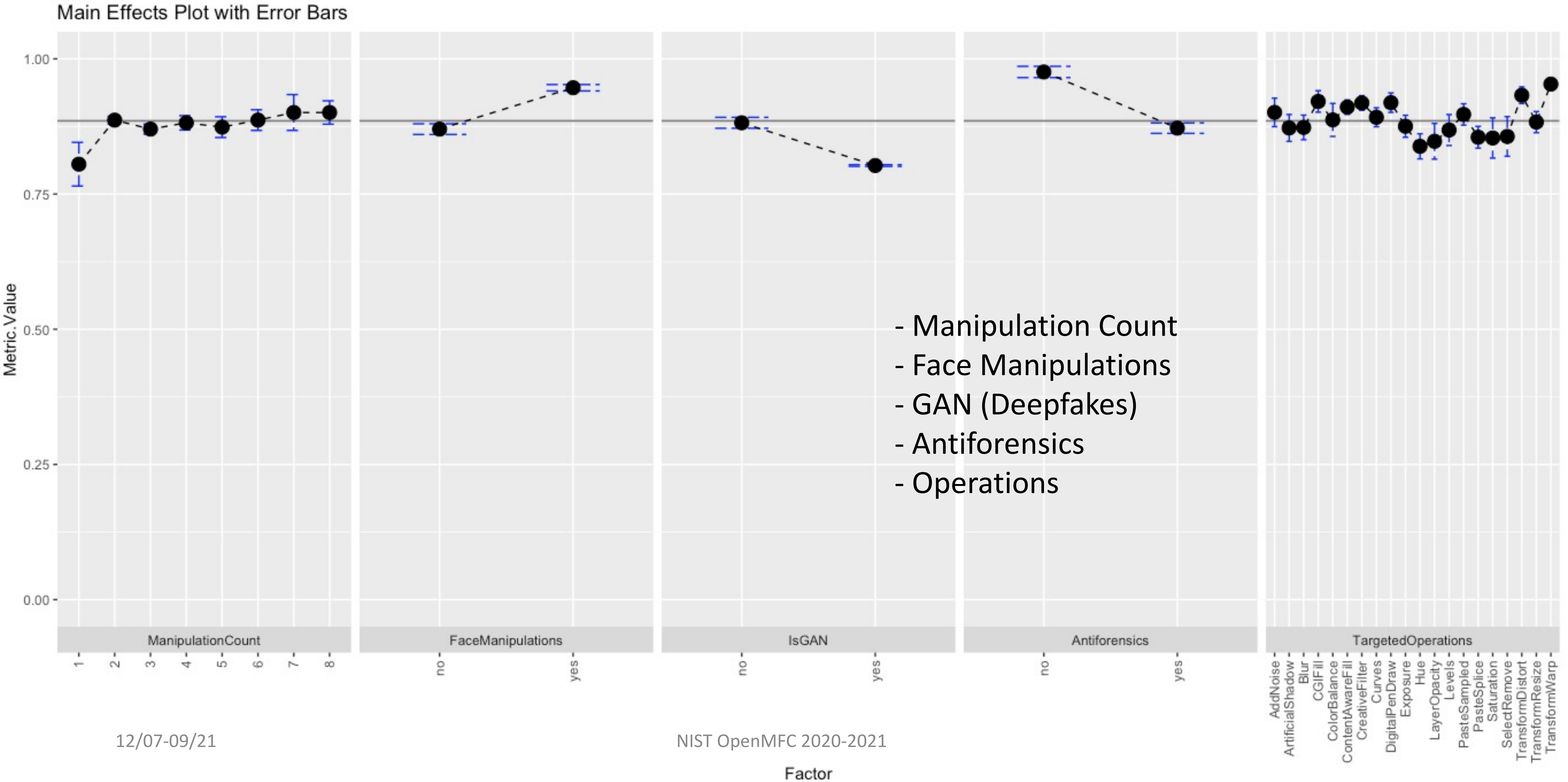
Web-based Media Level Analysis for Validation Set (Provenance)

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?



- Manipulation Count
- Face Manipulations
- GAN (Deepfakes)
- Antiforensics
- Operations

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