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

Discussion:
Holistic vs. Opt-In

Holistic vs. Opt-In Technologies

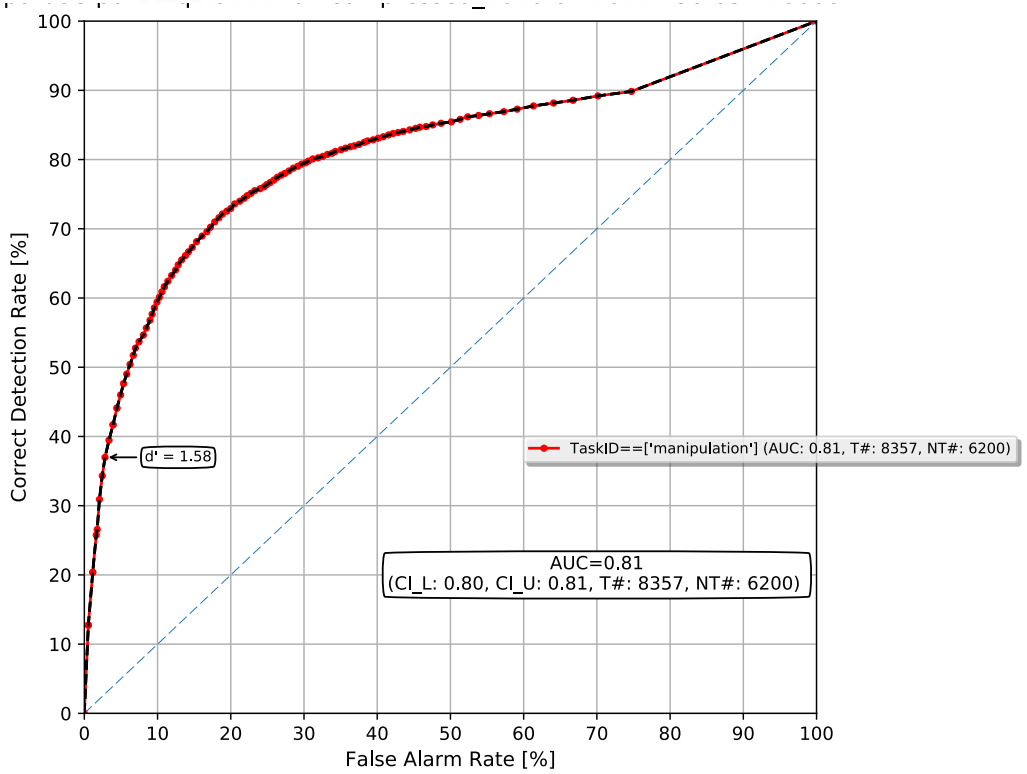
- Evaluation challenge:
  - Some media forensic systems have the option to return no response
    - E.g., face illumination consistency systems should not respond if no face was found in the image

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

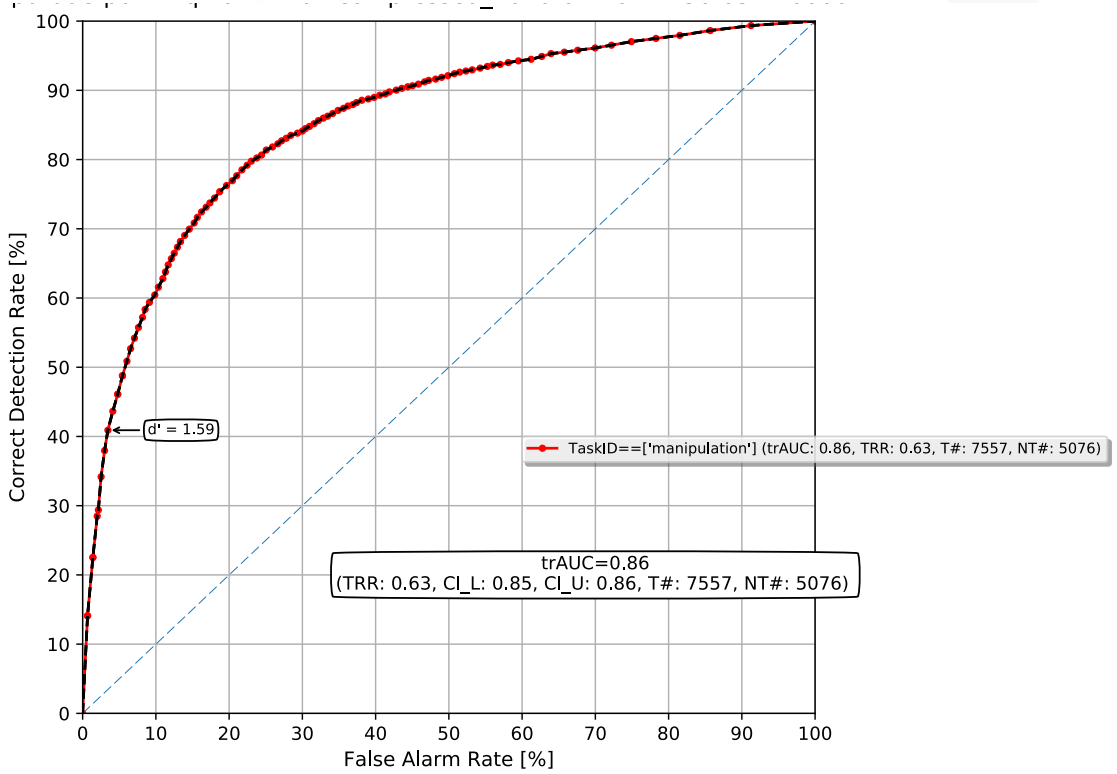
- NIST reports:
  - Holistic performance measures: score all trials
  - Opt-In performance measures:
    - Trial Response Rate (TRR) – Percent of Processed, NonProcessed, and FailedValidation images
    - Performance measures excluding opt'd out probes

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

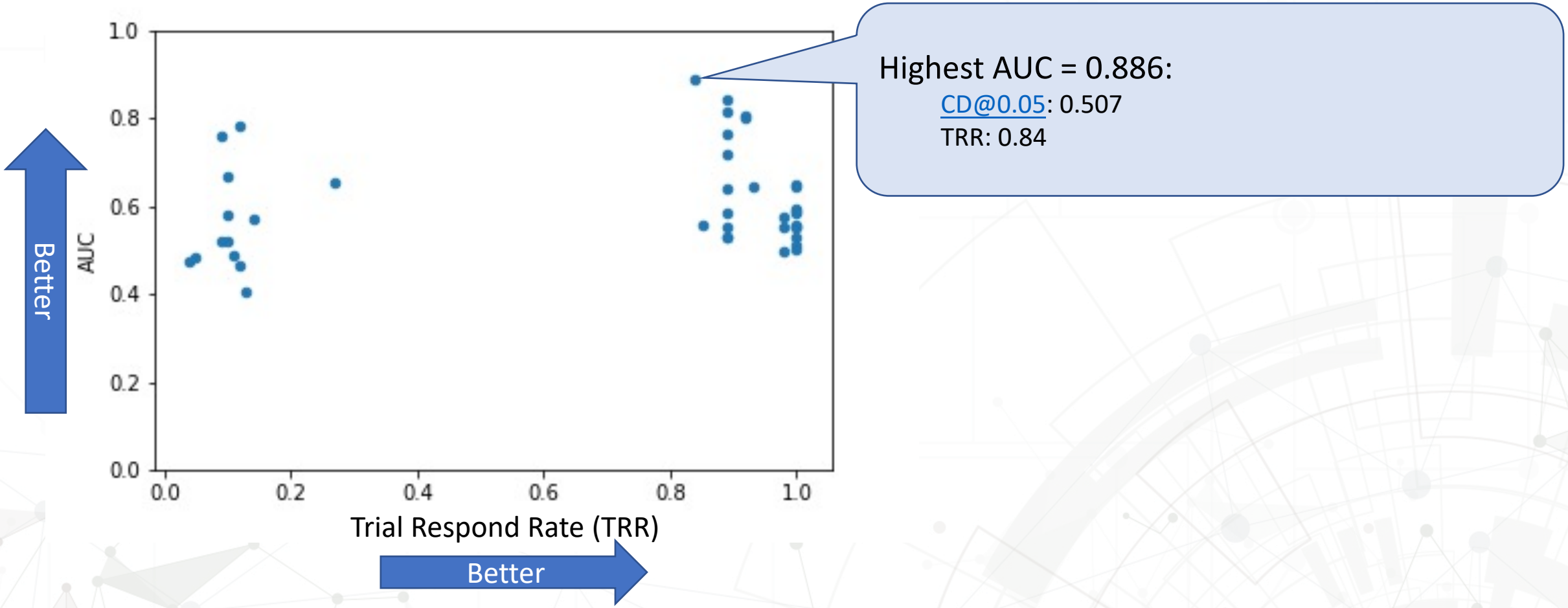


(a) Holistic



(b) Opt-In

MediFor IMD: Opt-In



Highest AUC = 0.886:

CD@0.05: 0.507

TRR: 0.84

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 techinques locally/globally	530
NonGAN-CGI	Any operations that use non-GAN CGI	309
Distortion	Deformation of images	918
Remove	Remove a set of pixels.	833
Face Manipulation	Any manipulation done to a face.	22
All	All data without selective scoring	5750

Selective Scoring for System Performance Analysis

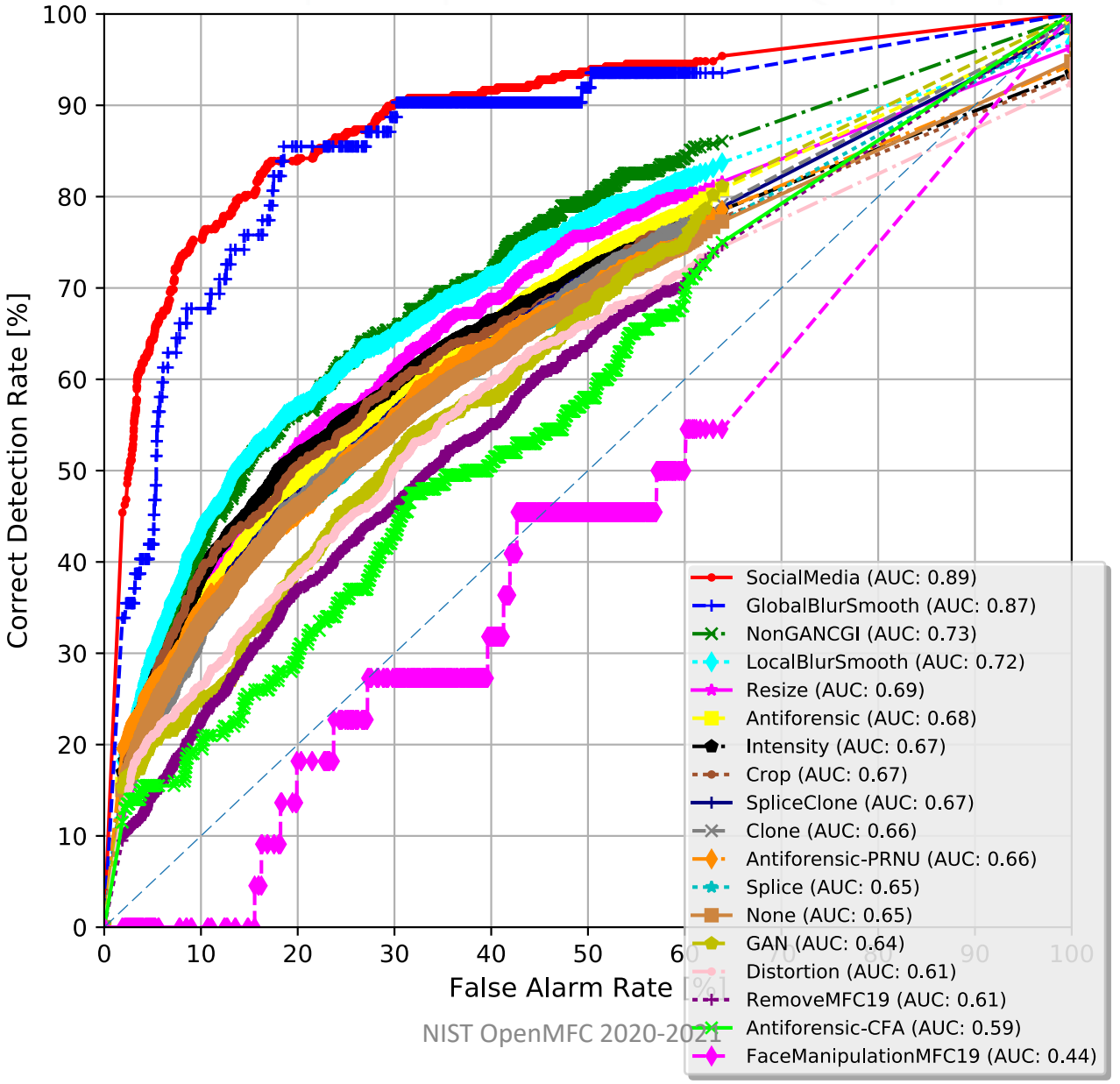


Image Detection Selective Scoring – Top 10 system AUC

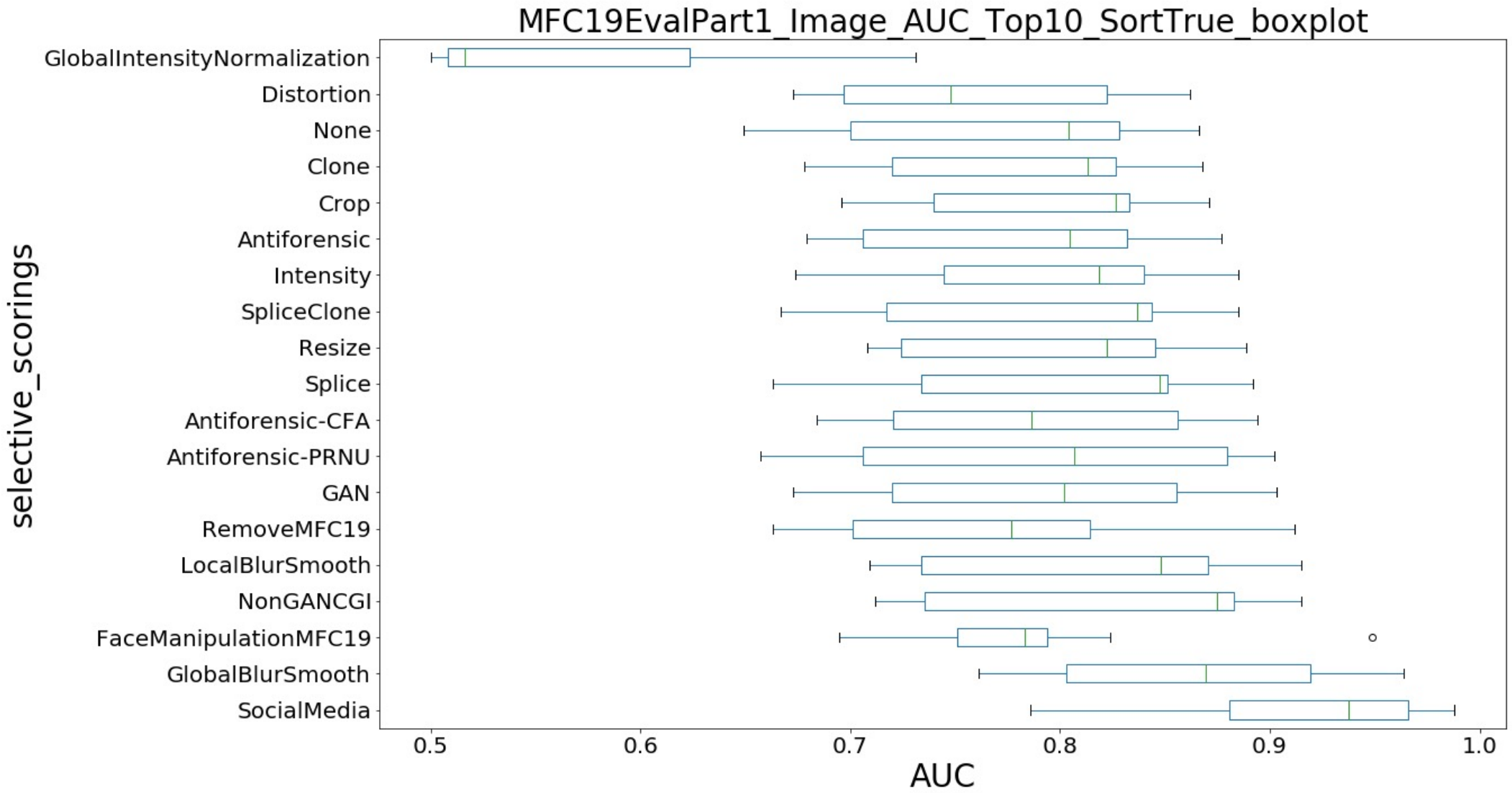
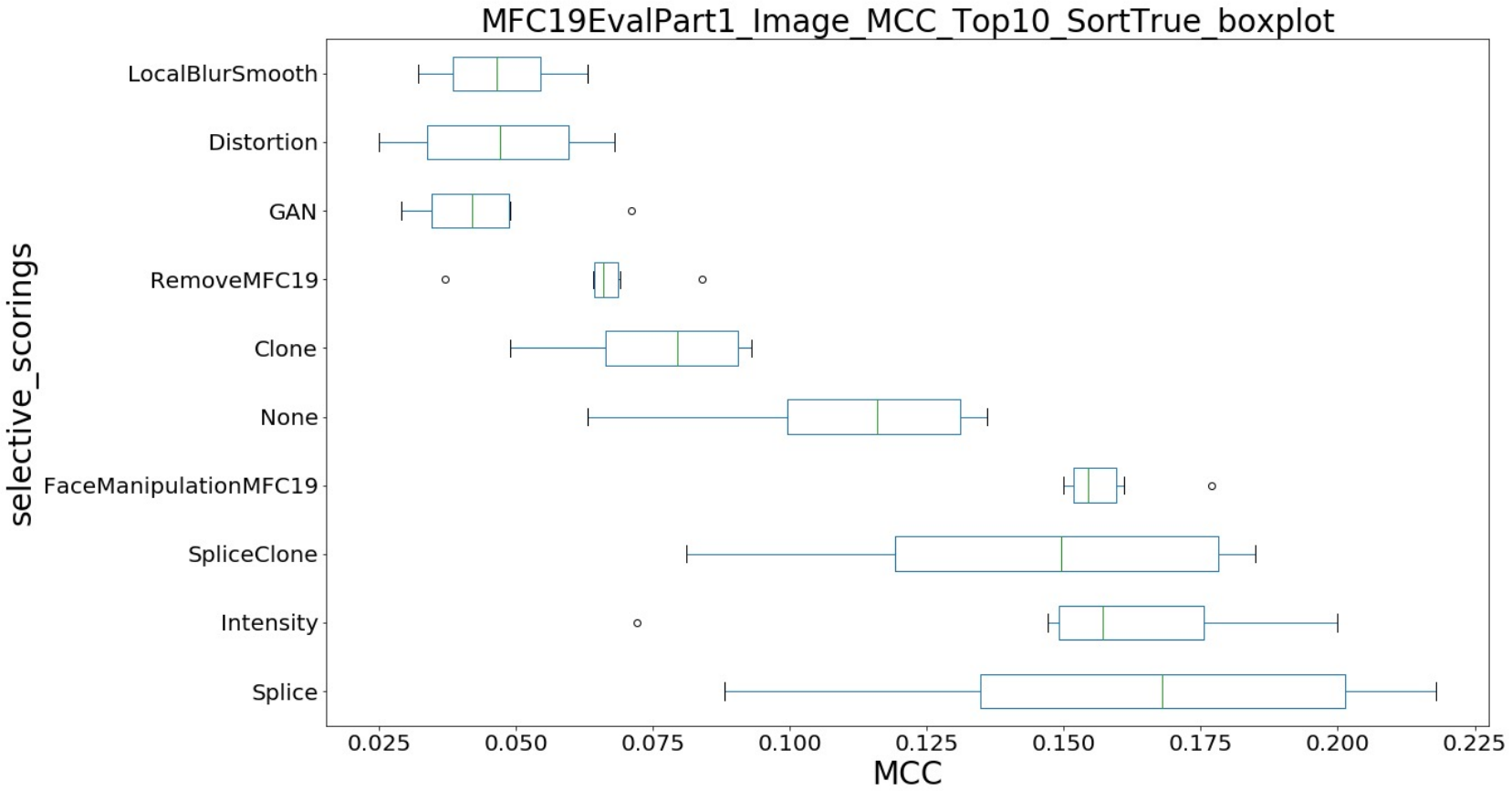


Image Localization Selective Scoring – Top 10 system MCC



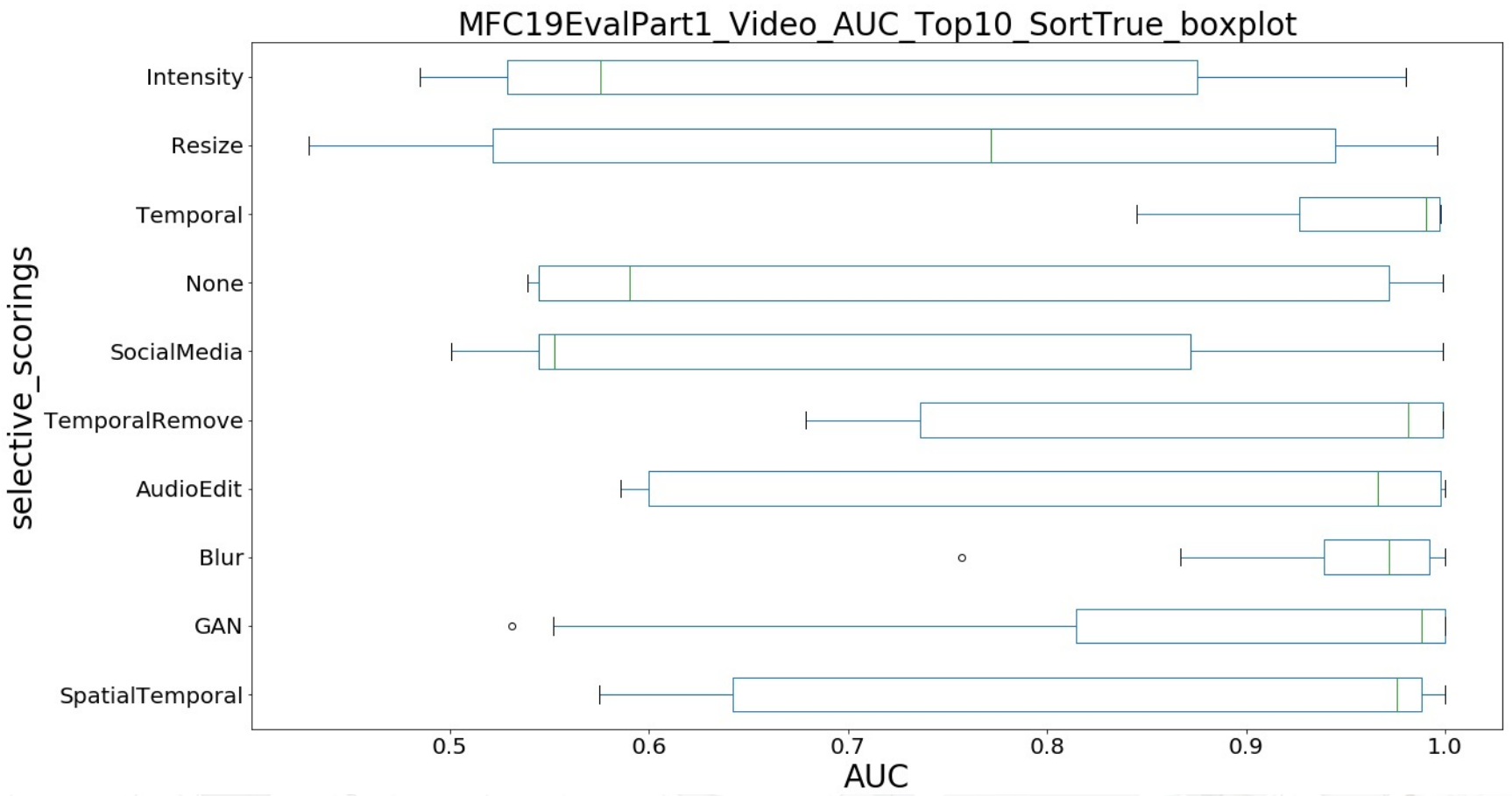
Video Selective Scoring

- 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!