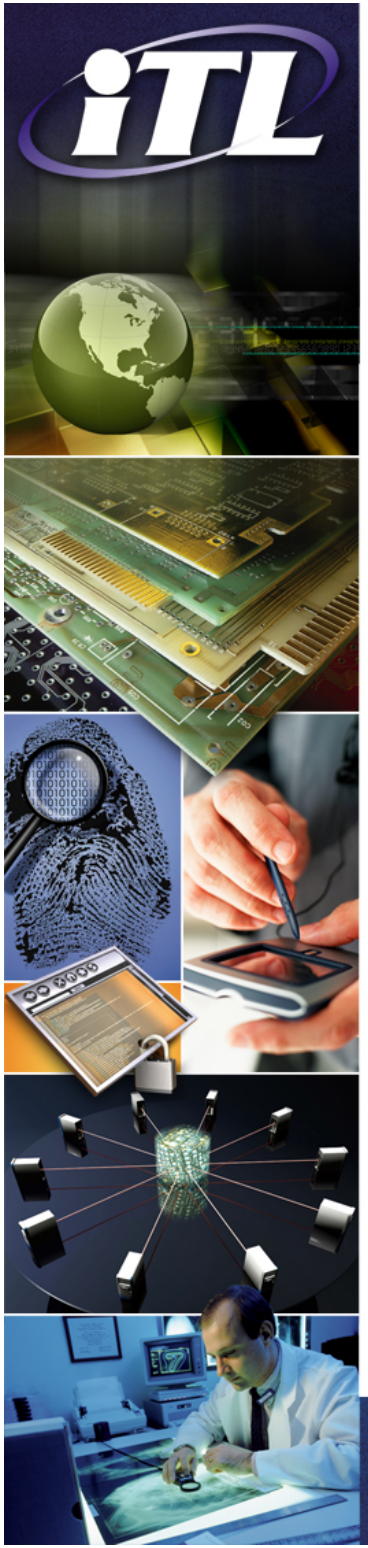


## Media Forensics

# The 2017 ~~Nimble~~ Challenge Evaluation: Results and Future Directions

CVPR Workshop on Media Forensics
July 26, 2017

**Jonathan Fiscus**
jfiscus@nist.gov

Dr. Haiying Guan ,Dr. Yooyoung Lee,
Dr. Amy Yates, Andrew Delgado, Daniel Zhou, David Joy, August Pereira

Multimodal Information Group
Information Access Division
Information Technology Laboratory
National Institute of Standards and Technology (NIST)

# You Will Learn About:

- The motivation behind the Nimble ‘17 evaluation
- The results of the baseline evaluation for four evaluation tasks
- How the Media Forensics Challenge supports many research goals
- How to participate in the 2018 Media Forensics Challenge

# Media Forensics

- Digital media manipulation is entertaining
  - Social media filters
  - CGI/Movies
- Digital media manipulation is nefarious
  - Fraud
  - Disinformation

# Media Forensics:

## Fictitious Insurance Fraud Example

**Claim:**

Jack's Excavating failed to protect their work site from traffic allowing Mr. Smith to drive his car into the work zone crashing into a ditch on the 12th of December in Clarion, PA.

# Translating the Use Case Into Research Tasks

- Is the image manipulated?
- Where do the manipulations spatially occur?
- What operations were performed?
- Is there an original image?
- Are there related images?
- Is the image consistent with the camera?
- Are there known examples of vehicles?
- Is the image consistent with the reported date and location?

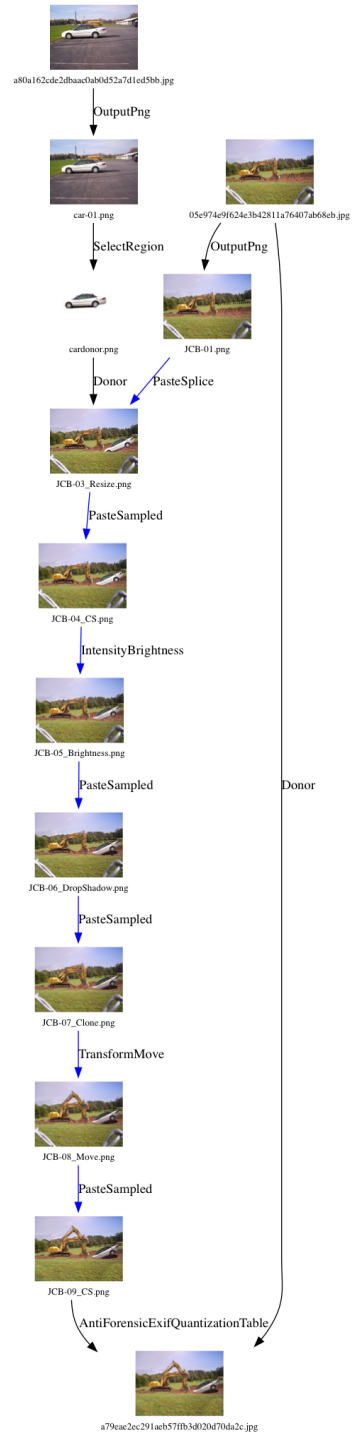

The evaluation series strives to support many aspects through detailed annotation and failure analysis

# Our Approach to Media Forensics Technology Development

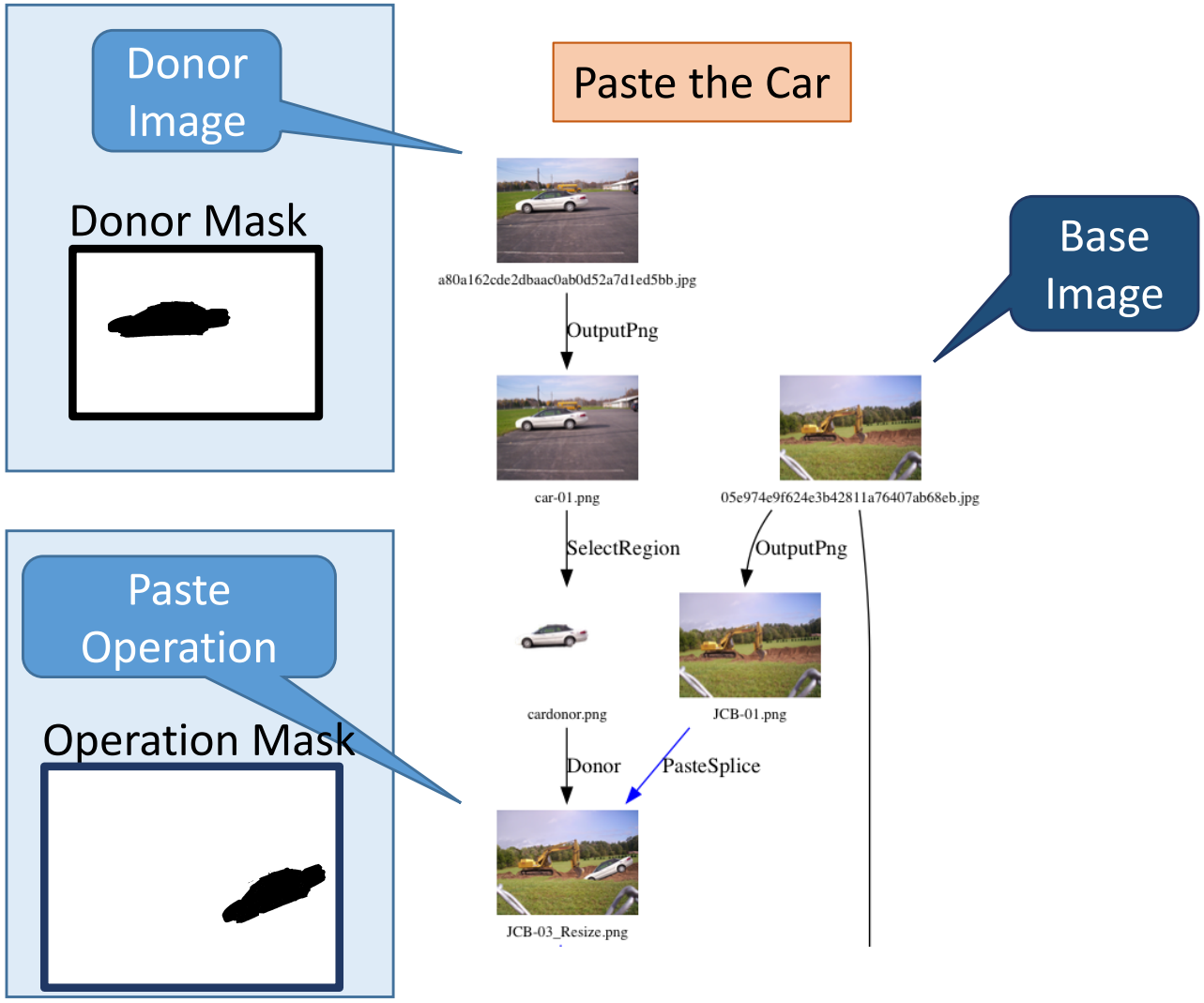
- Develop an expressive manipulation annotation record capable of supporting research and analysis
- Develop evaluation tasks and performance metrics that both explore component and end-to-end technologies
- Develop data sets to support research, development, and evaluations
- Administer a multi-year evaluation series to support long-term research

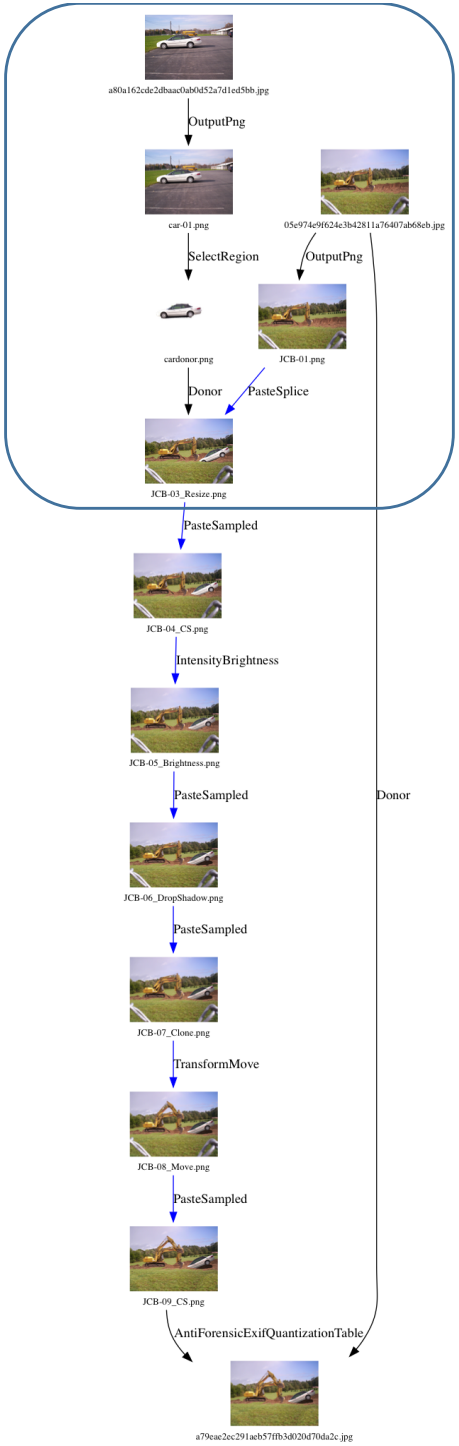
# Manipulation Annotation:

- Manipulation operations are recorded in graphs
  - Graph formalism defined for the Nimble ‘16 Data set
  - PAR Government Inc. extended the formalism creating manipulation ”journals”
- Masks collected for incoming and outgoing links

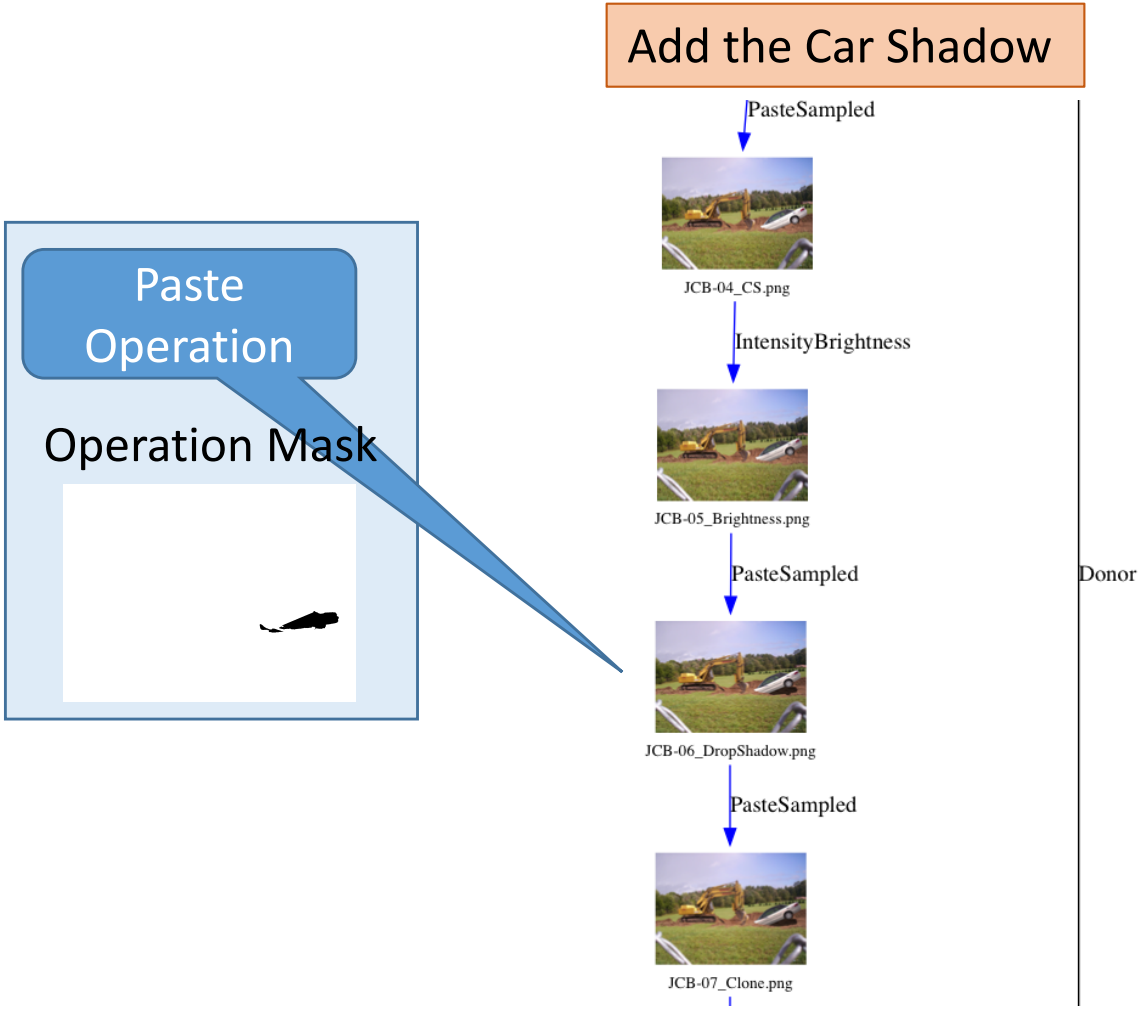


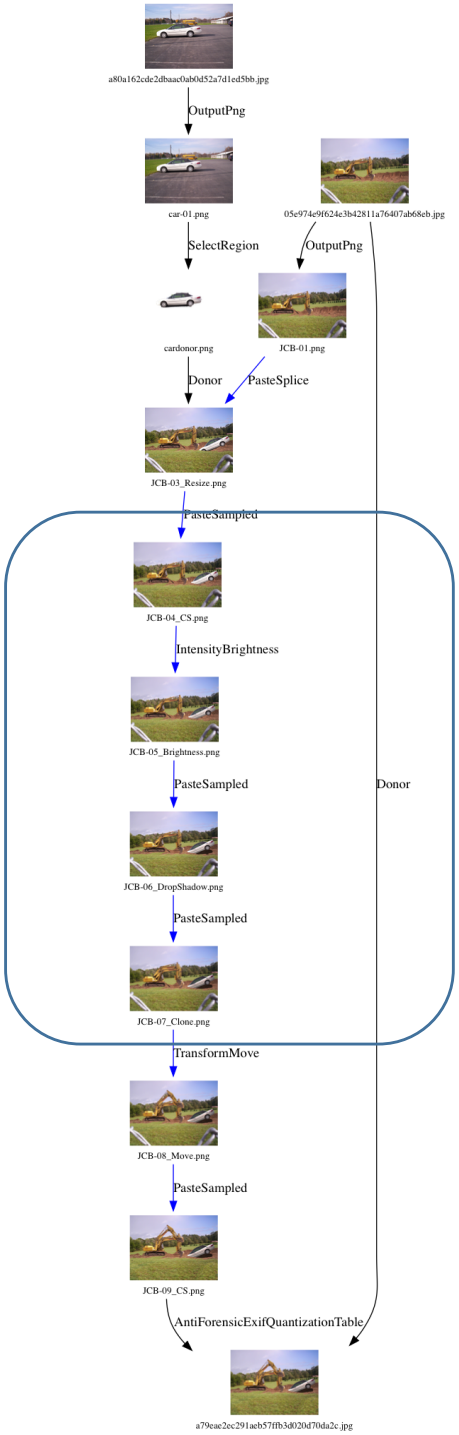
# Manipulation Annotation:



# Manipulation Annotation:



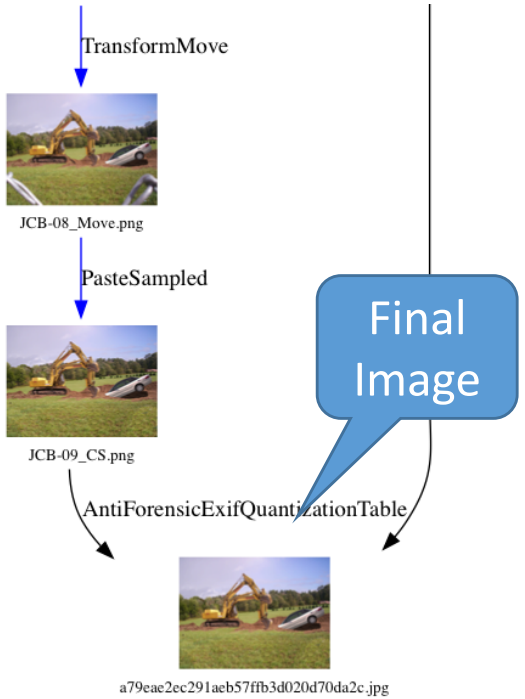



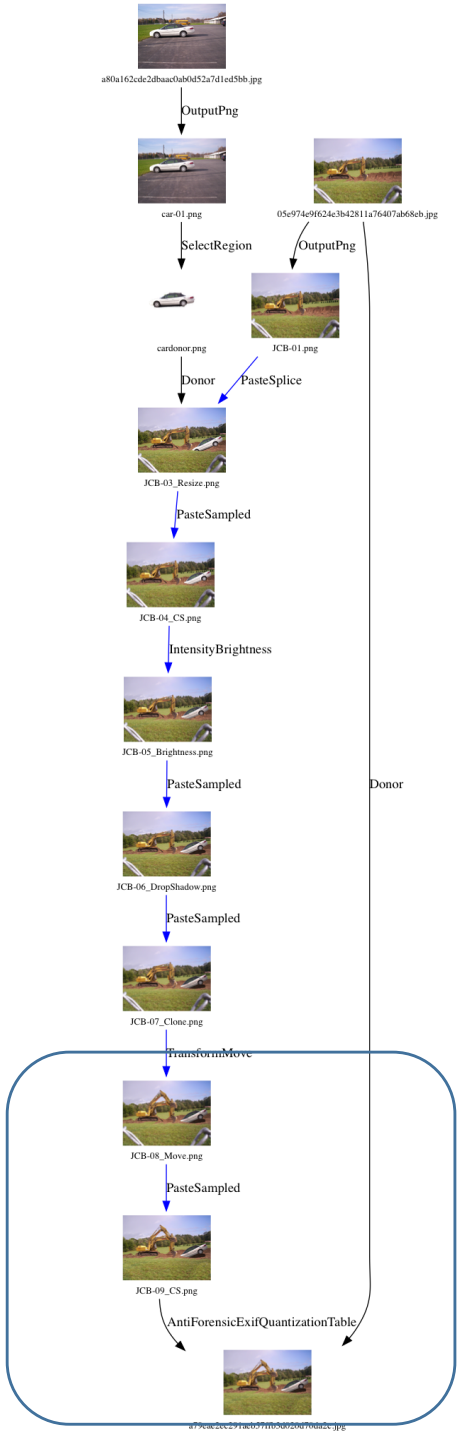
# Manipulation Annotation:

Copy the Backhoe Arm and Hide the Fence

Move Operation

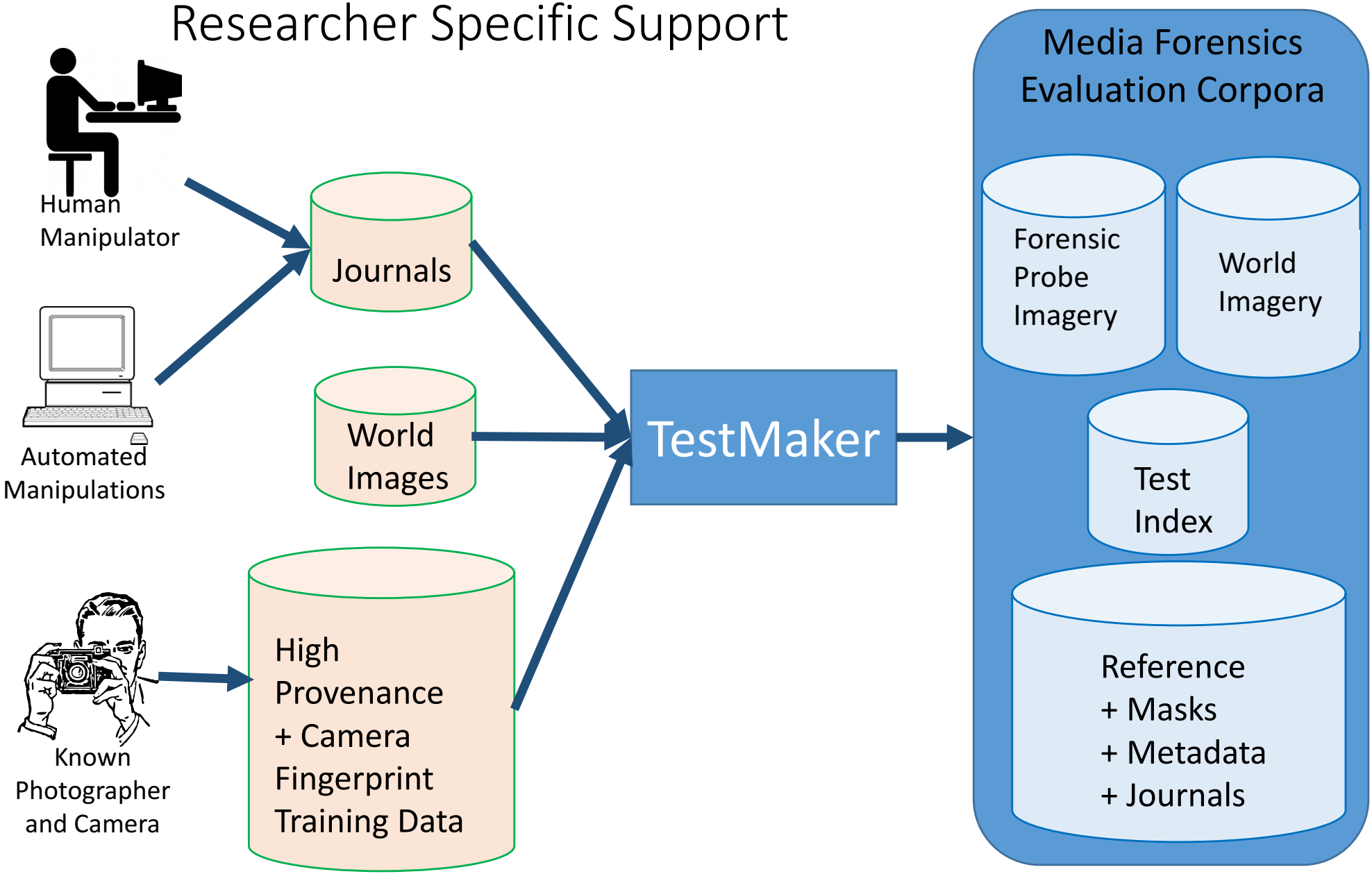
Operation Mask



# Data Set Production Data Flow:

## Researcher Specific Support



# Nimble Challenge 2017 Tasks and Definitions

- Manipulation Detection and Localization (MDL)
  - Image
  - Video (detection only)
- Splice Detection and Localization (SDL)
- Provenance Filtering (PF)
- Provenance Graph Building (PGB)

## Definitions:

**Probe**: the image or video being forensically analyzed

**Detection**: determine IF the probe was manipulated

**Localization**: determine WHERE the probe was manipulated

**Filtering**: Find imagery 'related' to the probe

**Graph Building:** construct the phylogeny graph of the probe

# Researcher Flexibility: System OptIn and Selective Scoring

Media Forensics techniques often address a specific manipulation type, sources, etc.

- System OptIn Protocol
  - The “OptIn” Protocol allows developer/system to:
    - Determine if a response is appropriate given ‘only the imagery and imagery metadata’
    - Communicate which probes were not processed and why
  - Score reporting
    - Trial Response Rate – Fraction of probes for which the system responded
    - Performance measures on the subset of trials.
- Selective Scoring – two approaches
  - Developer declares the type of operation detected by the system
    - E.g., this is a local blur detection system
  - Performed by NIST as a data analysis technique using metadata to condition analysis, i.e., manipulations of a certain type, etc.

# 2017 Nimble Challenge Participates Overview

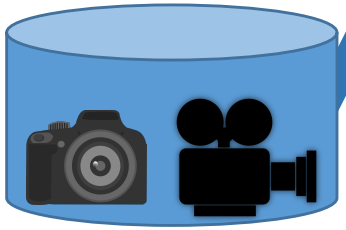
| Team Abb. | Organization ID | MDL (image/video) | SDL | PF | PG |
|---|---|---|---|---|---|
| BIN | Binghamton University | 1 | - | - | - |
| FIB | Honeywell ACS Laboratories | 1 | 1 | - | - |
| KIT | Kitware<br>UC Berkeley<br>Dartmouth College<br>University at Albany, SUNY | 4 + 1(video) | - | 1 | - |
| MAY | MAYACHITRA<br>Naval Air Warfare Center, China Lake<br>UC Riverside | 9 | - | - | - |
| PUR | Purdue<br>Politecnico di Milano, Italy<br>University of Siena<br>Univ. of Notre Dame; University of Campinas, Brazil | 5 | - | 5 | 4 |
| SRI-TA2 | SRI International, Princeton (Ajay Divakaran) | 1 | - | - | - |
| SRPPRI | SRI International, Princeton (Jeffrey Lubin) | 1+1(video) | - | - | - |
| UMD | University of Maryland, College Park | 1 | - | - | - |
| UNIFI | University of Florence, FENCE, Prato, Italy | 3 | 2 | - | - |
| USCISI | University of Southern California, ISI | 5 | 1 | 1 | 1 |
| 10 teams | 19 organizations, 49 systems | 31 + 2(video) | 4 | 7 | 5 |

# Manipulation Detection and Localization Evaluation Task

## System Input

Image(s) + (metadata)

Camera Fingerprint DB



**Algorithm**

Opt Out

Processed

## System Output

e.g., "No illuminated faces in the image for the system to analyze"

## Metrics

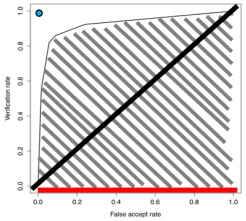
"Not Scored"

## Detection

Confidence score

27.58

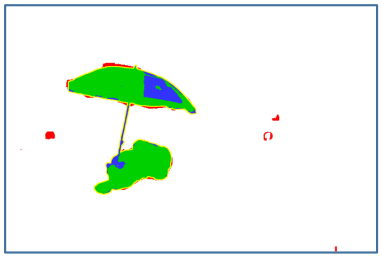
Receiver operating characteristic



Area Under the Curve: 0.85 (AUC)

## Localization

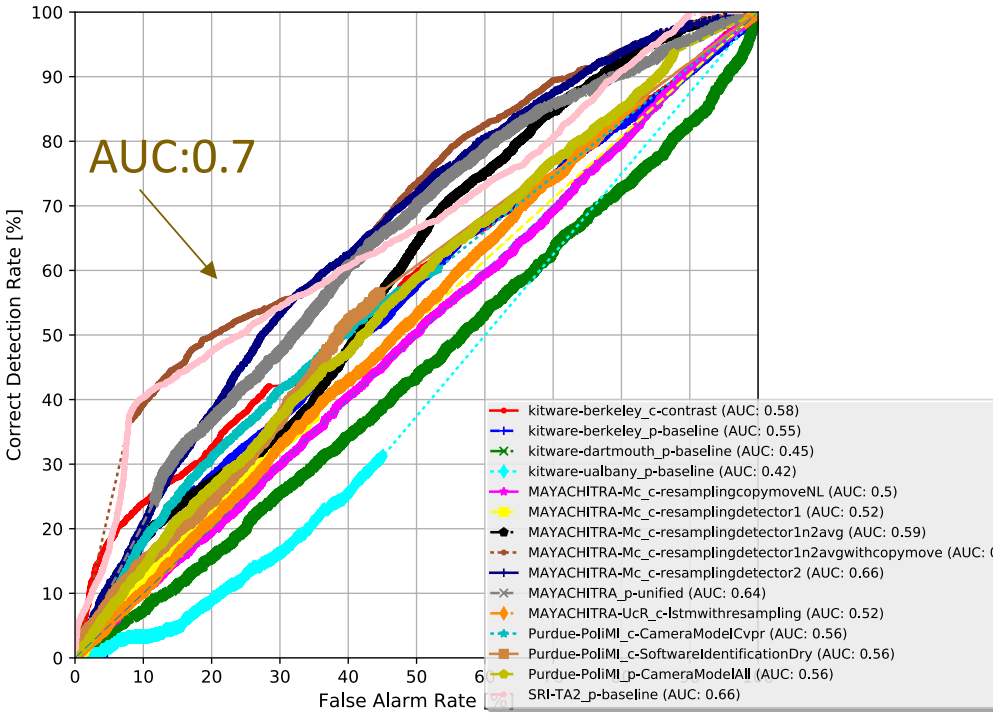
Probe Mask
(If a manipulation)



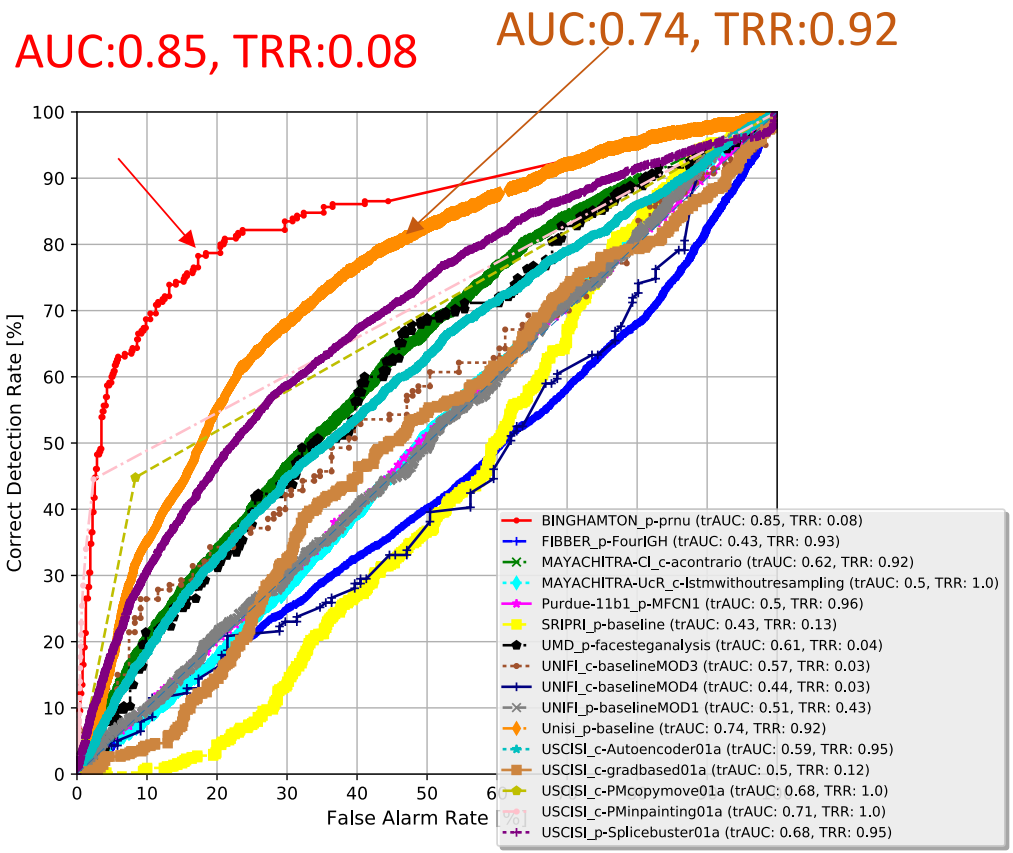
Maximum Matthews Correlation Coefficient
0.873343237591

# NC17 Image Manipulation Detection Results

## Systems That Processed All Probes

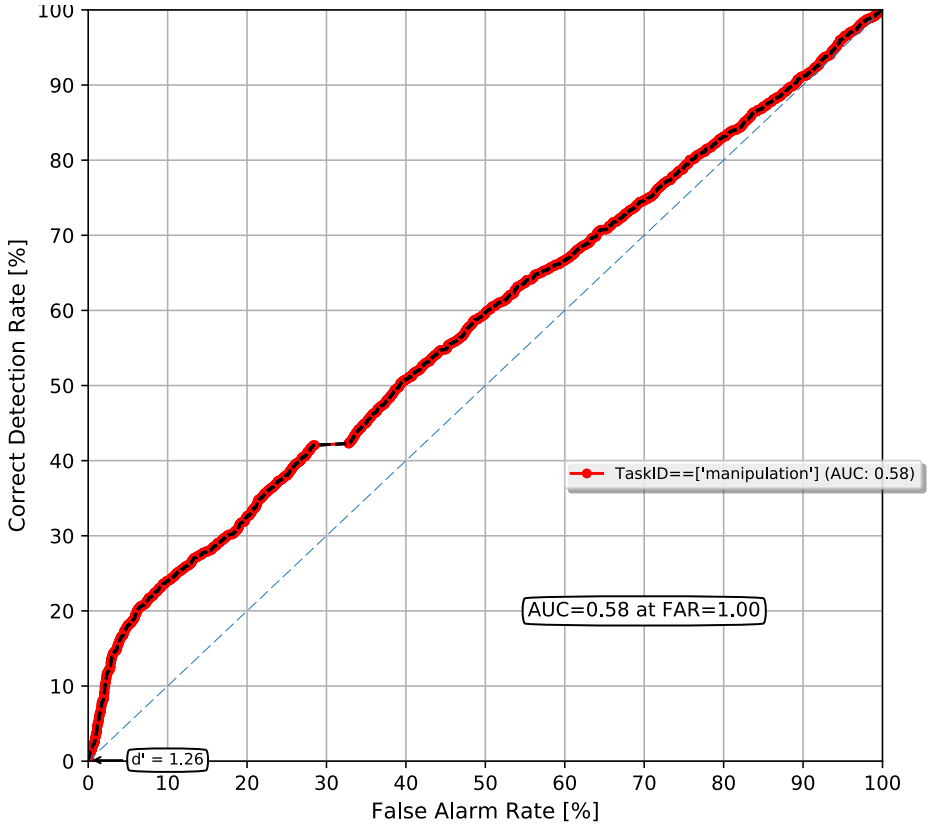



## Systems that "Opted In" to Process Some of the Probes

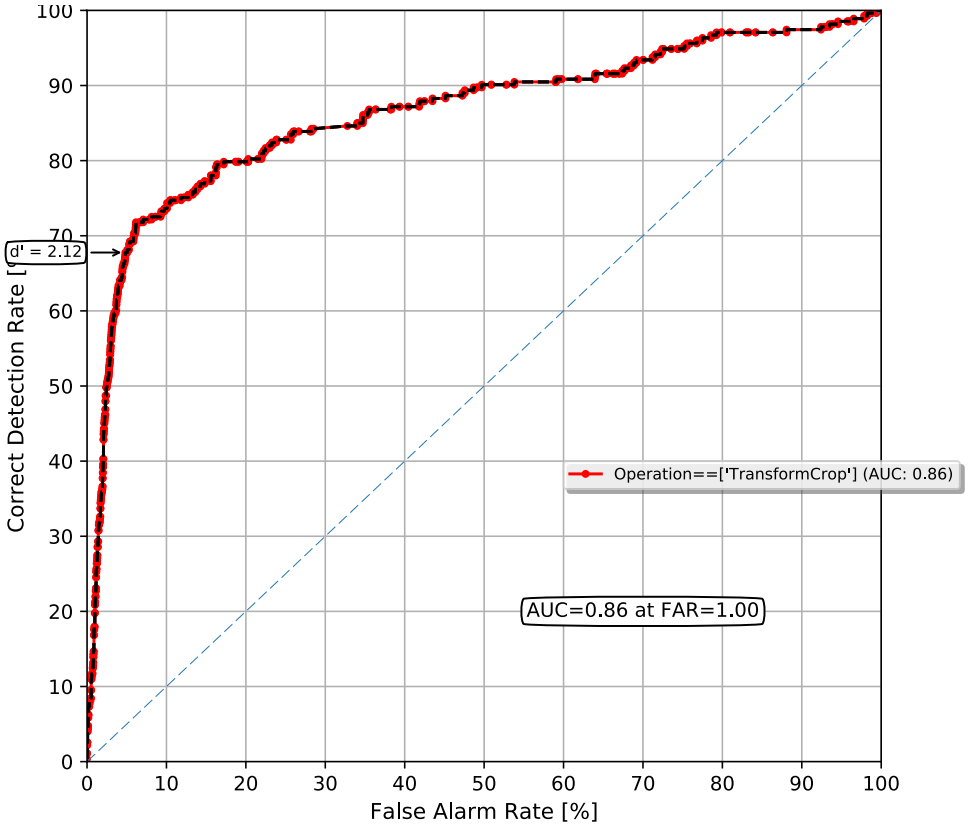


# Selective Scoring Results:
# All Operation vs. Crop-Only Probes

All Probes - AUC: 0.58




Crop-Only Probes - AUC: 0.86



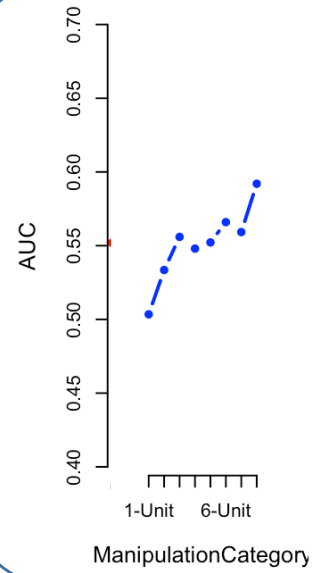
# Sensitivity Factor Analysis:

## Manipulation Detection Performance of Primary All Manipulation Systems

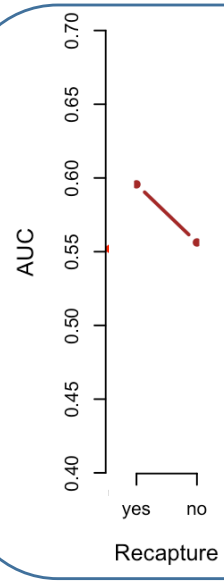
- Journals contain a wealth of information about the manipulations
- Selective Scoring provides the mechanism to study the effect of operations, metadata, etc. on performance
- This study measured the effect (the range of AUC performance across teams) for 25 metadata factors

### Top 17 of 25 Factors Sorted by Effect

| Factors | Effect |
|---|---|
| **ManipulationCategory** | **0.089** |
| Operation | 0.077 |
| BrowserUnit | 0.050 |
| OperationArgument | 0.043 |
| **Recapture** | **0.039** |
| SeamCarving | 0.026 |
| ImageCompressionTable | 0.026 |
| Natural Scene | 0.019 |
| SemanticRepurposing | 0.018 |
| CompositePixelSize | 0.016 |
| AntiforensicApplied | 0.016 |
| JournalSource | 0.014 |
| AntiforensicNoiseRestoration | 0.014 |
| AntiforensicAddCamFingerprintPRNU | 0.011 |
| Purpose | 0.010 |
| People | 0.007 |
| SemanticRestaging | 0.003 |



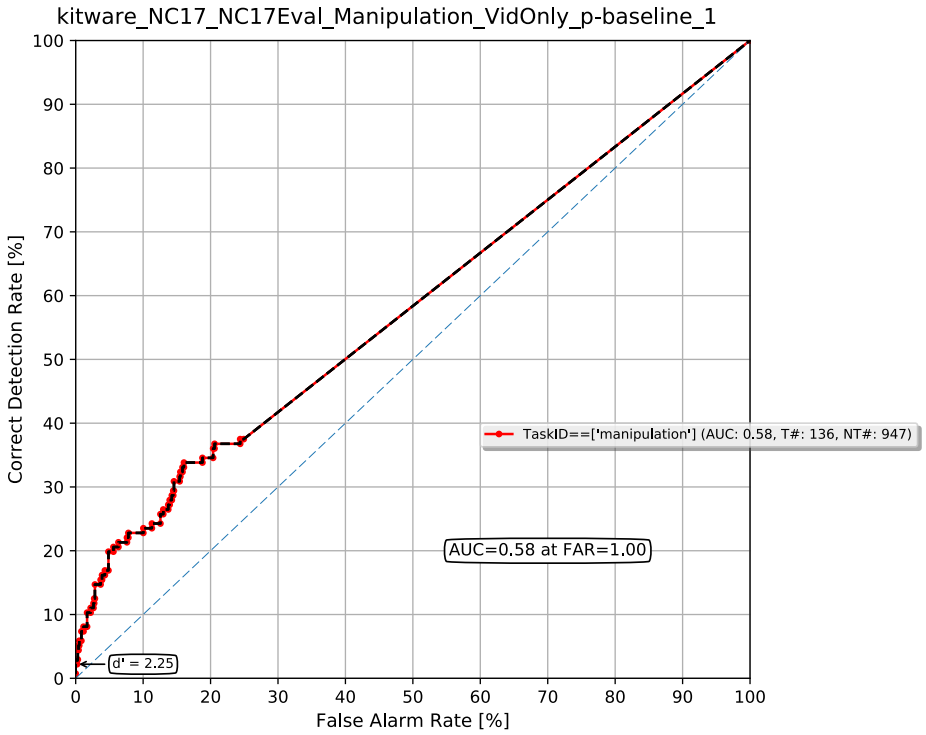

**Manipulation Category**
The number of manipulations applied to the image that changed the image in some manner



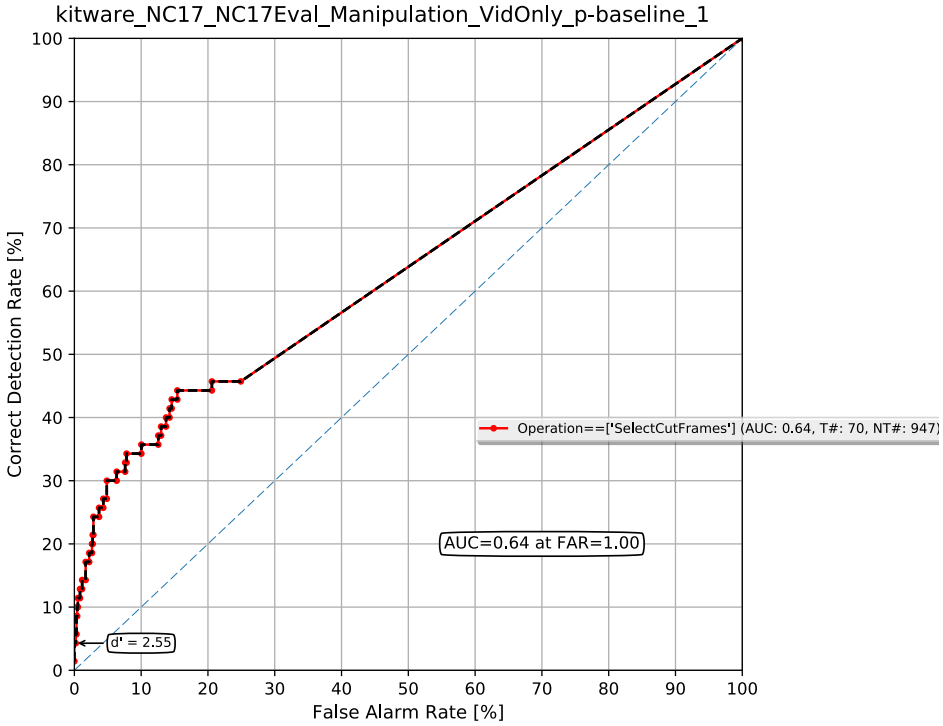

**Recapture**
The image was recaptured (i.e., a scan of a print of the image)

# Video Manipulation Detection:
## Kitware -- All Manipulations vs. Drop Frame Probes

All operations; AUC = 0.58




Selective Scoring
Drop Frame; AUC = 0.64



# NC17 Image Manipulation Localization Results (11 teams, 16 systems)

| Team | System | All-MMCC | TR-MMCC | Trial Response Rate |
|---|---|---|---|---|
| BINGHAMTON | p-prnu_1 | | 0.1853 | 0.1000 |
| FIBBER | p-FourIGH_1 | | 0.0365 | 0.9886 |
| MAYACHITRA-Cl | c-acontrario_3 | | 0.0345 | 0.9945 |
| MAYACHITRA-Mc | c-resamplingdetector1_3 | 0.0202 | | |
| MAYACHITRA-UcR | c-lstmwithoutresampling_2 | | 0.0035 | 0.9975 |
| Purdue-11b1 | p-MFCN1_1 | | 0.0596 | 0.9980 |
| SRI-TA2 | p-baseline_1 | 0.0887 | | |
| SRIPRI | p-baseline_1 | | 0.0831 | 0.1870 |
| UMD | p-facesteganalysis_1 | | 0.1876 | 0.1054 |
| UNIFI | c-baselineMOD3_1 | | 0.2241 | 0.0686 |
| | c-baselineMOD4_1 | | 0.2237 | 0.0681 |
| USCISI | c-Autoencoder01a_1 | | 0.1893 | 0.9727 |
| | c-PMcopymove01a_1 | | 0.1317 | 0.9995 |
| | c-PMinpainting01a_1 | | 0.1209 | 0.9995 |
| | c-gradbased01a_1 | | 0.1957 | 0.2337 |
| | p-Splicebuster01a_1 | | 0.1991 | 0.9727 |

Matthews Correlation Coefficient:

(MCC)

$$\frac{TP * TN - FP * FN}{\sqrt{(TP + FP)(TP + FN)(TN + FP)(TN + FN)}}$$

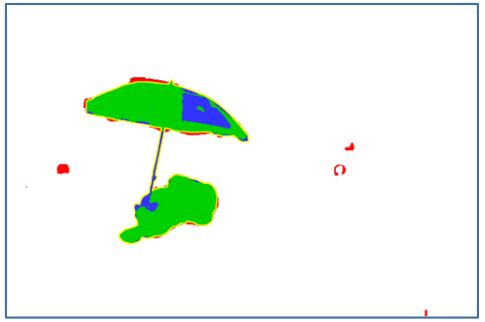
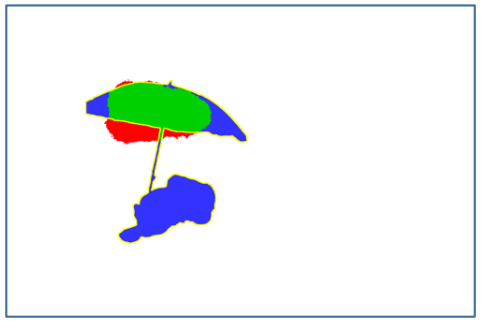
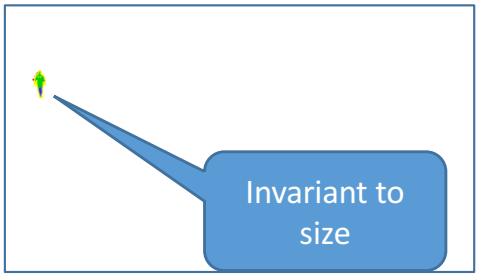
- MCC=1 → perfect correlation
- MCC=0 → no correlation or no output (by convention)
- MCC=-1 → perfect anti-correlation

Maximum MCC=

$$argmax_{\Theta}\ (MCC(\Theta))$$

- All-MMCC – Maximum MCC average over all true manipulations
- TR-MMCC – Maximum MCC average over Opted In true manipulations

# NC17 Image Manipulation Localization - All operation example (1)

| | | | |
|---|---|---|---|
| | • Black – Manipulation<br>• Yellow - No-Score | • Green - True Positives<br>• Red - False Alarm.<br>• White - True Negative<br>• Blue - False Negative | |
| Composite | Binarized Reference | SystemID1 | SystemID2 |
| 21a1b6501b9c0d84fa46ad6eddf8bbe4 | | MMCC: 0.87 | MMCC: 0.57 |
|  |  |  |  |
| fb8785800546e9602ef35c7ee0cee8b7 | | MMCC: 0.84 | MMCC: 0 |
|  |  |  |  |

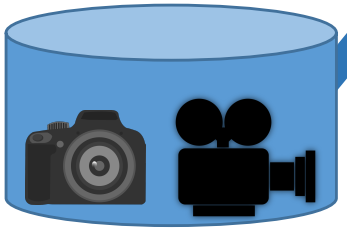
# Splice Detection and Localization Evaluation Task

## System Input

Image(s) + (metadata)


Probe Image


Donor image



Camera Fingerprint DB

Opt Out

**Algorithm**

Processed

## System Output

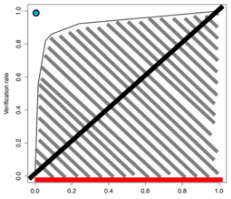
e.g., "No illuminated faces in the image for the system to analyze"

## Metrics

"Not Scored"

## Detection

Confidence score
97.86

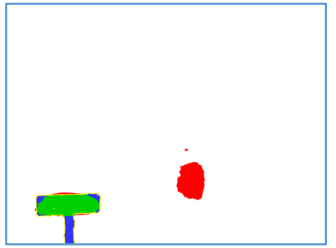


ROC

AUC: 0.85

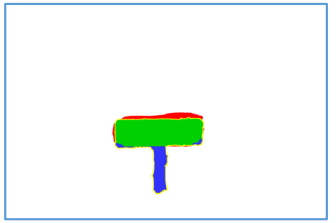
## Localization

System probe mask



Probe Max MCC: 0.636

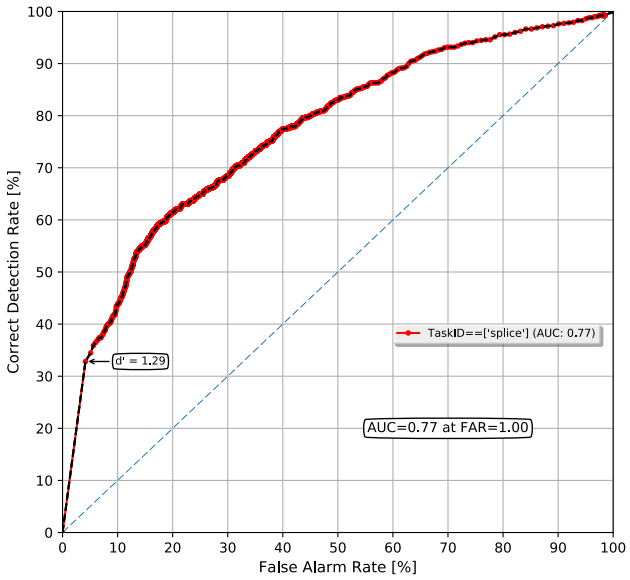
System donor mask



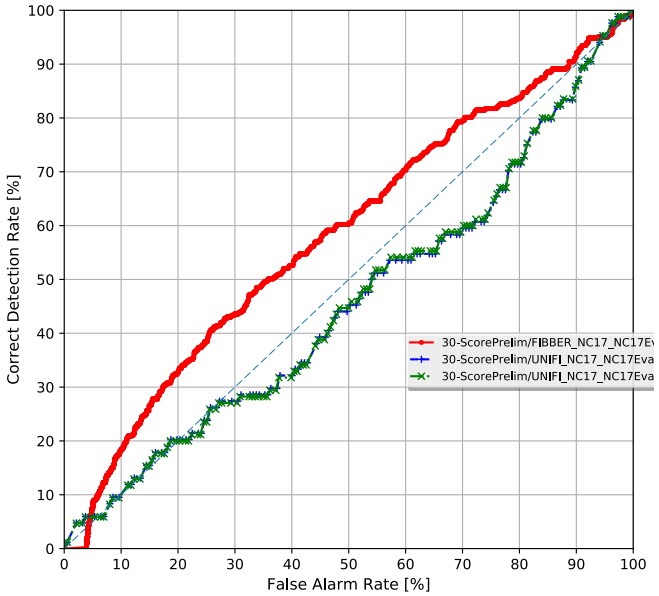
Donor Max MCC: 0.816

# NC17 Splice Manipulation Results
## Detection ROC and Localization MaxMCC

Systems That Processed All Probes



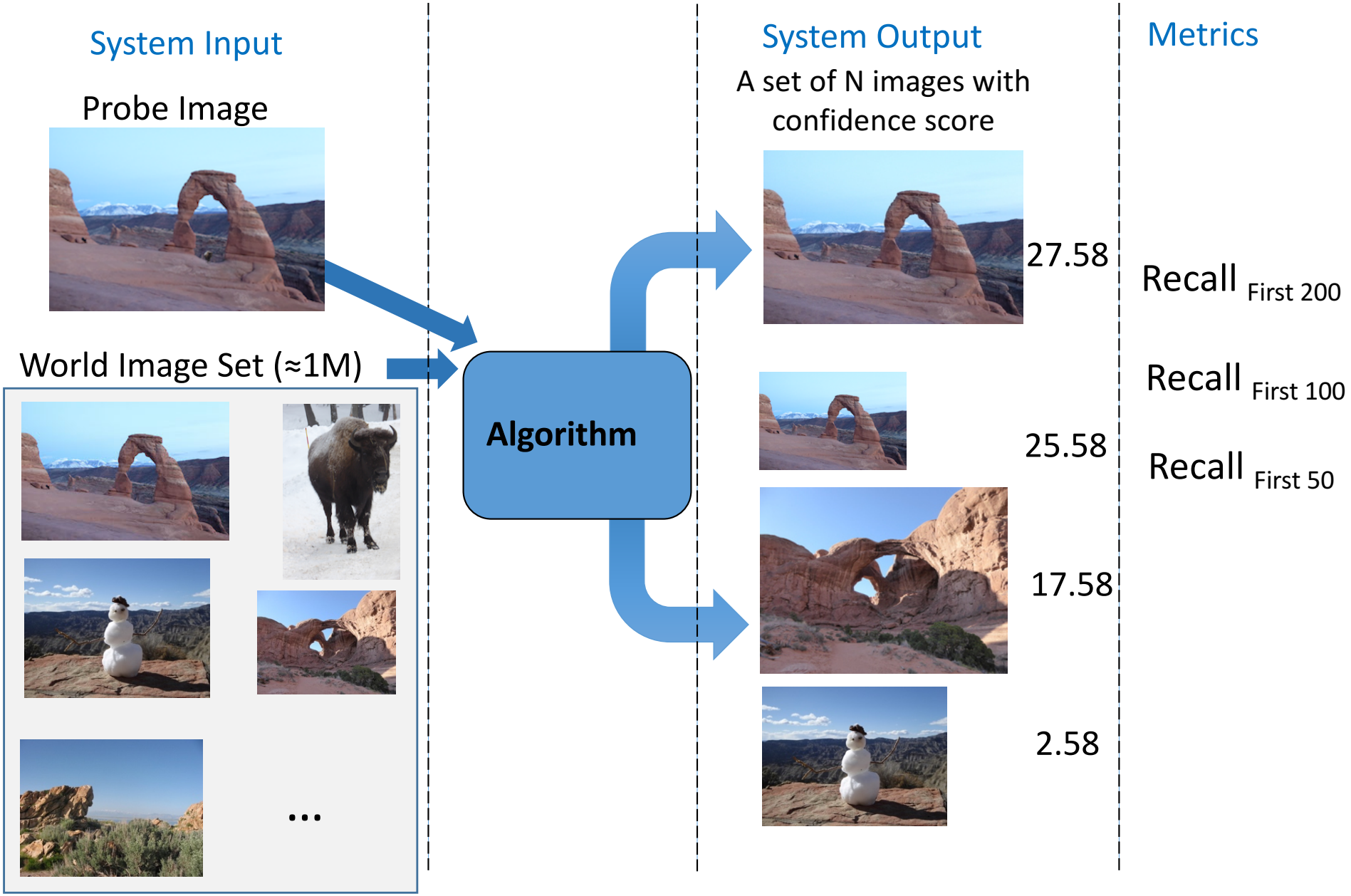

Systems that "Opted In" to Process Some of the Probes



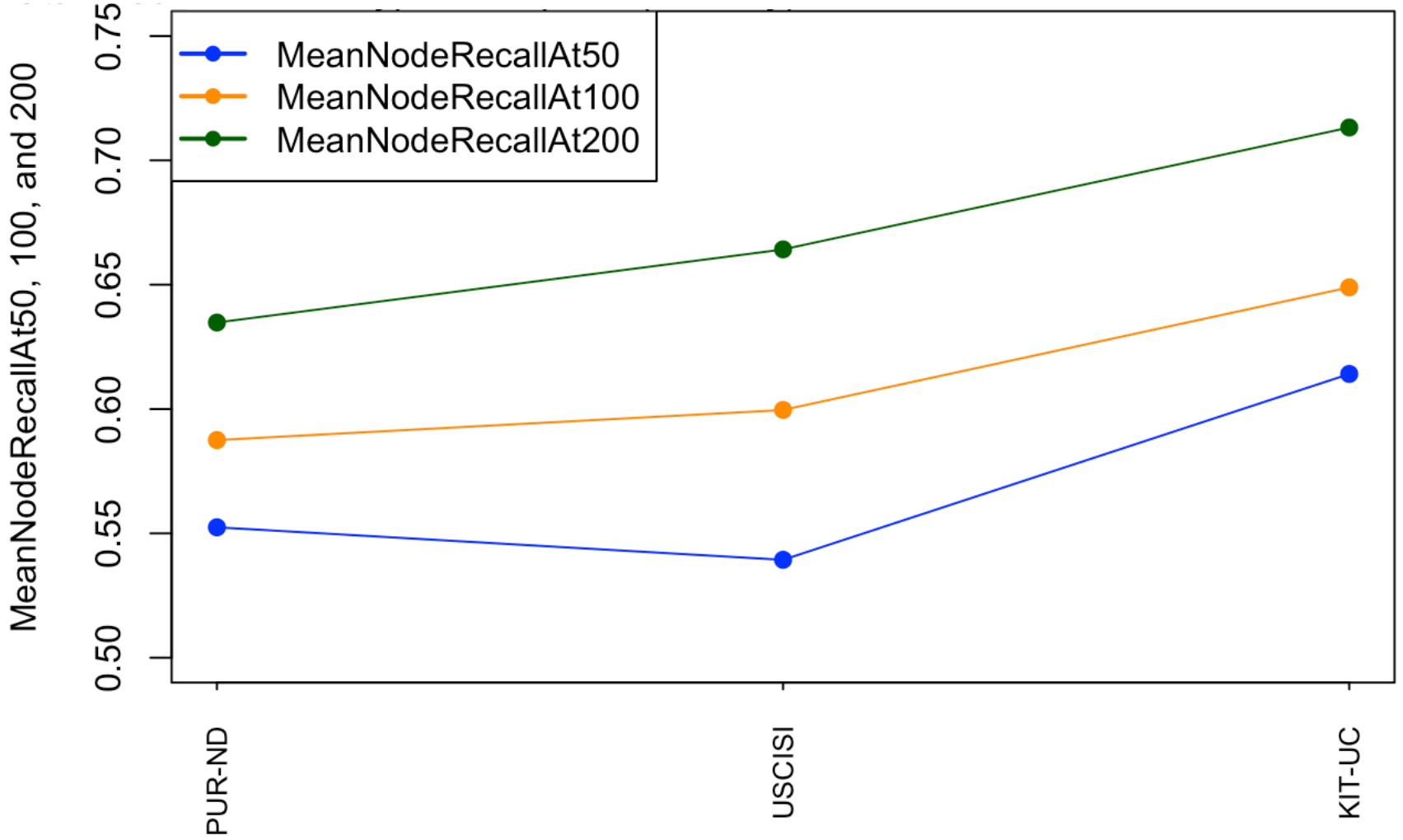

| | | Donor TRR | Donor trMMCC | Probe TRR | Probe trMMCC |
|---|---|---|---|---|---|
| UNIFI | c-baselineMOD4_1 | 0.0907 | 0.1010 | 0.0910 | 0.1940 |
| | p-baselineMOD3_1 | 0.0918 | 0.0998 | 0.0921 | 0.1916 |
| USCISI | p-baseline_1 | 1.000 | 0.1862 | 1.0000 | 0.1740 |

OptIn

# Provenance Filtering Evaluation Task



# Recall Metric Comparison by Depth of Retrieval

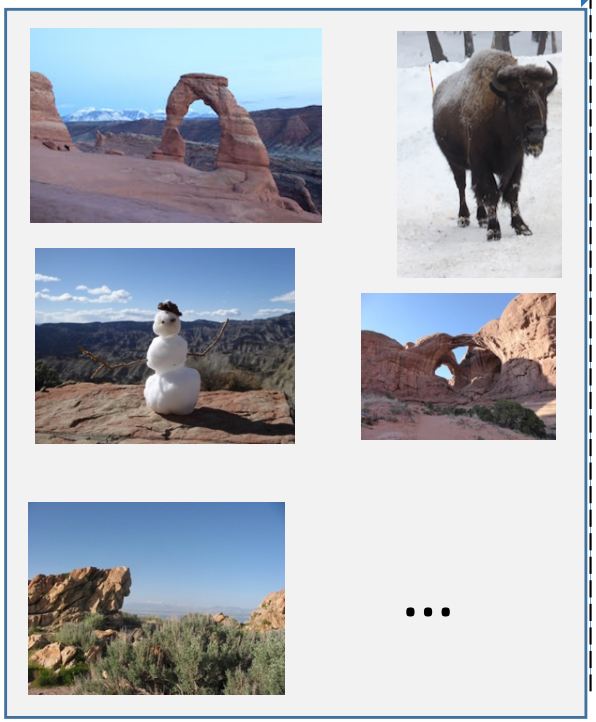


# Provenance Graph Building Evaluation Task

## System Input

Probe Image

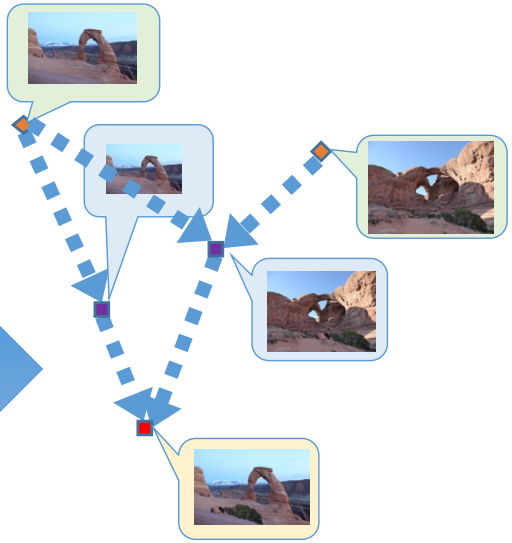
World Image Set (≈1M)



Algorithm

## System Output

A provenance graph



## Metrics

Graph Similarity

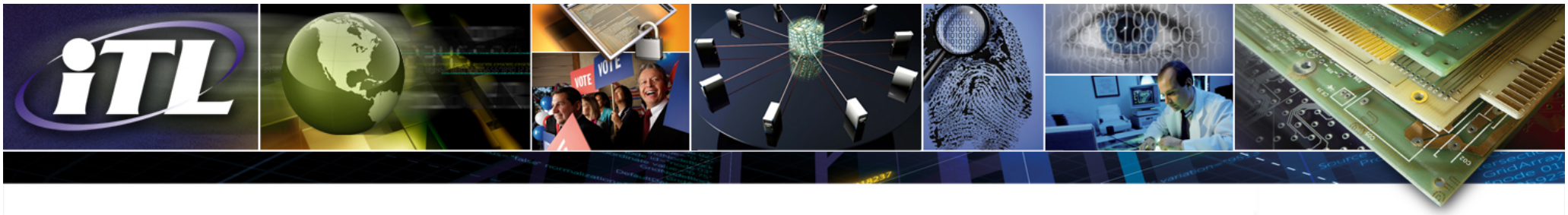
Generalized F-measure:

- Sim(nodes)
- Sim (links)
- Sim(nodes+links)

# Provenance Graph Building Task Evaluation Metrics

- Graph Similarity and Generalized F-measure
  - Overlap of nodes: $\mathrm{sim}_{\mathrm{NO}}(G_r, G_s) = 2\frac{|V_r \cap V_s|}{|V_r|+|V_s|}$
  - Overlap of links: $\mathrm{sim}_{\mathrm{LO}}(G_r, G_s) = 2\frac{|E_r \cap E_s|}{|E_r|+|E_s|}$
  - Overlap of node and links: $\mathrm{sim}_{\mathrm{NLO}}(G_r, G_s) = 2\frac{|V_r \cap V_s|+|E_r \cap E_s|}{|V_r|+|V_s|+|E_r|+|E_s|}$

| MeanNodeRecall | From Provenance Filtering |
|---|---|
| MeanSimNO | Similarity of Node Overlap for a Provenance Graph - Eval Plan Section 7.0 |
| MeanSimLO | Similarity of Link Overlap for a Provenance Graph - Eval Plan Section 7.0 |
| MeanSimNLO | Similarity of Link+Node Overlap for a Provenance Graph - Eval Plan Section 7.0 |



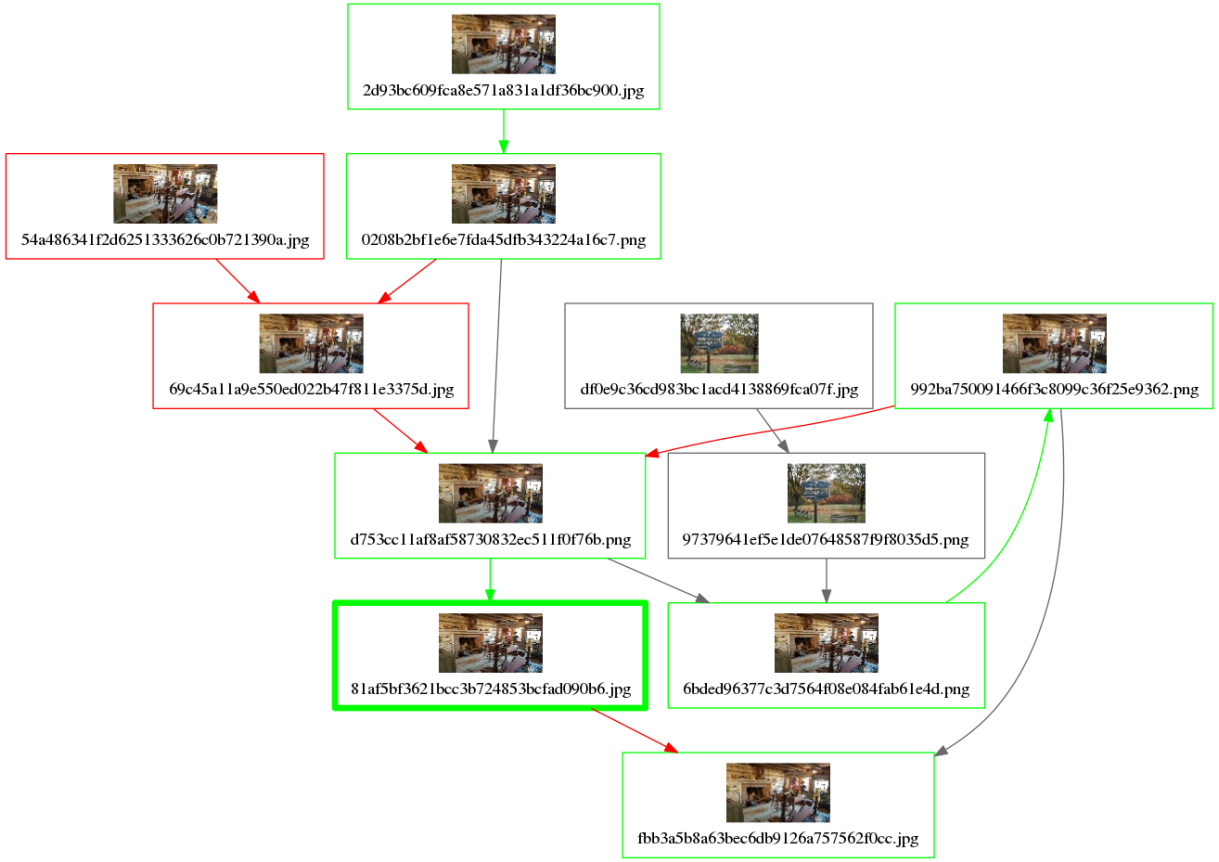

# NC2017 Provenance Graph Building Eval. Results

- 2 teams/organizations, 5 systems (end-to-end)

| Team/System | | Mean Node Recall | Mean Similarity | | |
|---|---|---|---|---|---|
| | | | Node Overlap | Link Overlap | Node and Link Overlap |
| NDPURDUE | c-contrast1_1 | 0.5249 | 0.5913 | 0.1812 | 0.3875 |
| | c-contrast2_1 | 0.5228 | 0.6124 | 0.2189 | 0.4170 |
| | c-contrast3_1 | 0.5246 | 0.5909 | 0.1809 | 0.3872 |
| | p-baseline_1 | 0.5230 | 0.6127 | 0.2085 | 0.4124 |
| USCISI | p-baseline_1 | 0.4786 | 0.4146 | 0.0776 | 0.2674 |

# Provenance Graph Evaluation Example: ND-Purdue, Baseline System

| | Mean Similarity | | |
|---|---|---|---|
| Mean Node Recall | Node Overlap | Link Overlap | Node and Link Overlap |
| 0.778 | 0.778 | 0.375 | 0.588 |




## Image Legend

- **Wide Green** image border - The Probe image.
- Green image border - Correctly included image.
- Red image border - False alarm image.
- Grey image border - Omitted provenance image (missed detection).

## Link Legend

- Green link - Correctly linked images.
- Red link - False alarm link.
- Grey link - Omitted link.

# Translating the Use Case Into Research Tasks

'17

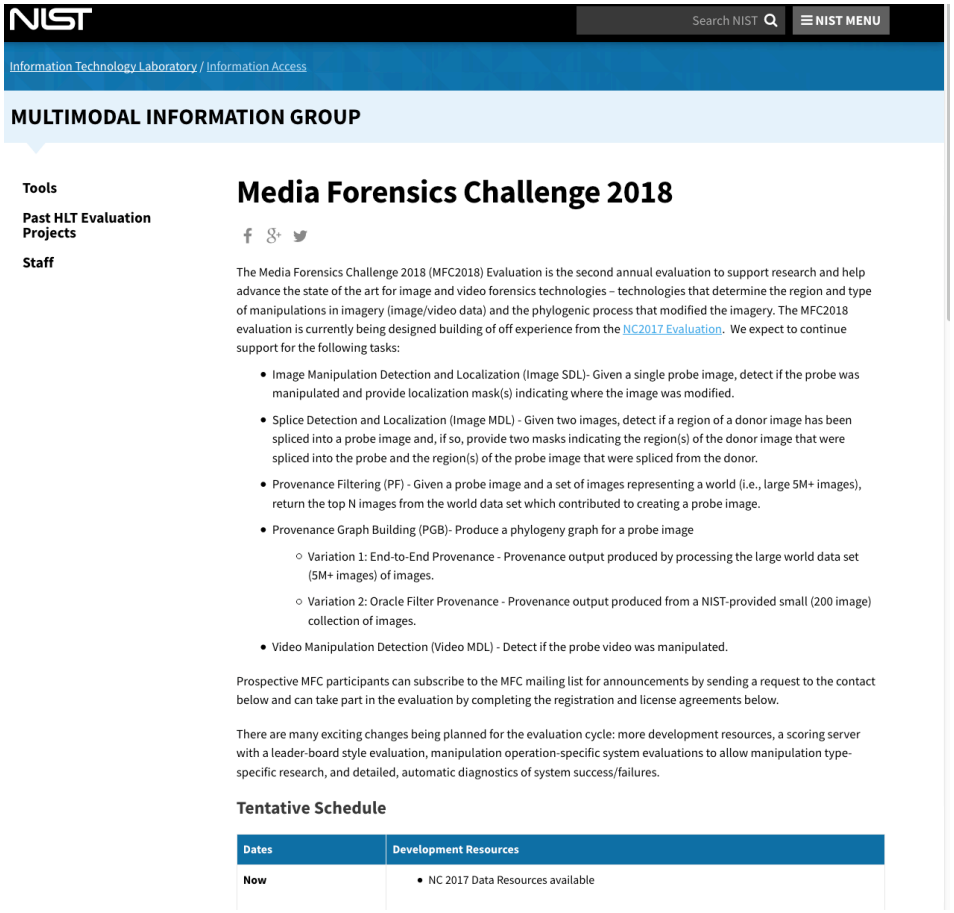
- Is the image manipulated?
- Where do the manipulations spatially occur?
- What operations were performed?
- Is there an original image?
- Are there related images?
- Is the image consistent with the camera?
- Are there known examples of vehicles?
- Is the image consistent with the reported date and location?

# Media Forensics Challenge ’18: Sign Me Up!

https://www.nist.gov/itl/iad/mig/media-forensics-challenge-2018

- Step 1: Complete agreements
- Step 2: Get data
- Step 3: Get evaluation tools
- Step 4: Build a system
- Step 5: Participate in the MFC ‘18 evaluation

- Step 6: Keep researching for MFC ‘19



NIST Search NIST NIST MENU

Information Technology Laboratory / Information Access

MULTIMODAL INFORMATION GROUP

Tools

Past HLT Evaluation Projects

Staff

## Media Forensics Challenge 2018

The Media Forensics Challenge 2018 (MFC2018) Evaluation is the second annual evaluation to support research and help advance the state of the art for image and video forensics technologies – technologies that determine the region and type of manipulations in imagery (image/video data) and the phylogenic process that modified the imagery. The MFC2018 evaluation is currently being designed building of off experience from the NC2017 Evaluation. We expect to continue support for the following tasks:

- Image Manipulation Detection and Localization (Image SDL)- Given a single probe image, detect if the probe was manipulated and provide localization mask(s) indicating where the image was modified.
- Splice Detection and Localization (Image MDL) - Given two images, detect if a region of a donor image has been spliced into a probe image and, if so, provide two masks indicating the region(s) of the donor image that were spliced into the probe and the region(s) of the probe image that were spliced from the donor.
- Provenance Filtering (PF) - Given a probe image and a set of images representing a world (i.e., large 5M+ images), return the top N images from the world data set which contributed to creating a probe image.
- Provenance Graph Building (PGB)- Produce a phylogeny graph for a probe image
  - Variation 1: End-to-End Provenance - Provenance output produced by processing the large world data set (5M+ images) of images.
  - Variation 2: Oracle Filter Provenance - Provenance output produced from a NIST-provided small (200 image) collection of images.
- Video Manipulation Detection (Video MDL) - Detect if the probe video was manipulated.

Prospective MFC participants can subscribe to the MFC mailing list for announcements by sending a request to the contact below and can take part in the evaluation by completing the registration and license agreements below.

There are many exciting changes being planned for the evaluation cycle: more development resources, a scoring server with a leader-board style evaluation, manipulation operation-specific system evaluations to allow manipulation type-specific research, and detailed, automatic diagnostics of system success/failures.

### Tentative Schedule

| Dates | Development Resources |
|---|---|
| Now | • NC 2017 Data Resources available |

# List of Data Sets Available to Participants

| Data Set Type | Data Set Name | Number of Forensic Probes | World Data Set Size | Data Size | Reference Annotations | Supported Tasks |
|---|---|---|---|---|---|---|
| Development | NC2016 – Both Nimble Science and Nimble Web | 624 | N/A | 4GB | Full | MDL |
| | NC'17 Development Image Data | 3,500 | 100,000 | 379 GB | Full | MDL, VMD, SDL, Prov |
| | NC'17 Development Video Data | 213 | | | | |
| | NC'18 Development Image and Video Data | TBD | TBD | TBD | Full | TBD |
| Past Evaluations | NC'17 Evaluation Images | 10,000 | 1,000,000 | 3.5TB | Full for 1/3 subset | MDL, SDL, Prov |
| | NC'17 Evaluation Videos | 1,000 | | 117GB | Full for 1/3 Subset | VMD |
| NC '18 Evaluation | NC'18 Evaluation Images | 50,0000 | 5,000,000 | ~ | | MDL, SDL, Prov |
| | NC'18 Evaluation Videos | 5,000 | | ~600GB | | VMD |

MDL: Manipulation Detection and Localization
SDL: Splice Detection and Localization
Prov: Provenance Filtering and Graph Building
VMD: Video Manipulation Detection

# MFC 2018 Changes

- Evaluation Task Changes
  - Provenance graphs with link operations
- New data resources
  - 2 additional development releases: Sept 30, Dec 31
  - Bigger evaluation collection
- Metric changes
  - Localization – Object/operation/sub-unit/region level scoring
  - Detection metrics focused on low false alarm
- Scoring Server
  - Leaderboard and blind evaluations
  - Developer-controlled selective scoring
  - Statistical system comparisons
- Semantic Integrity
  - Dave Doermann will present this later today

# Thank You for Your Attention!

NIST MediFor Team: medifor-nist@nist.gov

MFC '18 Web Site: https://www.nist.gov/itl/iad/mig/media-forensics-challenge-2018

**Disclaimer**

Any mention of commercial products or reference to commercial organizations in this report is for information only; it does not imply recommendation or endorsement by NIST nor does it imply that the products mentioned are necessarily the best available for the purpose.