usenix THE ADVANCED COMPUTING SYSTEMS ASSOCIATION
34TH USENIX SECURITY SYMPOSIUM

Scoop: Mitigation of Recapture Attacks on Provenance-Based Media Authentication

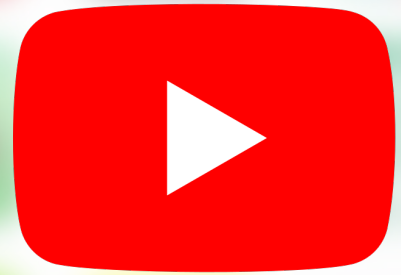
Yuxin (Myles) Liu,
Habiba Farrukh, Ardalan Amiri Sani, Sharad Agarwal, Gene Tsudik

TrussLab

UCI University of California, Irvine

Microsoft

In the digital world, information spreads fast



So does fake and fraudulent information

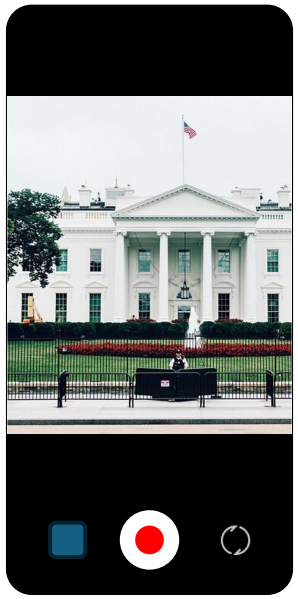
- Great advances in AI facilitate equally impressive advances in deep fakes

So does fake and fraudulent information

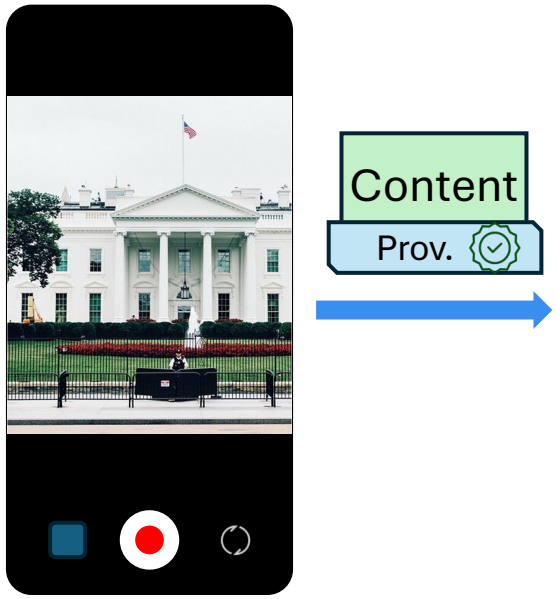
- Great advances in AI facilitate equally impressive advances in deep fakes

Detection-based approaches are an endless race that's unlikely to win

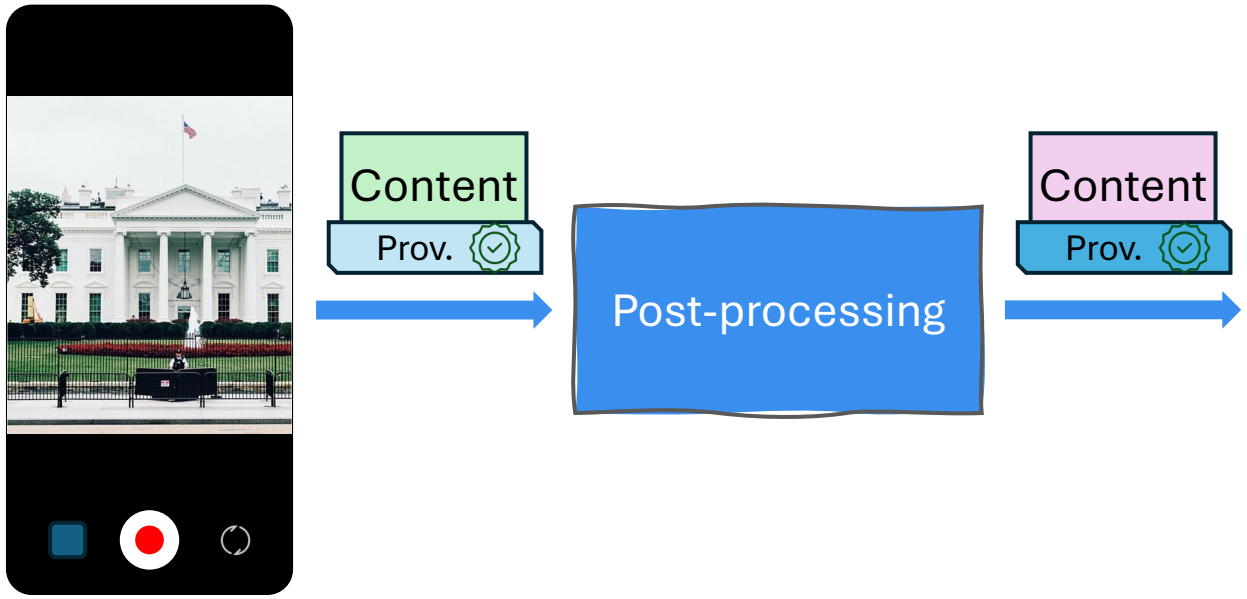
Provenance-based media authentication



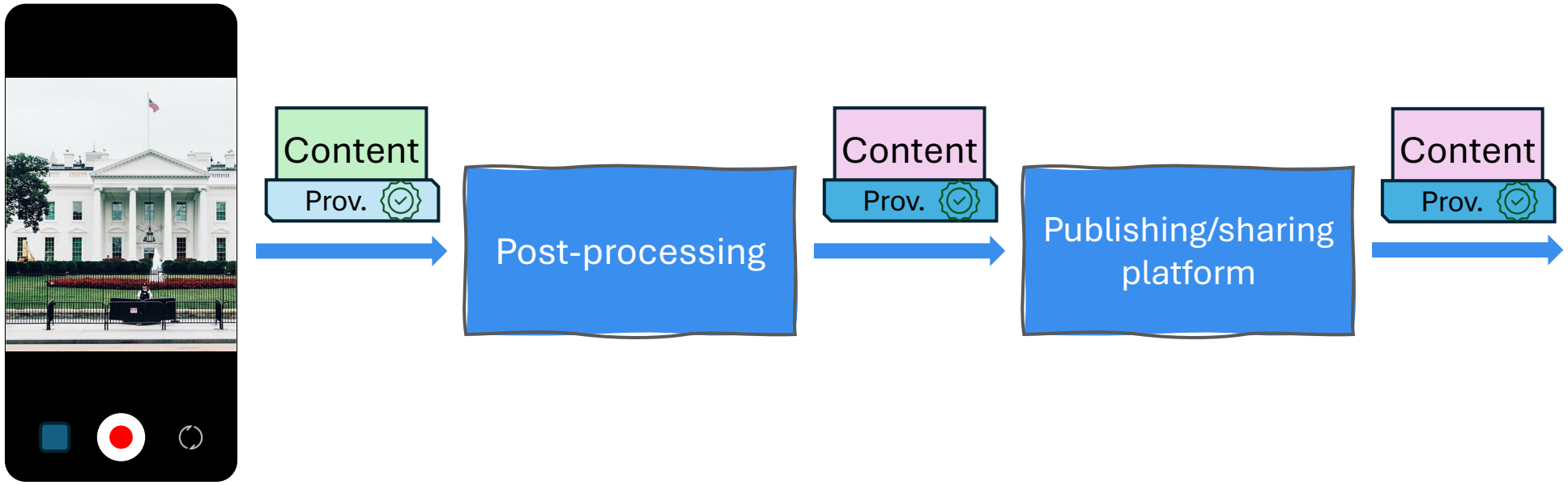
Provenance-based media authentication



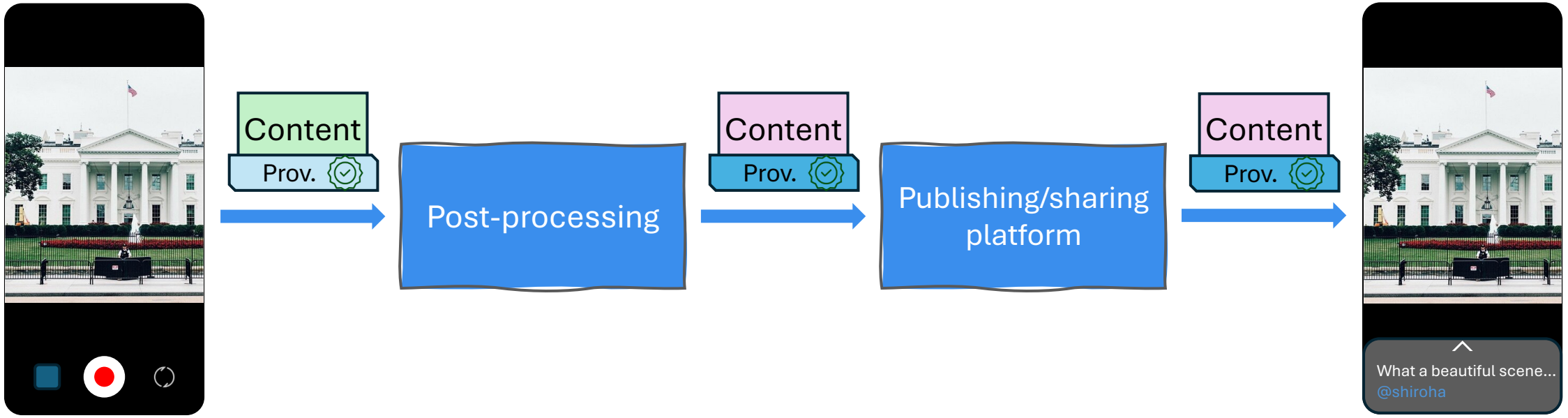
Provenance-based media authentication



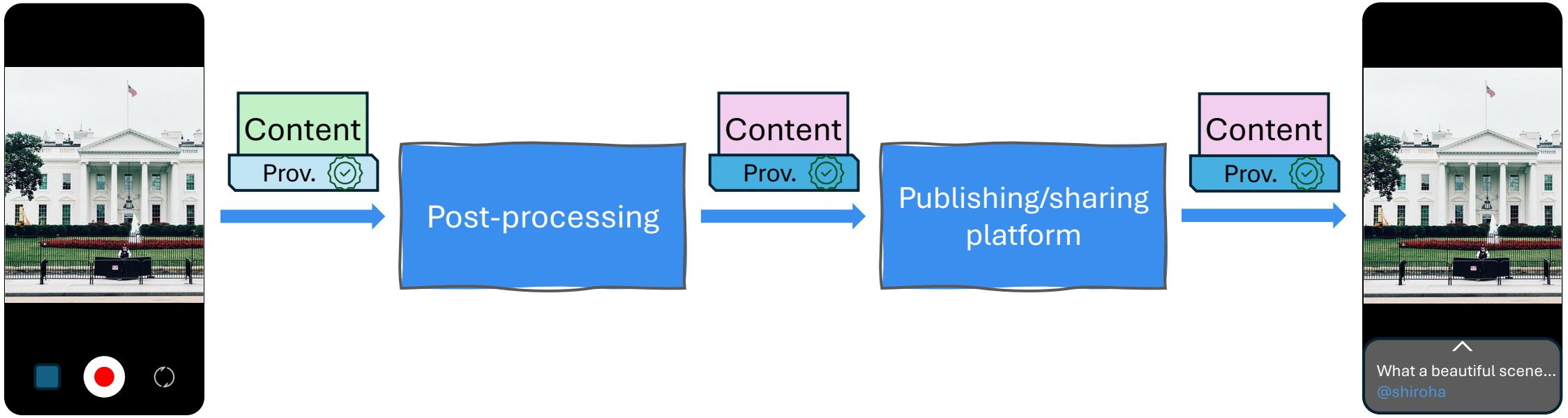
Provenance-based media authentication



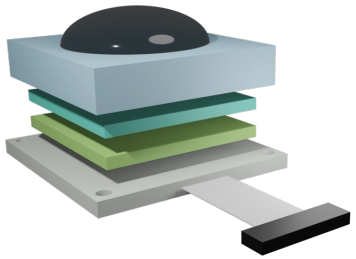
Provenance-based media authentication



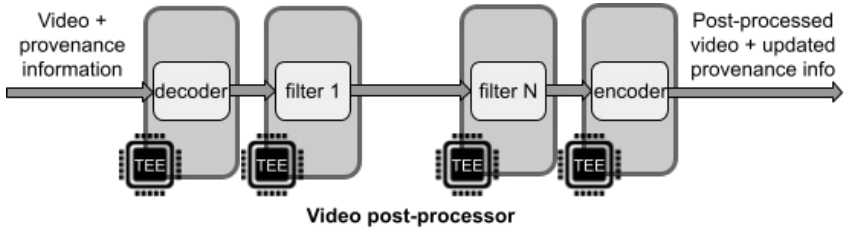
Provenance-based media authentication



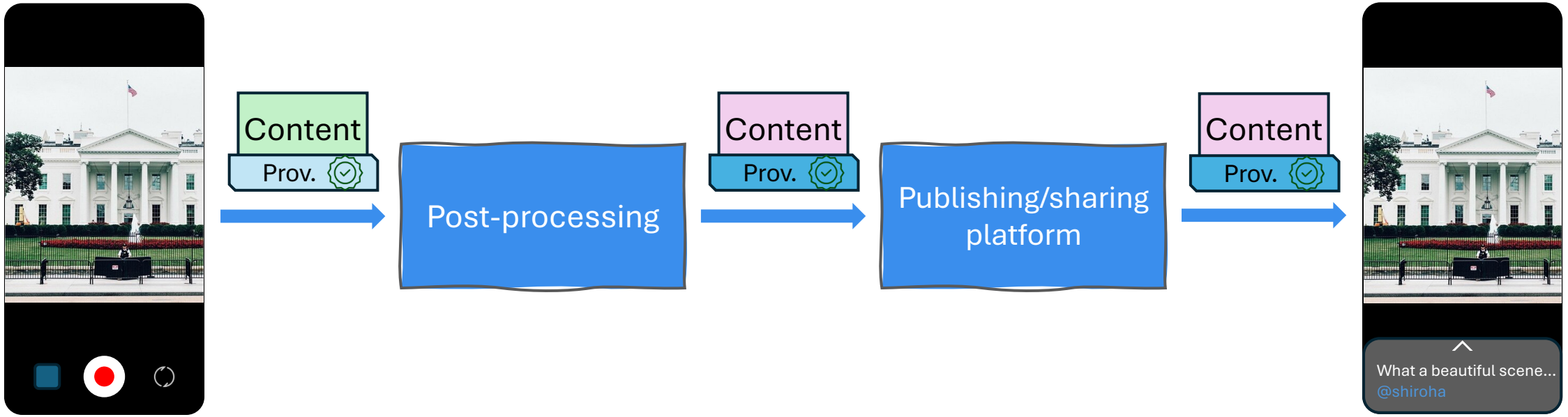
ProvCam (MobiCom '24)



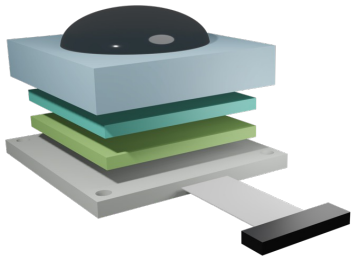
Vronicle (MobiSys '22)



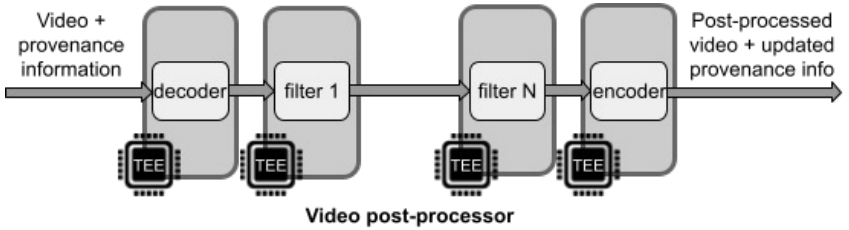
Provenance-based media authentication



ProvCam (MobiCom '24)



Vronicle (MobiSys '22)



Content Authenticity Initiative

C2PA

Adobe BBC Microsoft

...

Are we safe now?

Are we safe now?

- NO!

Challenge 1

- Provenance authenticates the origin and history of the content, but not the information within the content

Recapture attack in the provenance era

Recapture attack in the provenance era

Recapture attack in the provenance era

Recapture attack in the provenance era

Recapture attack in the provenance era

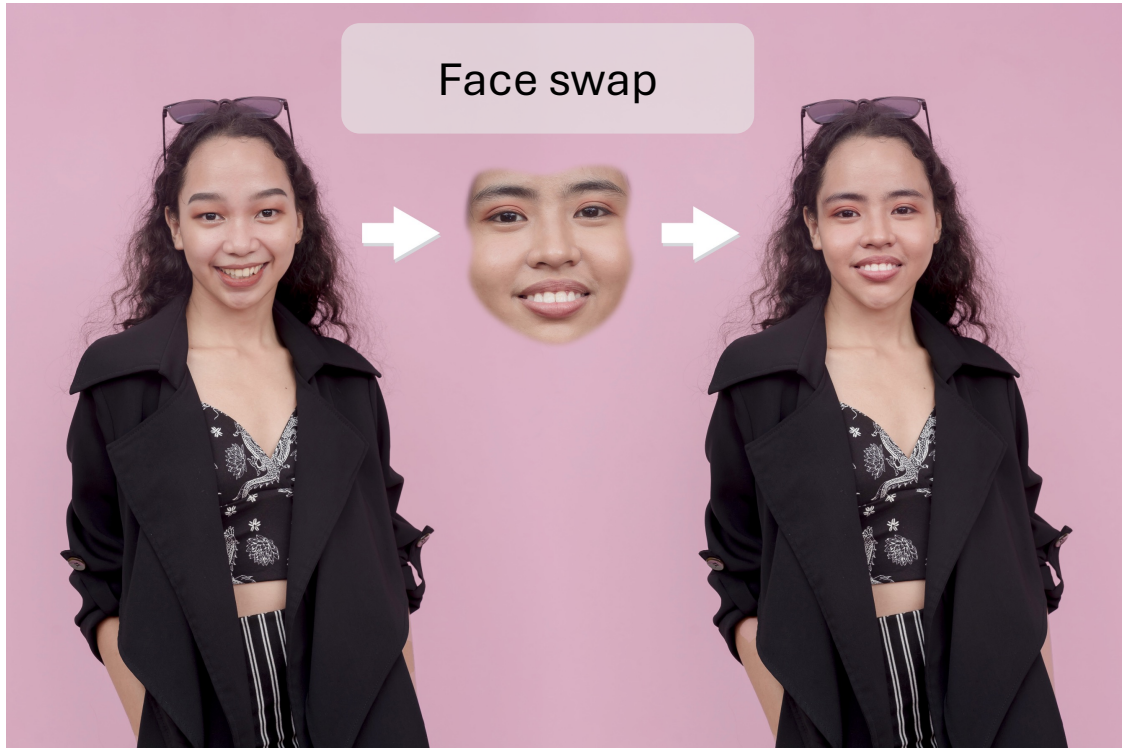
Recapture attack in the provenance era

Recapture attack in the provenance era

Classic recapture attack --- renewed

Classic recapture attack --- renewed

- Ease of content manipulation/fabrication (e.g., deepfake)



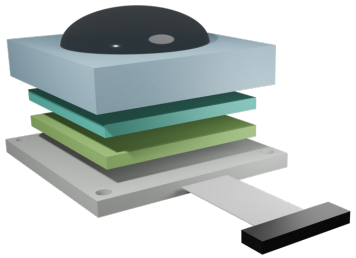
Classic recapture attack --- renewed

- Ease of content manipulation/fabrication (e.g., deepfake)
- Illusion of reality (e.g., advanced displays)

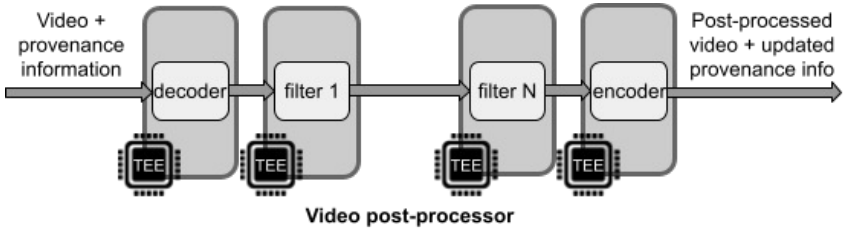
Classic recapture attack --- renewed

- Ease of content manipulation/fabrication (e.g., deepfake)
- Illusion of reality (e.g., advanced displays)
- False sense of trust (by provenance-based systems)

ProvCam (MobiCom ’24)



Vronicle (MobiSys ’22)



Content Authenticity Initiative

C2PA

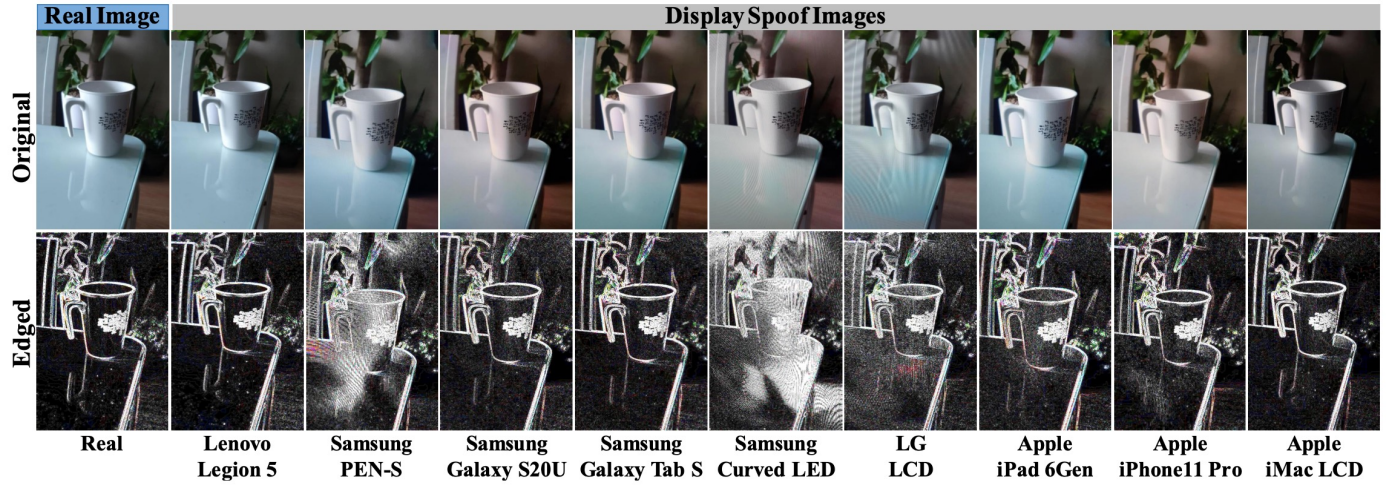
Adobe BBC Microsoft

…

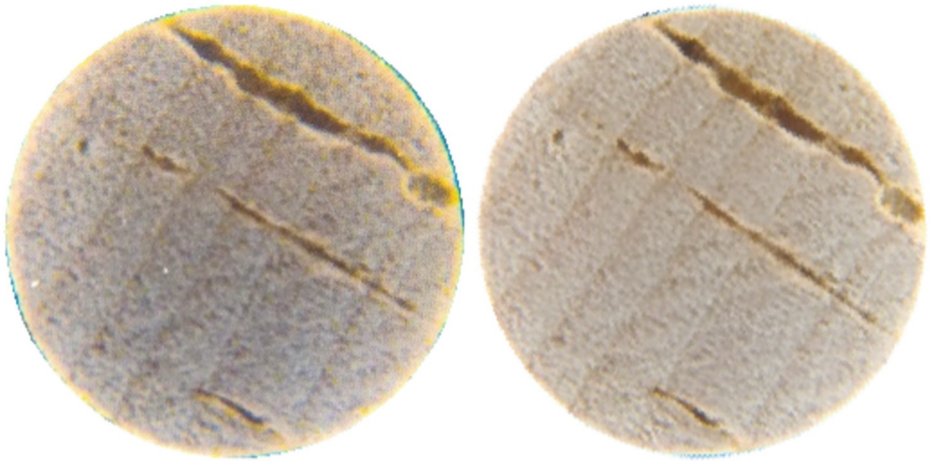
Any existing countermeasures?

Existing countermeasures

- Screen-based (e.g., LCD and OLED) recapture detection (e.g., using moiré pattern)
- Print-based (e.g., poster) recapture detection (e.g., using micro-surface pattern)



(Jeong et al.)



(Costa et al.)

Do they really work?

Do they really work?

- Not really

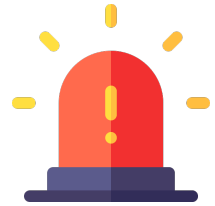
Do they really work?

- Not really
- Majority of them depend on prior knowledge of the recapture medium and camera

Do they really work?

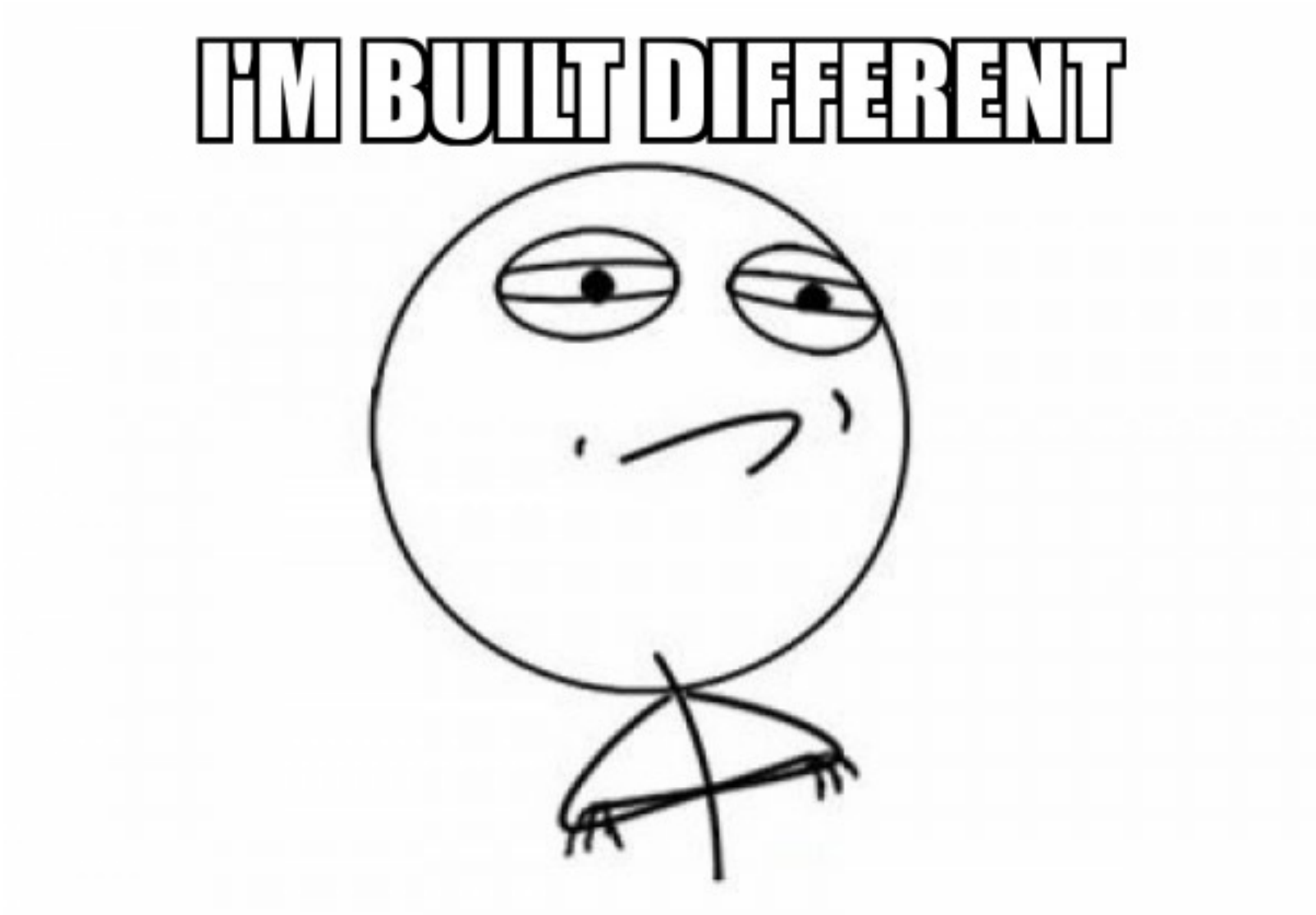
- Not really
- Majority of them depend on prior knowledge of the recapture medium and camera
- Attackers are **in control** here (for both the recapture medium and camera)

Spoiler alert



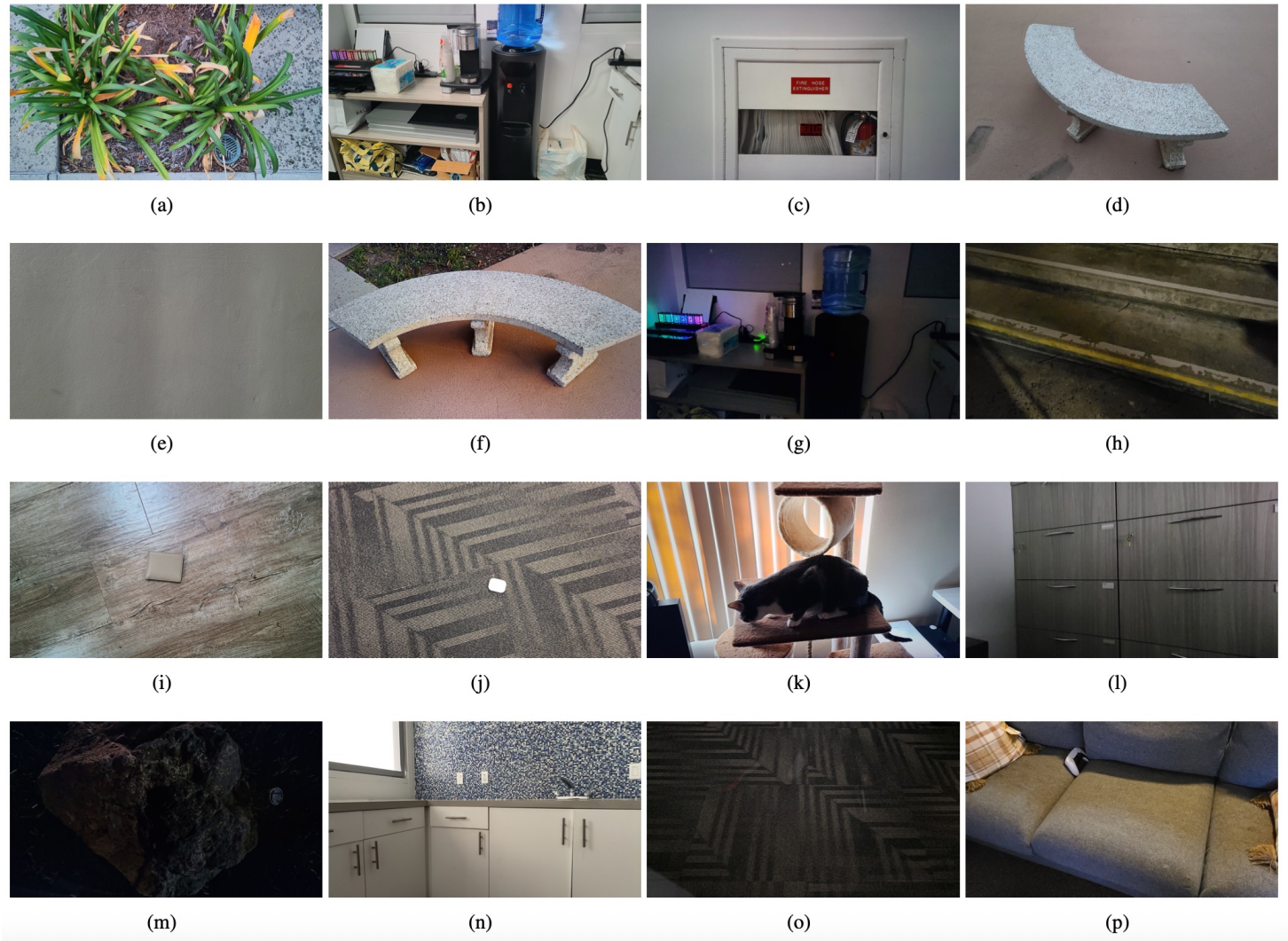
Chimera: Creating Digitally Signed Fake Photos by Fooling Image Recapture and Deepfake Detectors

Seongbin Park*, Alexander Vilesov*, Jinghui Zhang, Hossein Khalili,
Yuan Tian, Achuta Kadambi, Nader Sehatbakhsh
University of California, Los Angeles
{parkseongbin,vilesov}@ucla.edu


I'M BUILT DIFFERENT

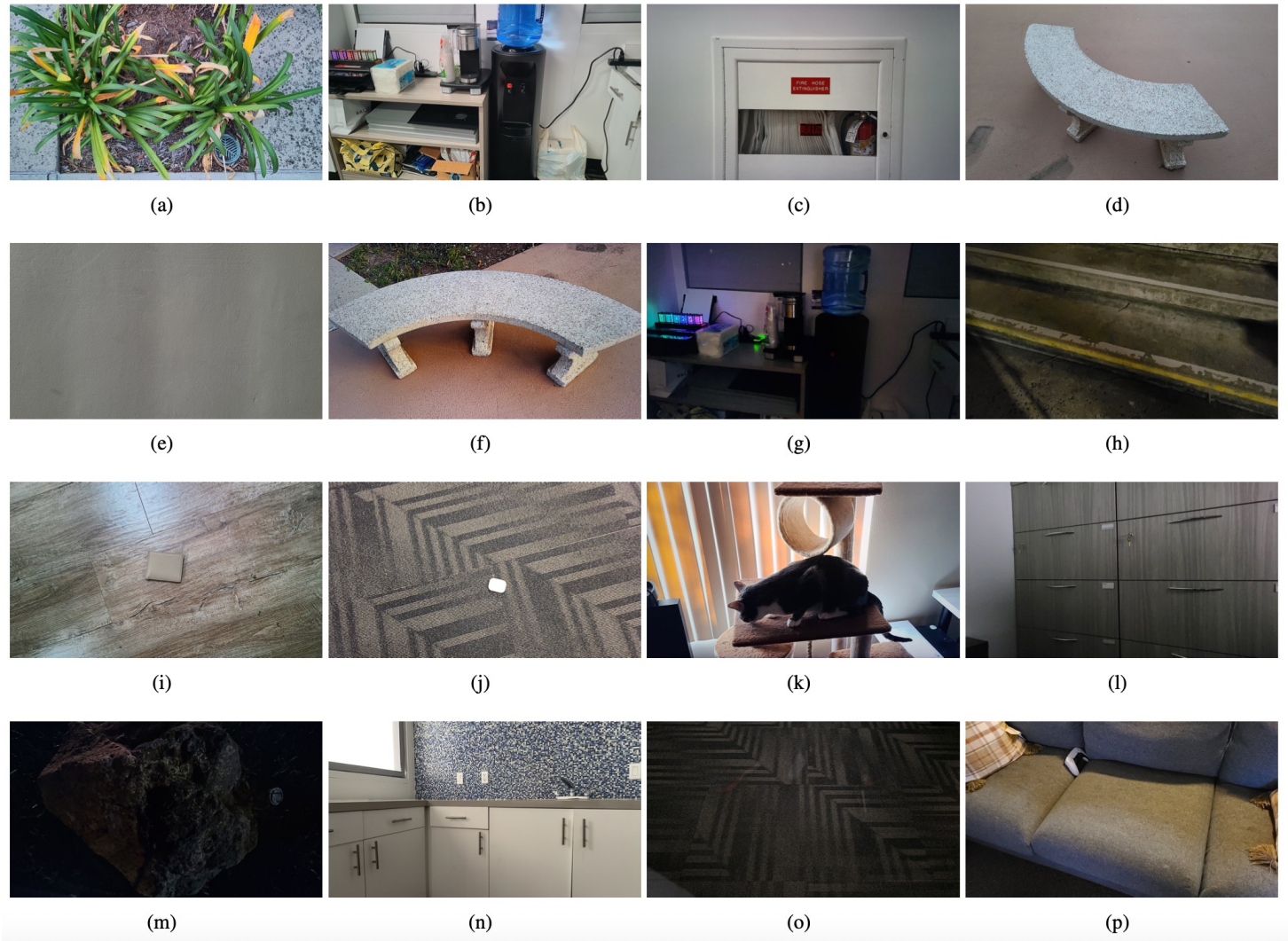
User study

- 16 photos (8 original / 8 recaptured)
- 43 adult participants
- All done on one 14-inch 2K display laptop



User study

- 16 photos (8 original / 8 recaptured)
- 43 adult participants
- All done on one 14-inch 2K display laptop
- Accuracy (correct classification) at 50.15% $(t(42) = 0.071, p = 0.944)$



So how do we tackle such an attack?

So how do we tackle such an attack?

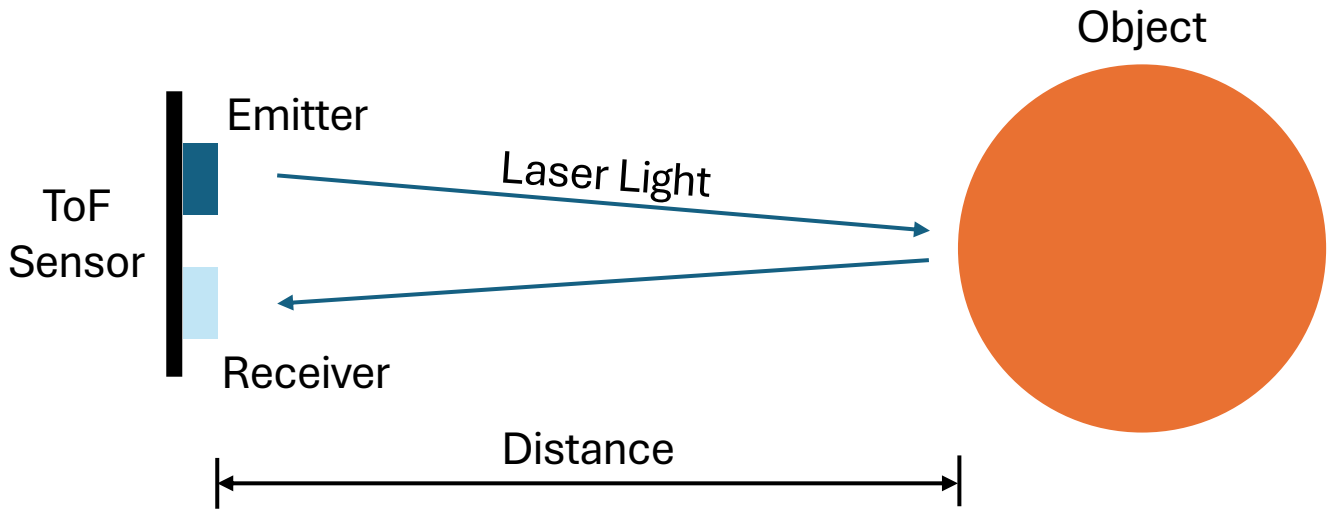
- Aftermath analysis-based approaches are an endless race that’s unlikely to win

So how do we tackle such an attack?

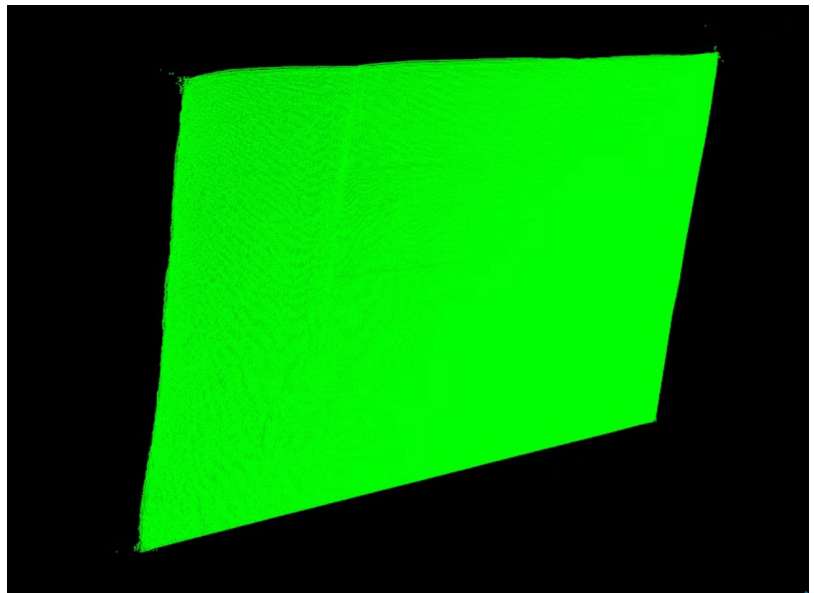
- Aftermath analysis-based approaches are an endless race that's unlikely to win
- Provenance can still help

Depth-enabled provenance

- Time-of-Flight depth sensors (dToF and iToF)
- Measure distance to a target
- Equipped on modern smartphones (e.g., Apple, Samsung, etc.)




Depth Prov.

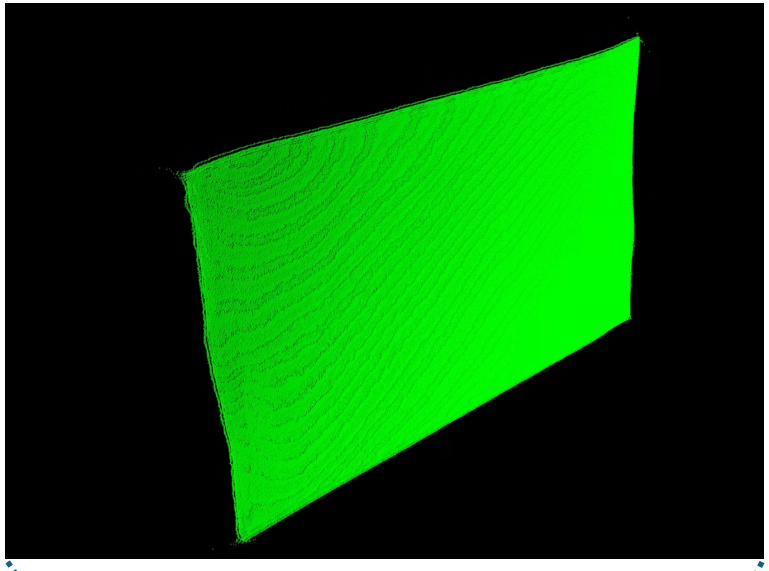


Depth Prov.

Challenge 2

- Not all flat surfaces are recaptures


Depth Prov.



Depth Prov.

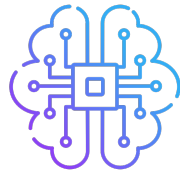
Catching misleading recaptures

Catching misleading recaptures

- Let’s compare human perception of depth with the provenance depth

Catching real recaptures

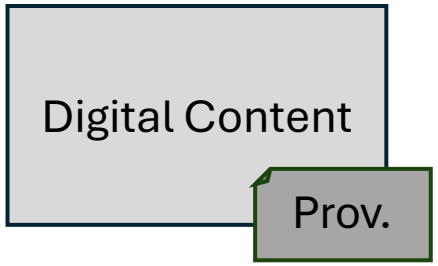
- Let’s compare human perception of depth with the provenance depth
  - Learning-based monocular depth estimation



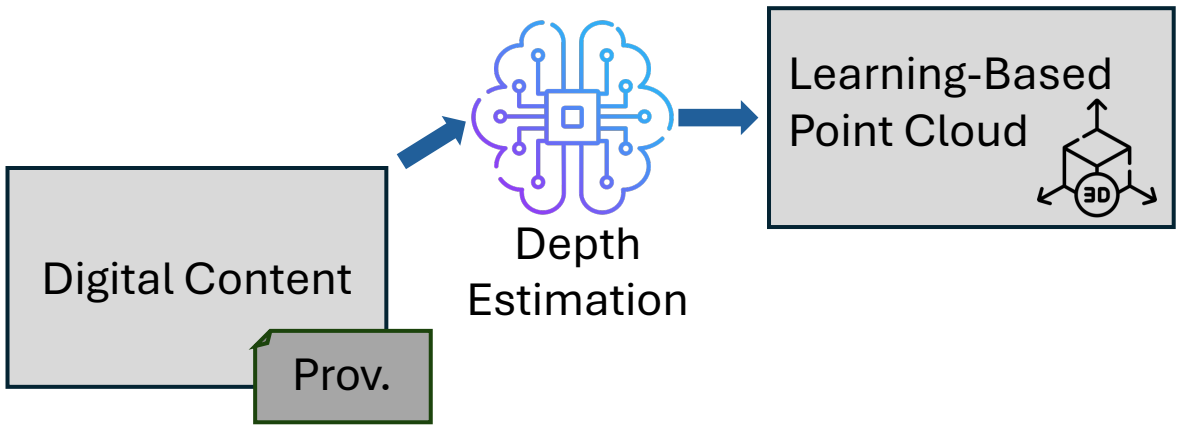
Introducing Scoop

- **In**: A photo/video with depth-enabled provenance
- **Out**: The same photo/video with misleading recapture regions highlighted

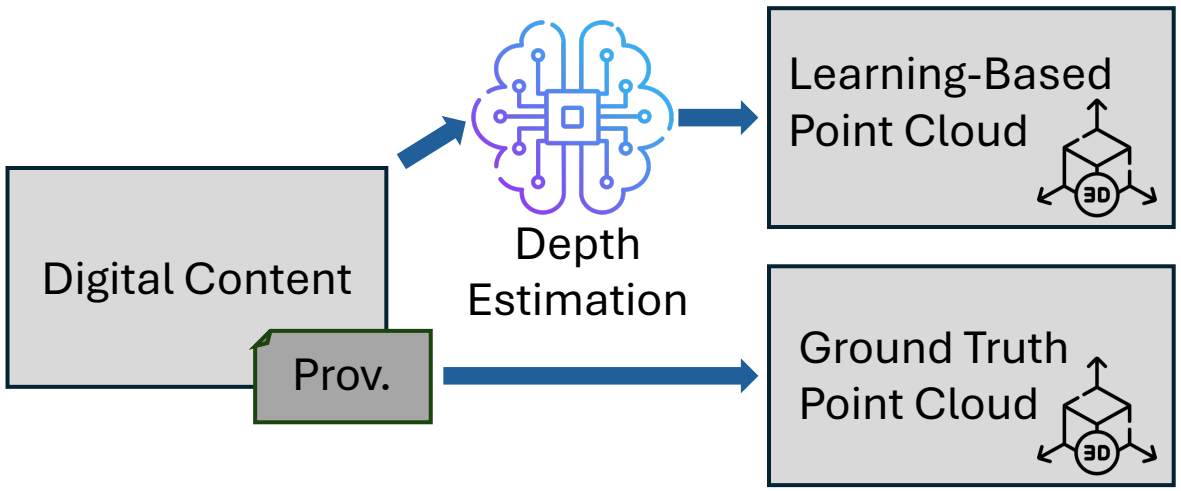
Scoop’s design



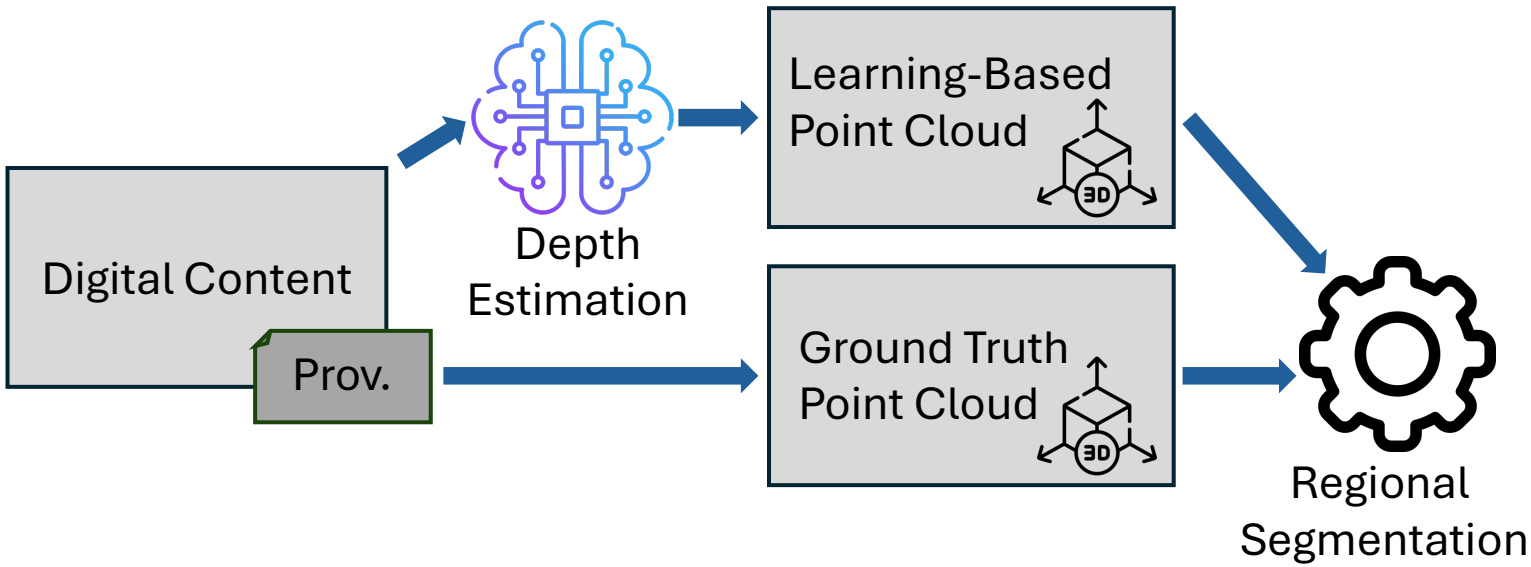
Scoop’s design



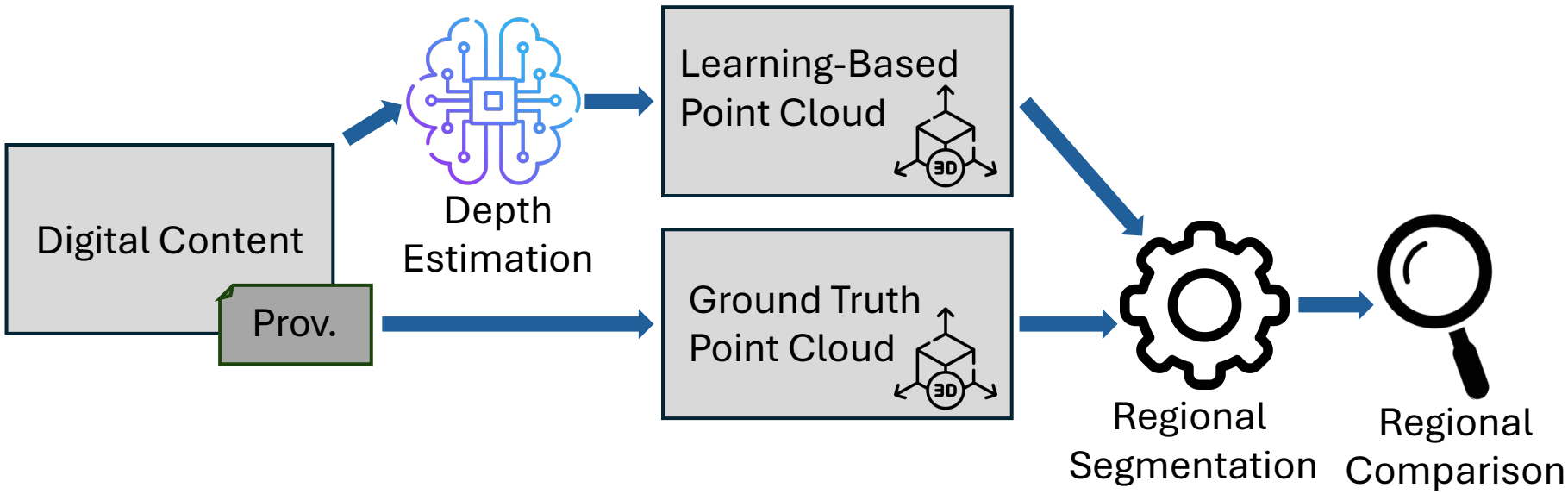
Scoop’s design



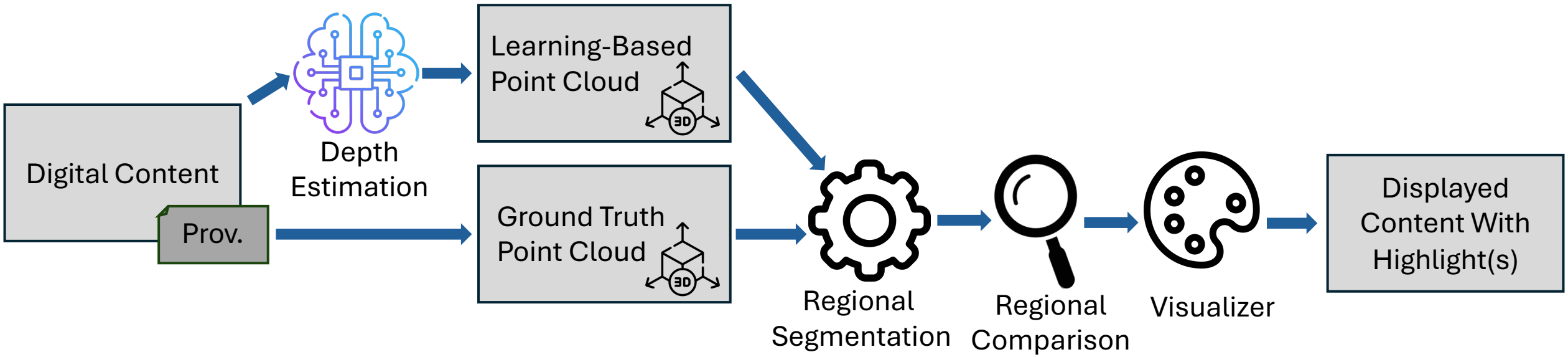
Scoop’s design



Scoop’s design



Scoop’s design



Scoop in action

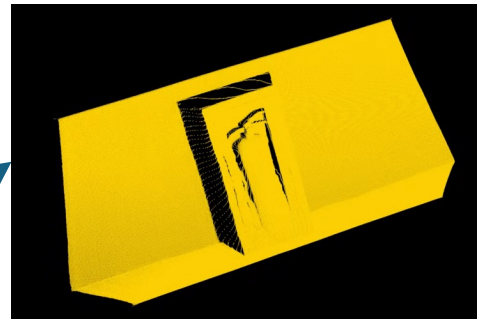
Captured photo
(w/ Depth Provenance)

Scoop in action

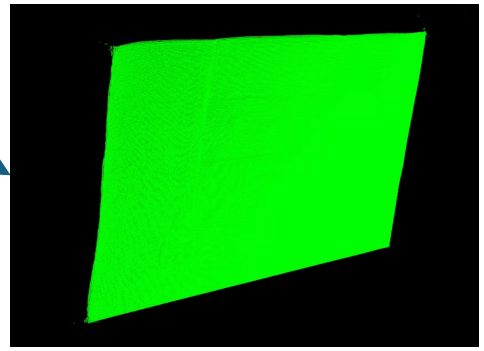
Captured photo
(w/ Depth Provenance)

Scoop in action

Captured photo
(w/ Depth Provenance)



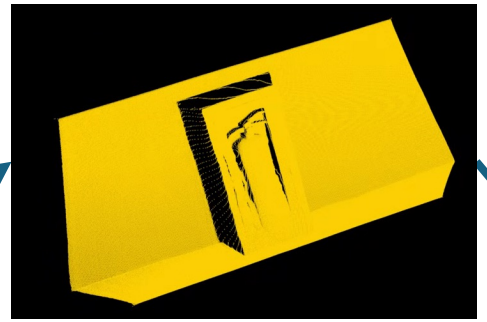
Learning-based point cloud



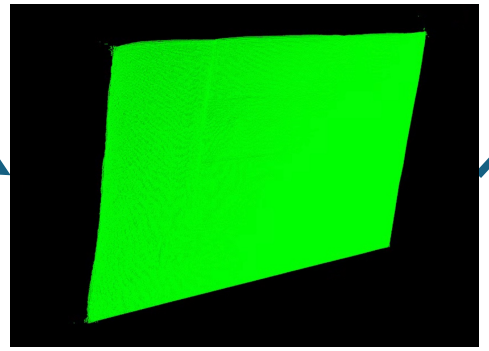
Ground truth point cloud

Scoop in action

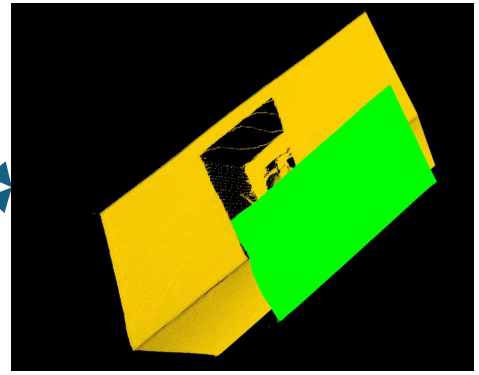
Captured photo
(w/ Depth Provenance)



Learning-based point cloud



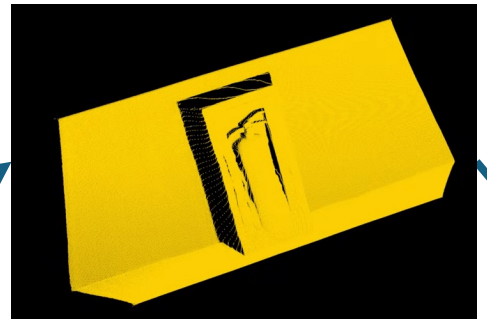
Ground truth point cloud



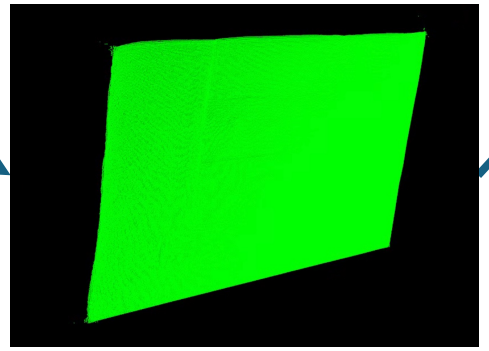
Comparison

Scoop in action

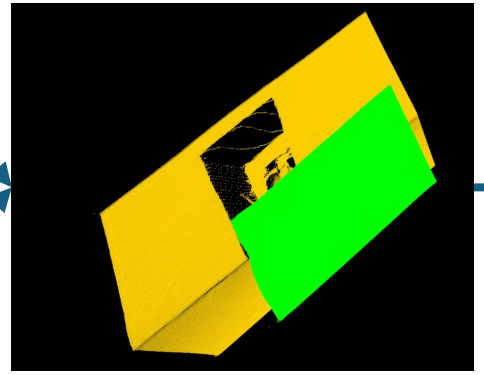
Captured photo
(w/ Depth Provenance)



Learning-based point cloud



Ground truth point cloud



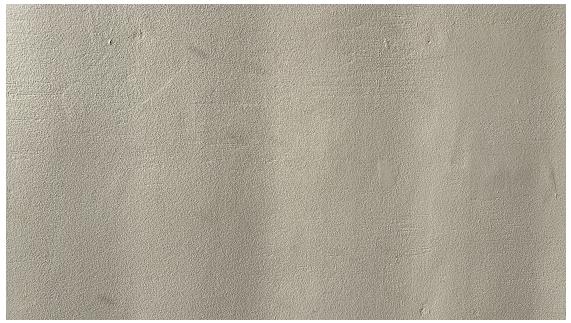
Comparison

Displayed photo
(w/ Highlight)

Scoop in action

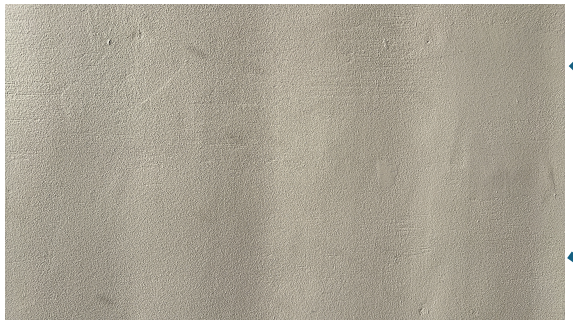
Captured photo
(w/ Depth Provenance)

Scoop in action

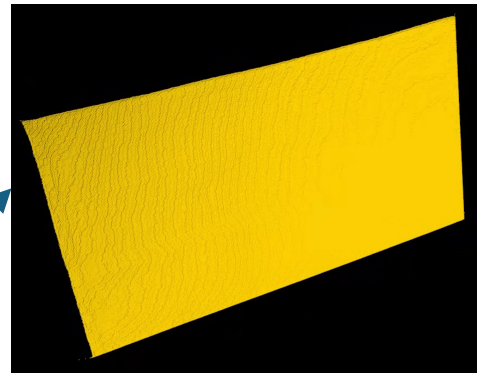


Captured photo
(w/ Depth Provenance)

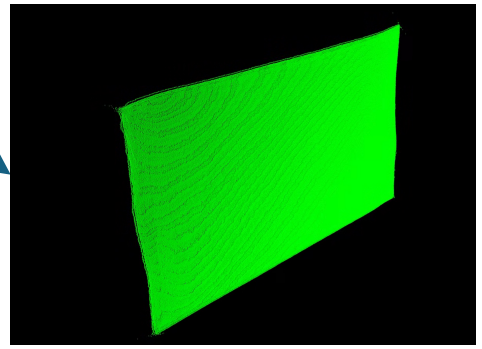
Scoop in action



Captured photo
(w/ Depth Provenance)

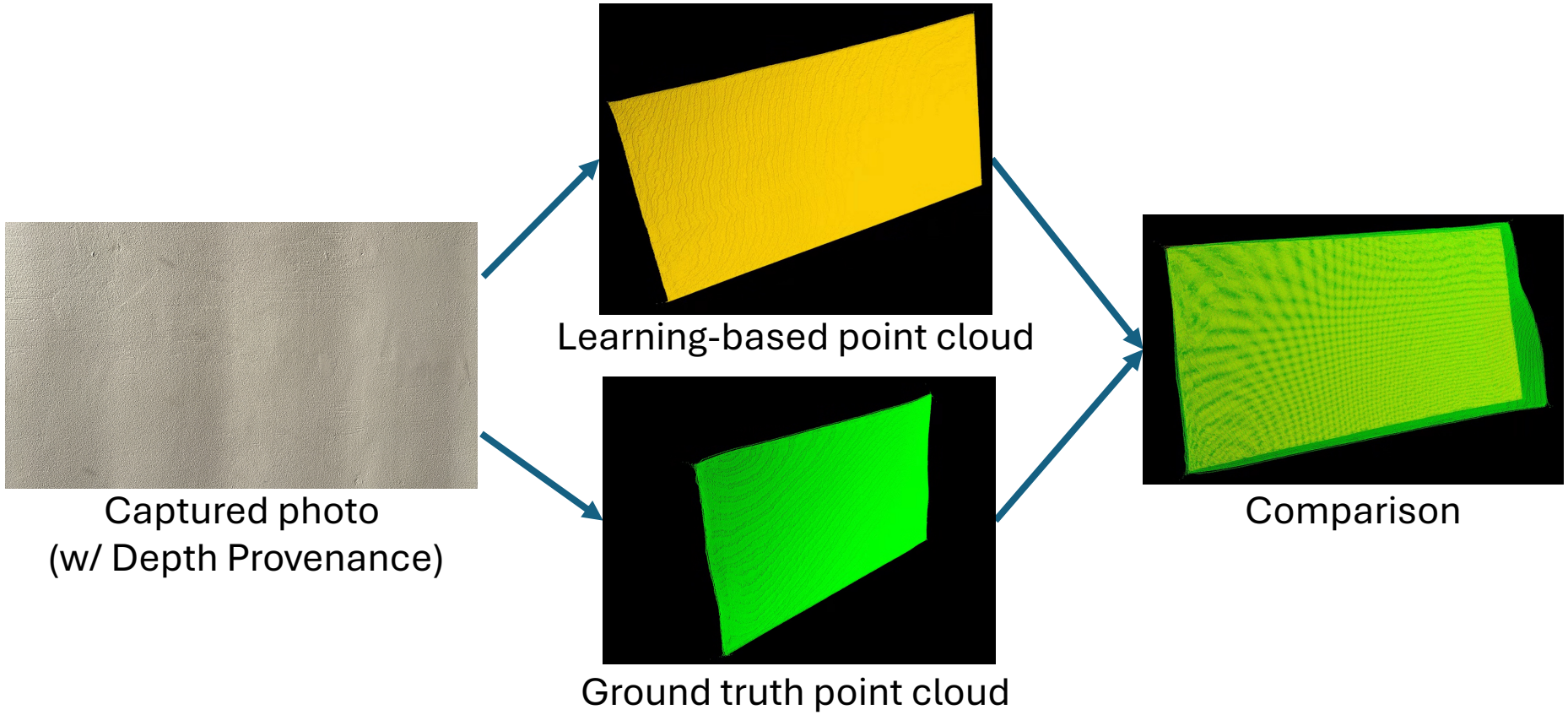


Learning-based point cloud

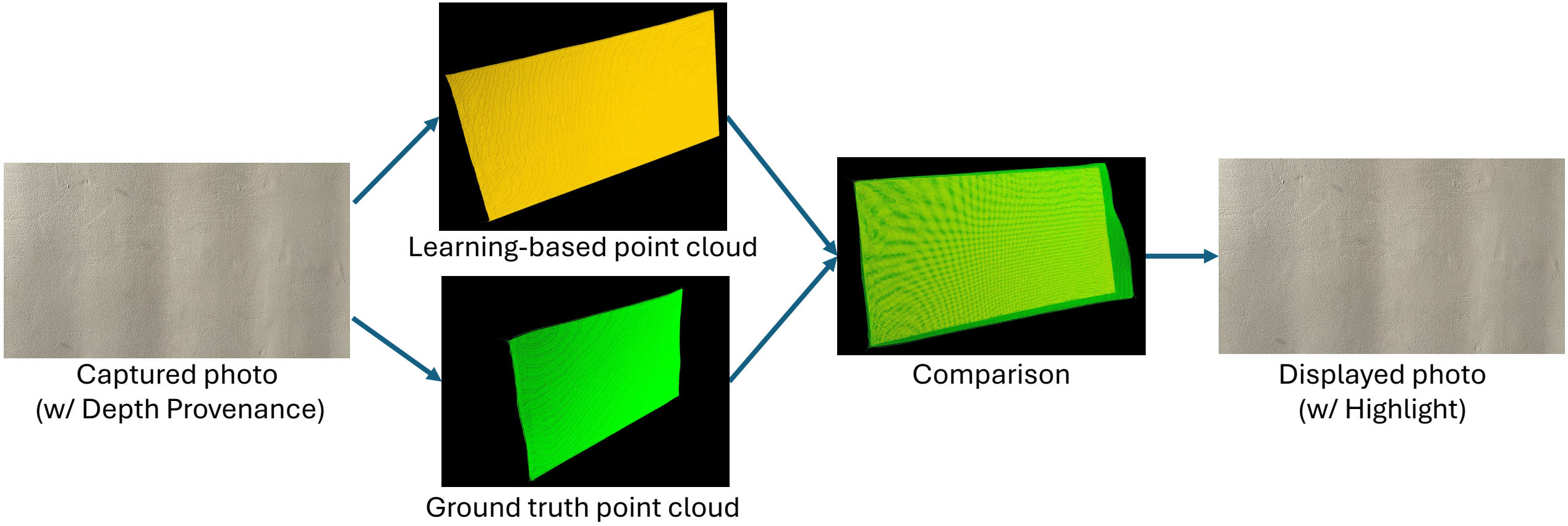


Ground truth point cloud

Scoop in action



Scoop in action



Scoop in action

Captured photo
(w/ Provenance)

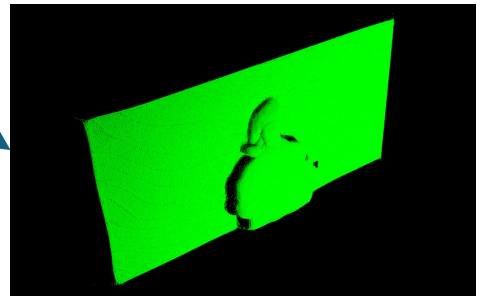
Scoop in action

Captured photo
(w/ Provenance)

Scoop in action

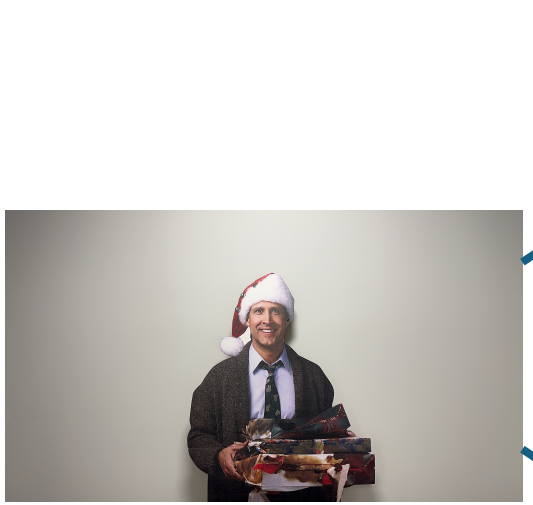
Captured photo
(w/ Provenance)

Learning-based point cloud

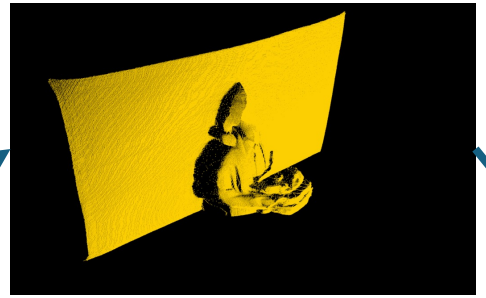


Ground truth point cloud

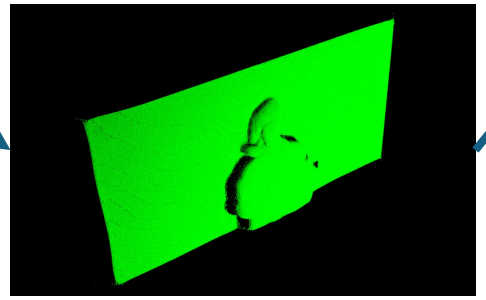
Scoop in action



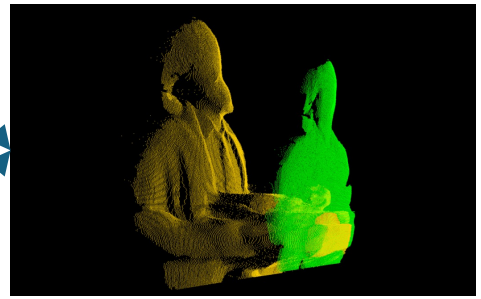
Captured photo
(w/ Provenance)



Learning-based point cloud

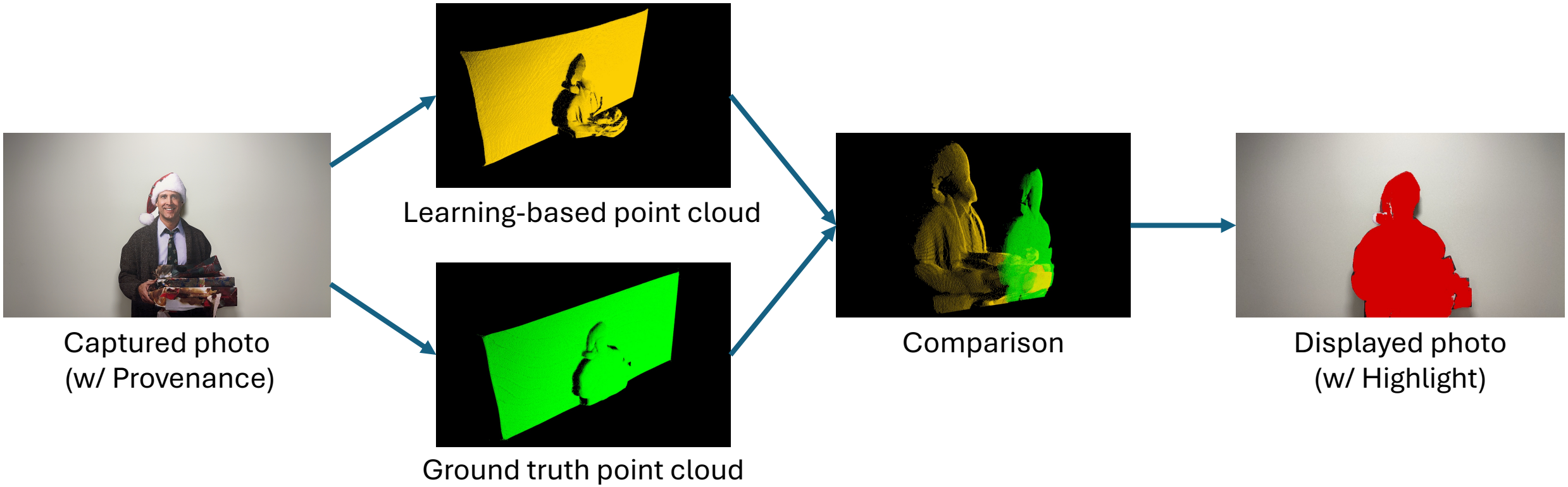


Ground truth point cloud



Comparison

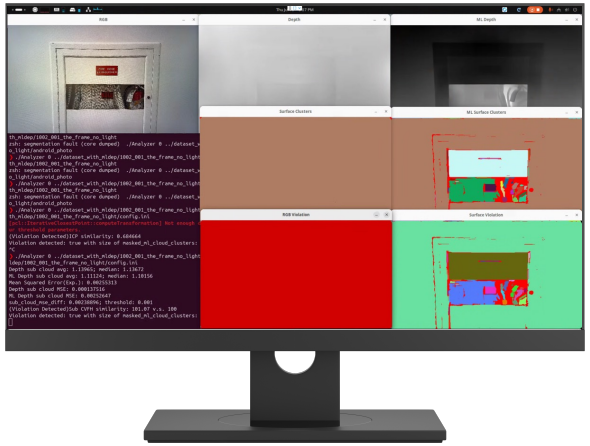
Scoop in action



Prototypes

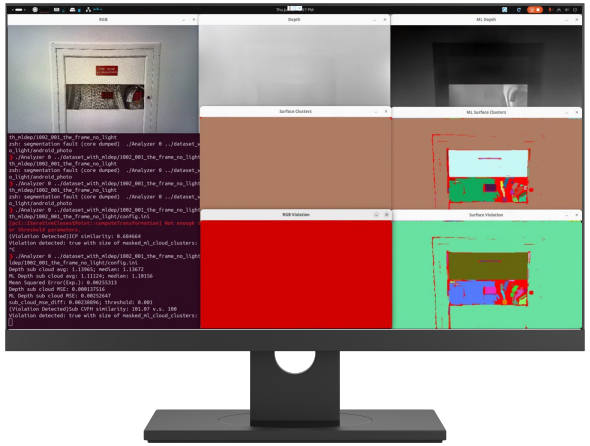
Prototypes

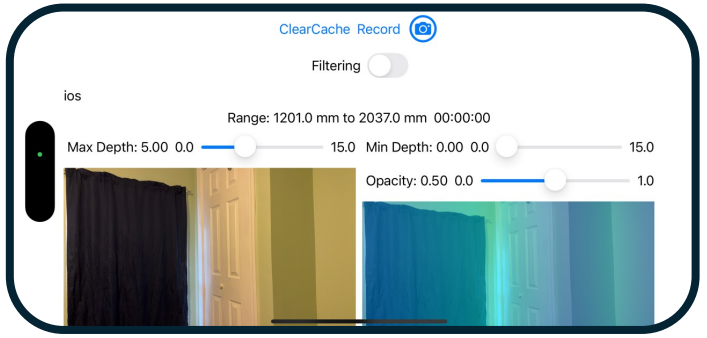
- A desktop Viewer application (OpenCV and Point Cloud Library)



Prototypes

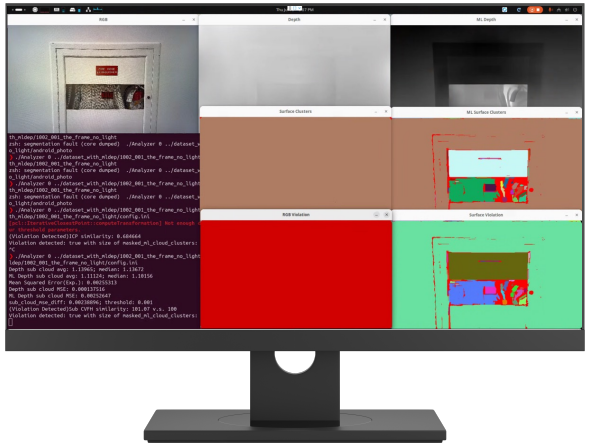
- A desktop Viewer application (OpenCV and Point Cloud Library)
- An iOS camera application

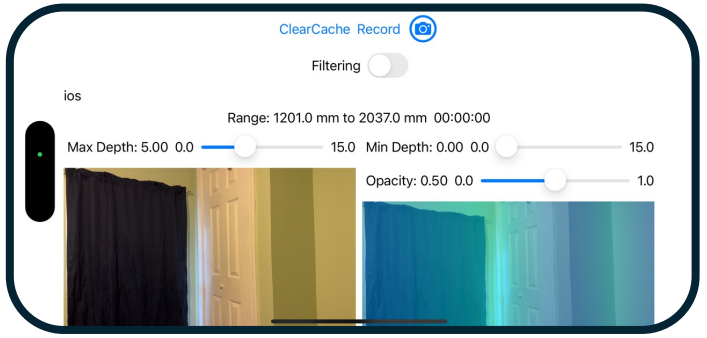


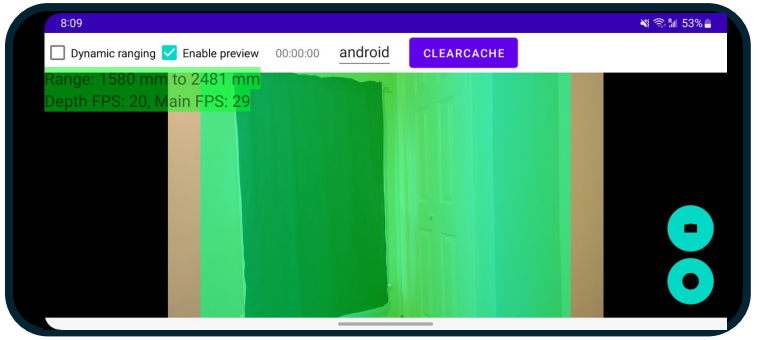


Prototypes

- A desktop Viewer application (OpenCV and Point Cloud Library)
- An iOS camera application
- An Android camera application







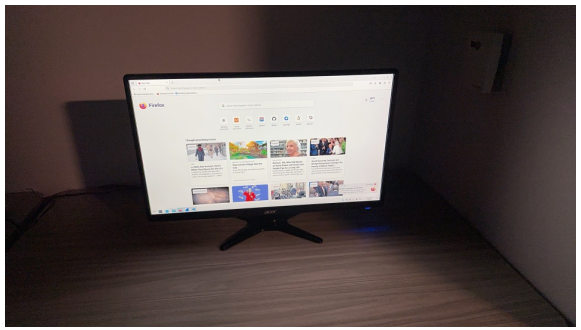
Evaluation

Challenge 3

- No existing dataset designed for evaluating systems like Scoop

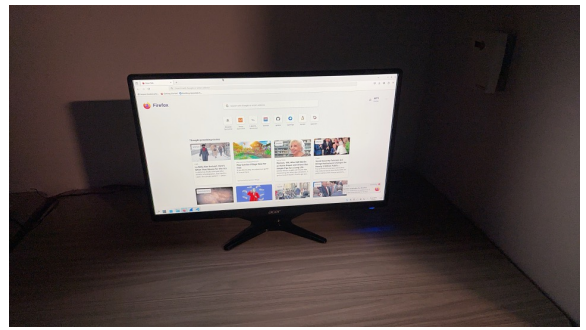
A first-of-its-kind dataset

- Captured using our two smartphone prototypes



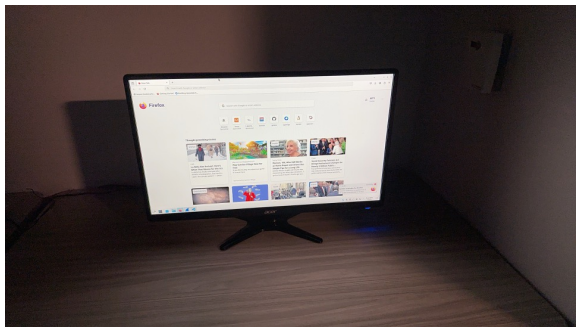
A first-of-its-kind dataset

- Captured using our two smartphone prototypes
- 122 unique data points
- 78 recapture scenarios (TVs, cardboard cutouts, projector, and mixed)



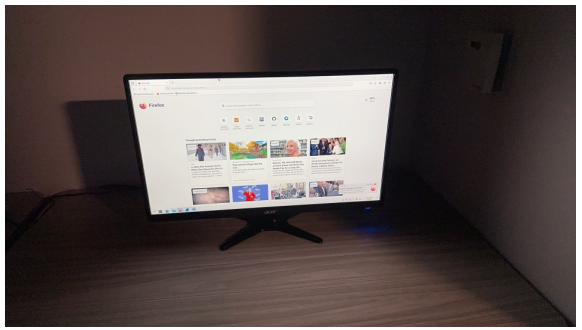
A first-of-its-kind dataset

- Captured using our two smartphone prototypes
- 122 unique data points
- 78 recapture scenarios (TVs, cardboard cutouts, projector, and mixed)
- Both photos and videos



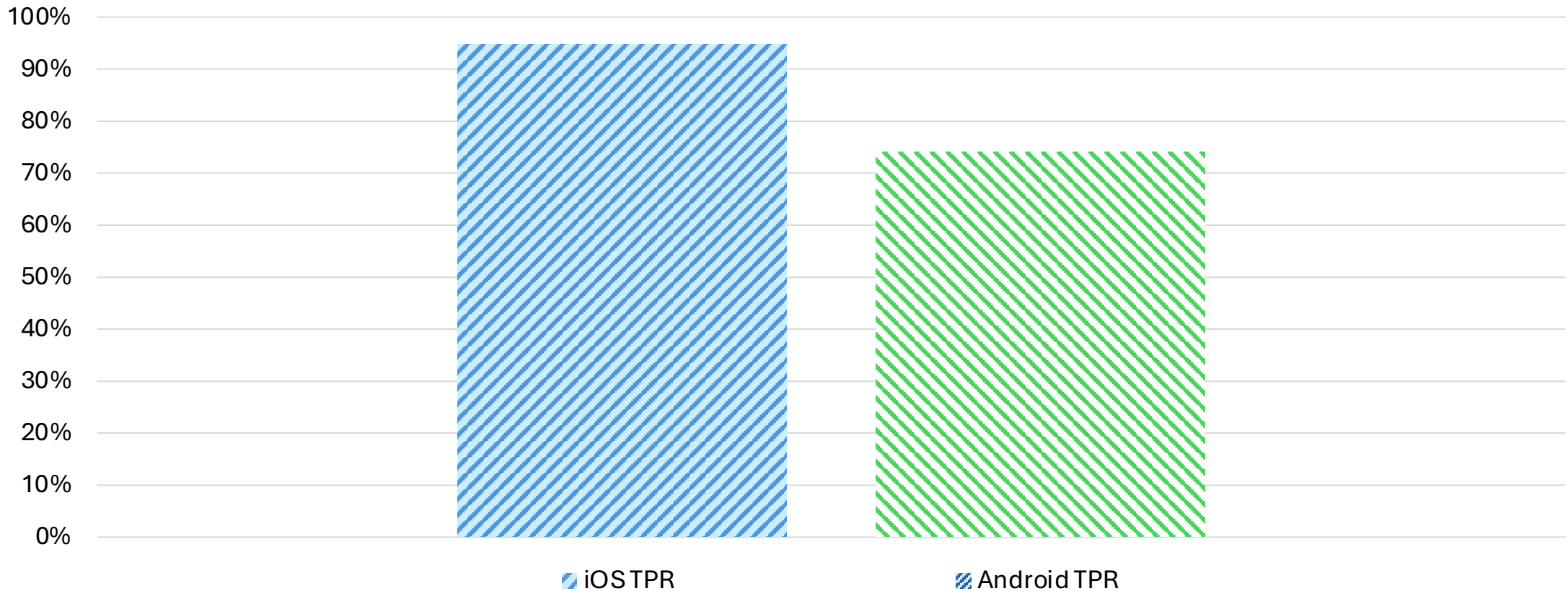
A first-of-its-kind dataset

- Captured using our two smartphone prototypes
- 122 unique data points
- 78 recapture scenarios (TVs, cardboard cutouts, projector, and mixed)
- Both photos and videos
- Cornerstone for evaluating systems such as Scoop



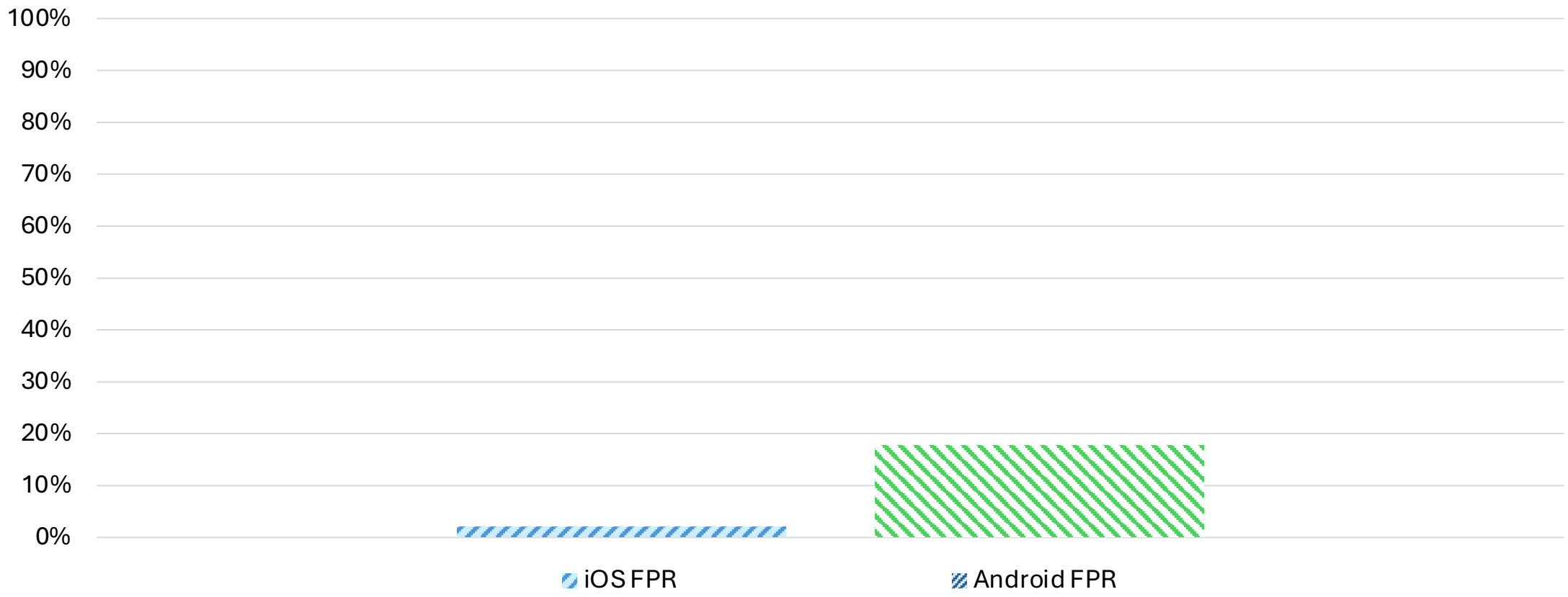
Effectiveness

True positive rate (TPR): Rate of correctly identified misleading recapture

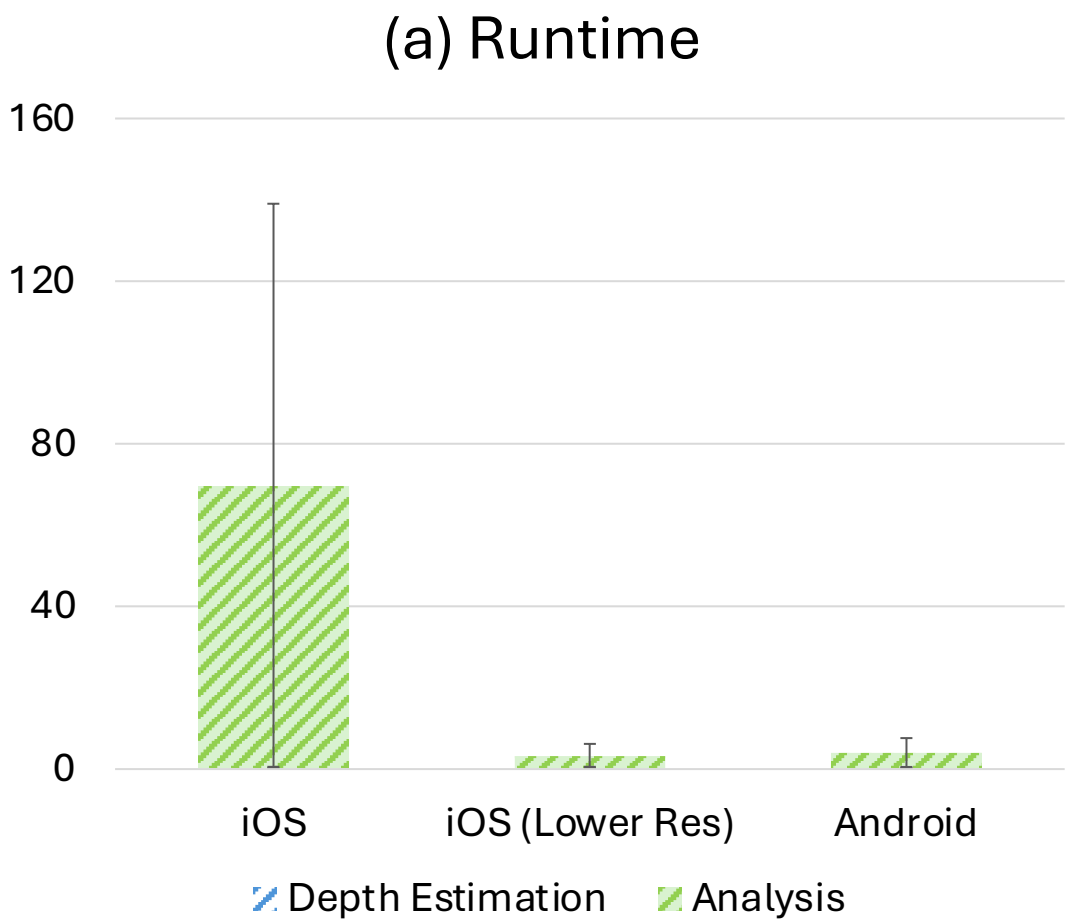


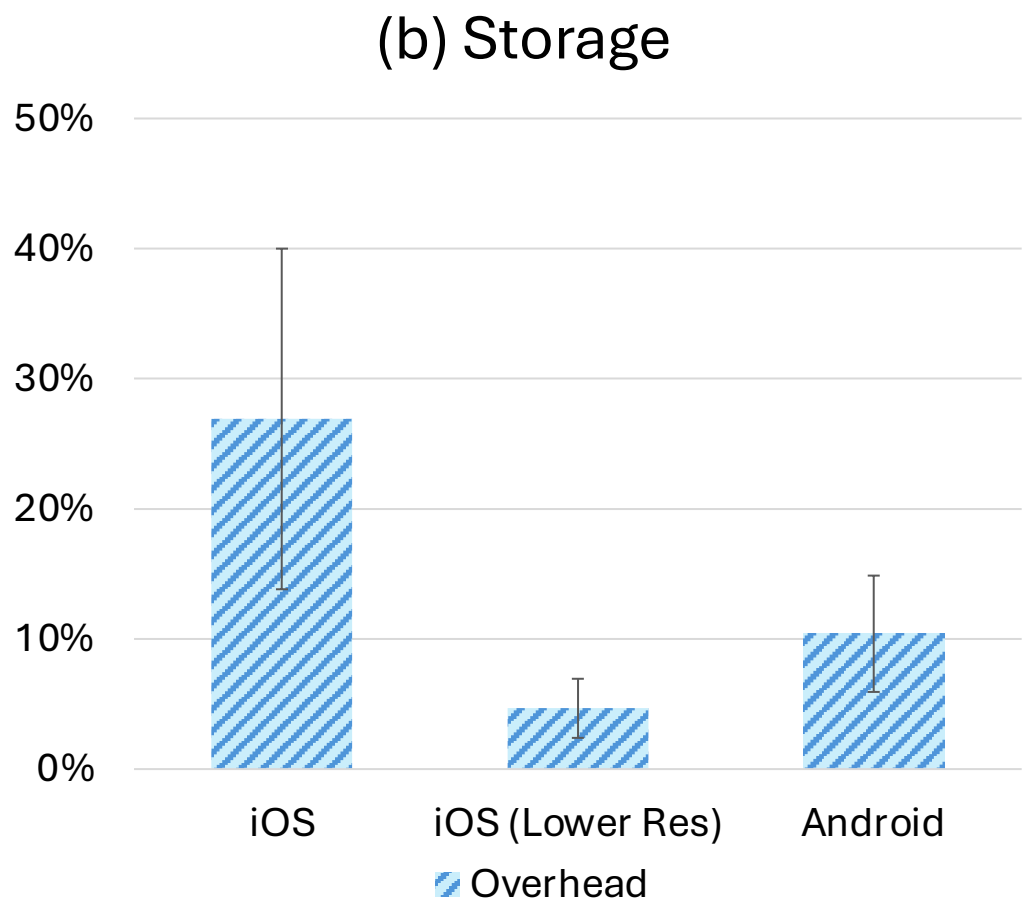
Effectiveness

False positive rate (FPR): Rate of incorrectly identified benign as recapture



Overhead


(a) Runtime
160
120
80
40
0
iOS
iOS (Lower Res)
Android
Depth Estimation
Analysis


(b) Storage
50%
40%
30%
20%
10%
0%
iOS
iOS (Lower Res)
Android
Overhead

Summary

- Shed light on an important vulnerability in existing provenance-based techniques

Summary

- Shed light on an important vulnerability in existing provenance-based techniques
- Design of Scoop, an effective countermeasure against recapture attacks

Summary

- Shed light on an important vulnerability in existing provenance-based techniques
- Design of Scoop, an effective countermeasure against recapture attacks
- A first-of-its-kind dataset for evaluating systems like Scoop

Summary

- Shed light on an important vulnerability in existing provenance-based techniques
- Design of Scoop, an effective countermeasure against recapture attacks
- A first-of-its-kind dataset for evaluating systems like Scoop
- Fully functional Scoop prototype based on commodity devices

Summary

Paper: **Code:**

- Shed light on an important vulnerability in existing provenance-based techniques
- Design of Scoop, an effective countermeasure against recapture attacks
- A first-of-its-kind dataset for evaluating systems like Scoop
- Fully functional Scoop prototype based on commodity devices

Thank you!

Questions? Feel free to email to yuxin.liu@uci.edu