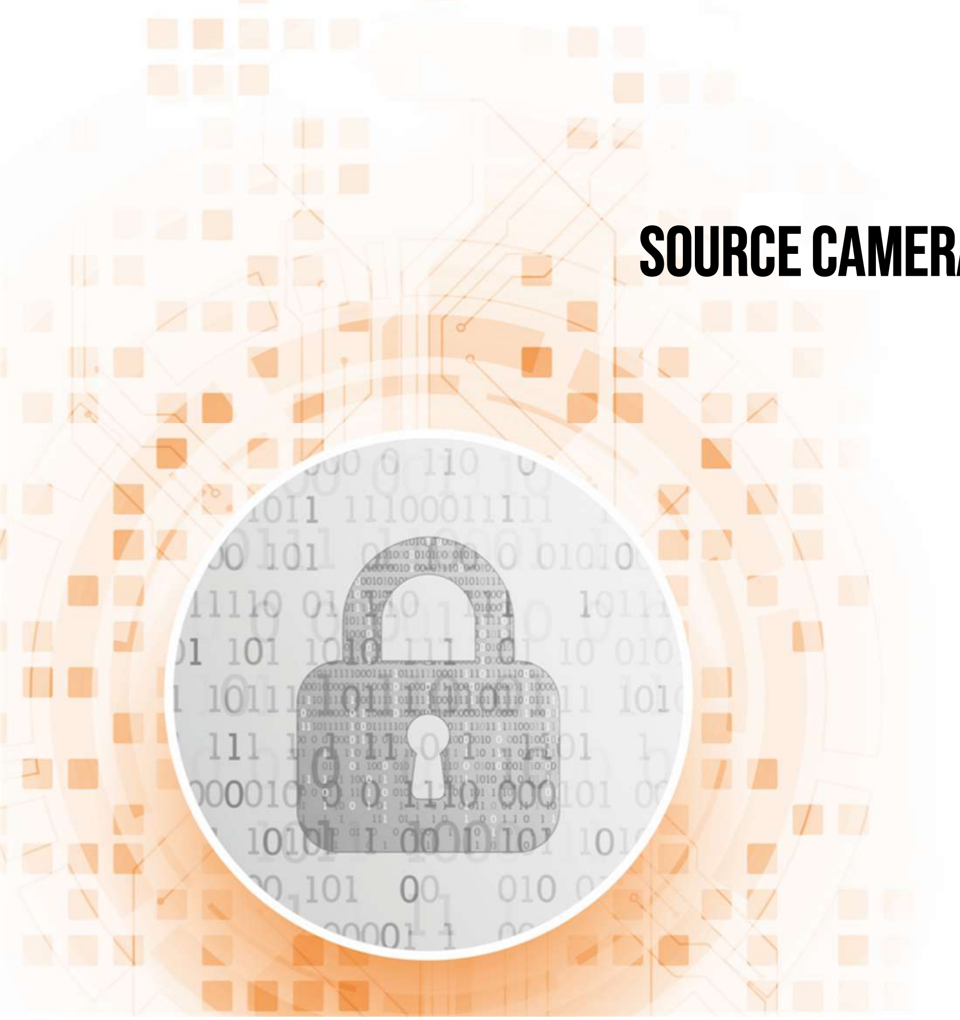
Research Institute
Cyber Defence
Universität der Bundeswehr München



METRICS MATTER

SOURCE CAMERA FORENSICS FOR LARGE-SCALE INVESTIGATIONS

Samantha Klier and Harald Baier

Research Institute CODE

University of the Bundeswehr Munich

Introduction | Source Camera Forensics and Use Cases

- Source Camera Forensics (SCF) links images to devices/models/brands
- Two phases: Investigation (screening) vs. Examination (verification)

INTRODUCTION | PROBLEM

- Sensor Pattern Noise (SPN) approach, developed and highly effective for verification…
- …also applied to large-scale screening, but…
- …requirements of the investigative phase have not been embraced:

Minimize evidence loss vs. False Positive Rate
Huge Image Sets vs. no efficiency concerns
No curation possible vs. problems with "post"-processed images

››› *Evaluation of 3 SCF techniques for investigation.*

RLW | USE CASE REQUIREMENTS

- **Examination | Verification:**
- Primary Aim → Minimize false convictions → False Acceptance Rate ↓
- Secondary Aim → Minimize false exonerations → False Negative Rate ↓
- **Investigation | Identification:**
- Primary Aim → Minimize Evidence Loss → True Positive Rate (Recall) ↑
- Secondary Aim → Maximize Data Reduction → Precision ↑

Only 5% of SCF approaches have been evaluated for Investigations, only 2 for images (evaluated on 2010's DIDB and not available)

RLW | Sensor Pattern Noise (SPN)

- Sensor Pattern Noise (SPN) approach: "gold standard" for Verification
- $N(I) = I - F(I) \rightarrow$ Camera "Fingerprint": average $N(I)$'s
- Cross correlate $N(I)$ with Camera Fingerprint
- Calculate PCE $> 60 \rightarrow$ Match
- 2009:
  - False Acceptance Rate of $2.4 * 10^{-5}$,
  - False Negative Rate of < 0.0238
- 2021: Concerns raised for bokeh images & several smartphone models

RLW | COMPARE

- Efficient SPN derivative: Computational & storage costs of classic SPN is a problem for large scale applications
- Extract noise residuals (e.g. acc. SPN approach)
- Divide noise residuals into sub-matrices
- Save only trace for each sub-matrix → constant compact size of e.g. 640x480px
- Use compact representation for comparison steps (e.g. acc. SPN approach)
- BUT: evaluated in terms of ROC/AUC (=TPR/FPR)

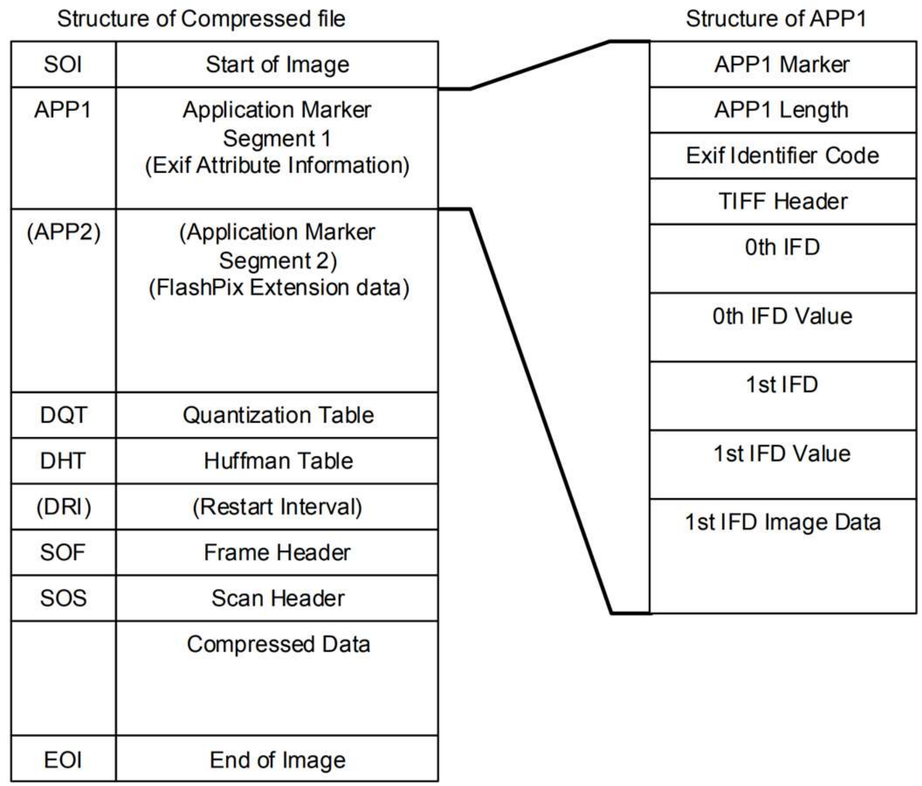
RLW | MEDIA SOURCE SIMILARITY HASHING (MSSH)

- based on JPEG structural information
- Extract JPEG and APP1 tags, build 2-grams, save in a set, SD by concatenation

```
C0C4C4C4C4DDD8E1D900DAD9DBC0DBDBDDDAE0DBE1E0

010001010101010201020100F010F0110011001120112011A
↪011A011B011B01280128013101310132013202130213 87
↪69876988258825A40BA40B0180
```

- Unify sets of several images to get *source SD*
- BUT: evaluated in terms of ROC/AUC (=TPR/FPR)



METHODOLOGY

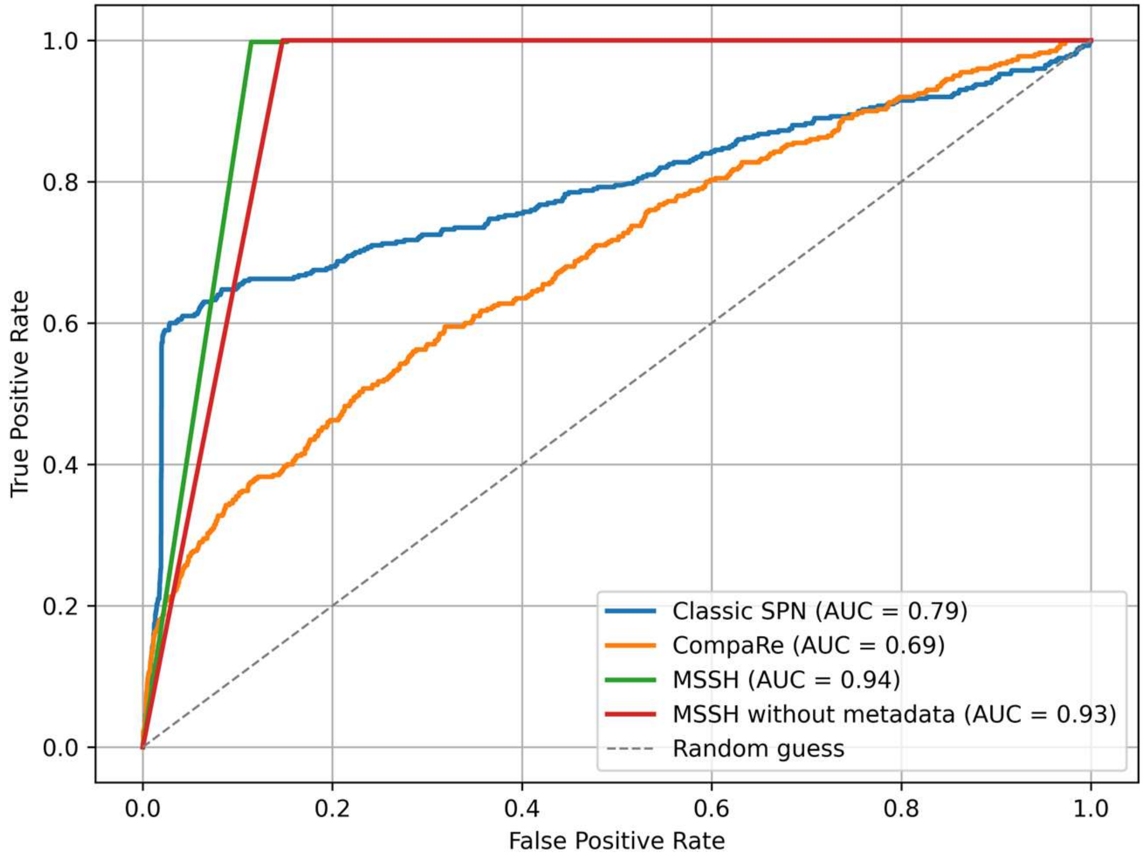
- Selected approaches:
  - Classic SPN:
    „gold standard", made for verification
  - CompaRe:
    by ROC/AUC superior to other efficient SPN approaches
  - MSSH:
    dedicated for large scale applications, superior to CompaRe by ROC/AUC

PrnuModernDevices Data Set	
Date of Publication	2021
Number of Devices	22
Number of Unique Models	17
Number of Images	550
Number of JPEGs	520
Types of Images	flat, natural, bokeh
Number of Images for Reference Generation	154
Number of Images for Evaluation	366

- Anti-Forensic: MSSH & Metadata
- Evaluation: PRC, ROC/AUC *to devices*
- Execution:
  - commodity hardware, single-threaded, no optimization

RESULTS | COMMON ROC/AUC

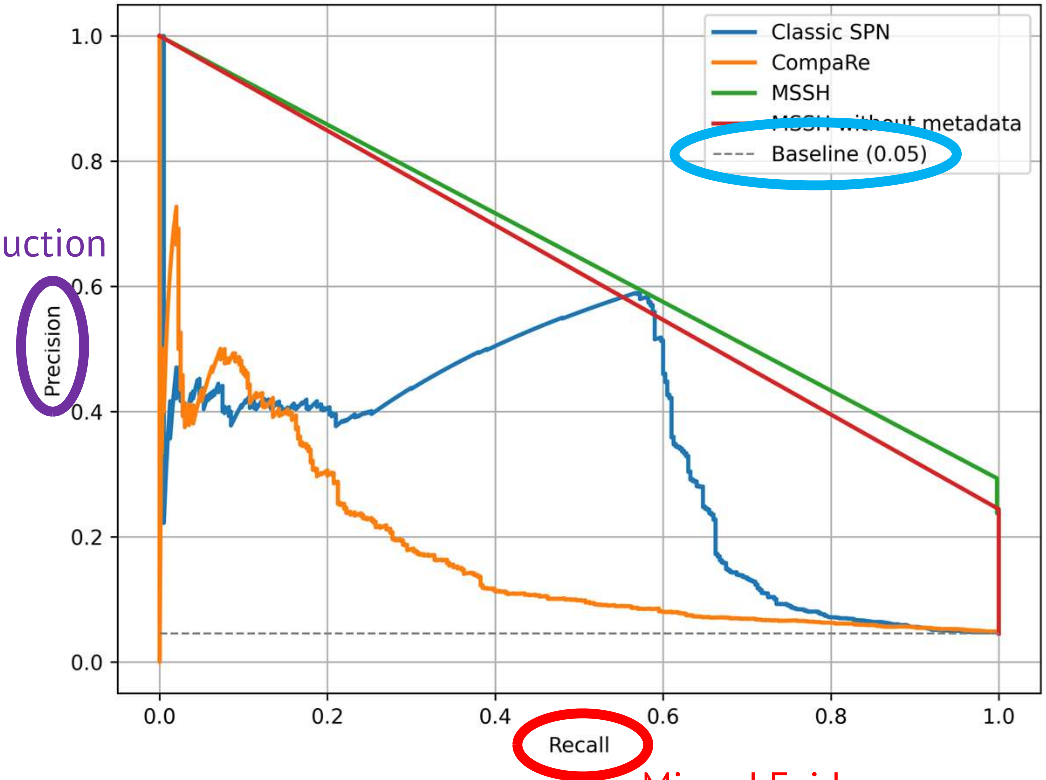
- **SPN / CompaRe:**
  below expectations?
- **MSSH:**
  robust without metadata



RESULTS | OVERALL PERFORMANCE

- **CompaRe:**
  - Precision & Recall low
  - **Threshold:**
    PREC/REC trade of possible, fast decline
  - **In practice:**
    threshold adaptable to case,
    e.g. 40% evidence missed

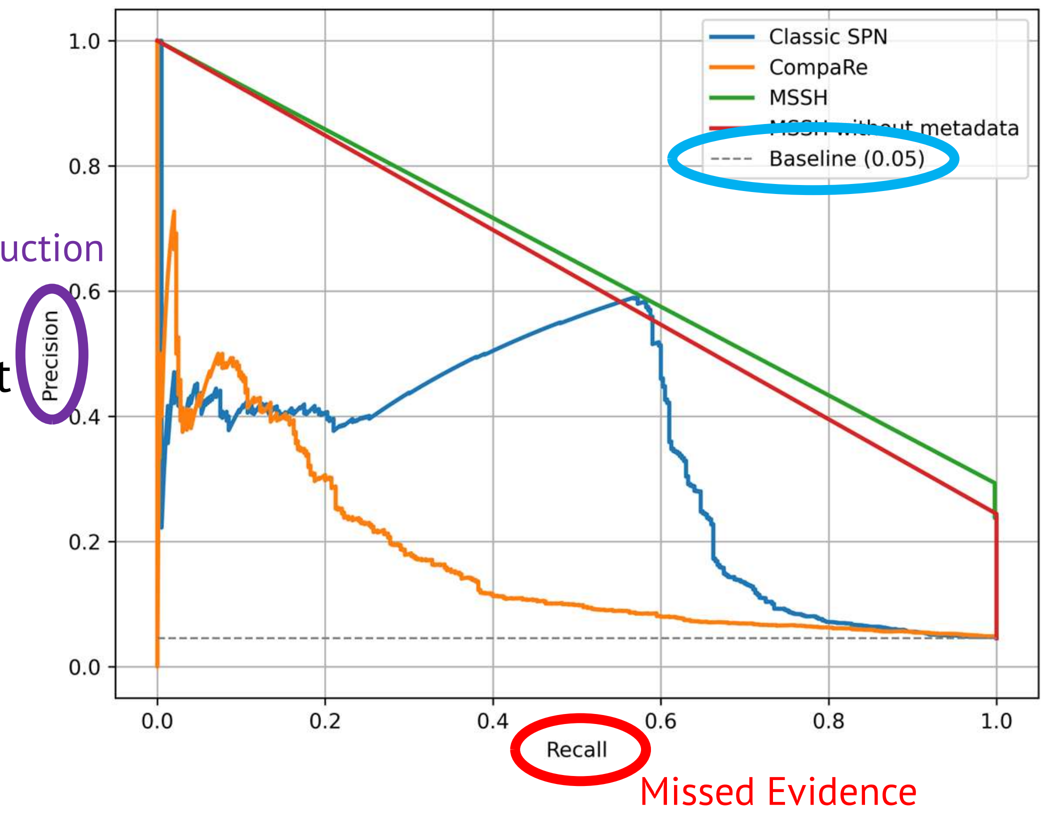
Data Reduction



Missed Evidence

RESULTS | OVERALL PERFORMANCE

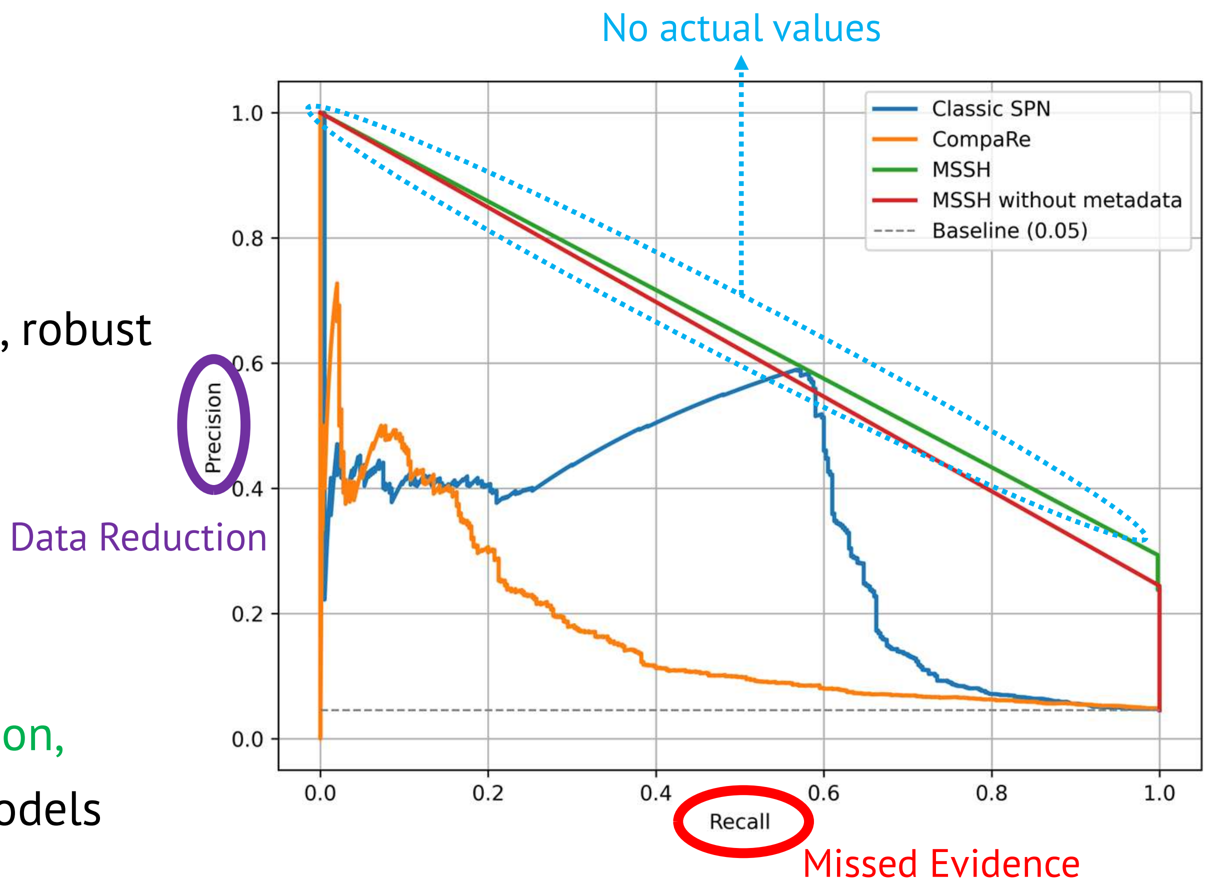
- **SPN:**
  - highest Precision (max. ~0.6) & Recall achievable
  - **Threshold:** adaptable, no trade off, but “sweet spot”
  - **In practice:** half of the evidence missed, unstable



RESULTS | OVERALL PERFORMANCE

- **MSSH:**
  - perfect Recall possible, Precision low (max. ~0.25), robust without metadata
  - **Threshold:** minimal effect
  - **In practice:** no adaptability to case, complete evidence retention,

→ Reliability? Devices vs. Models

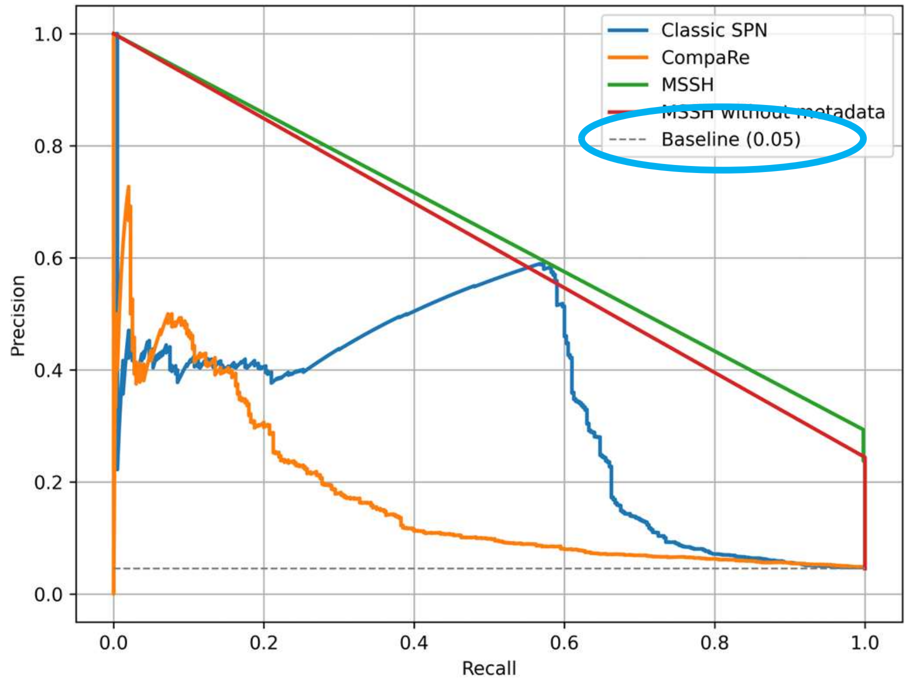


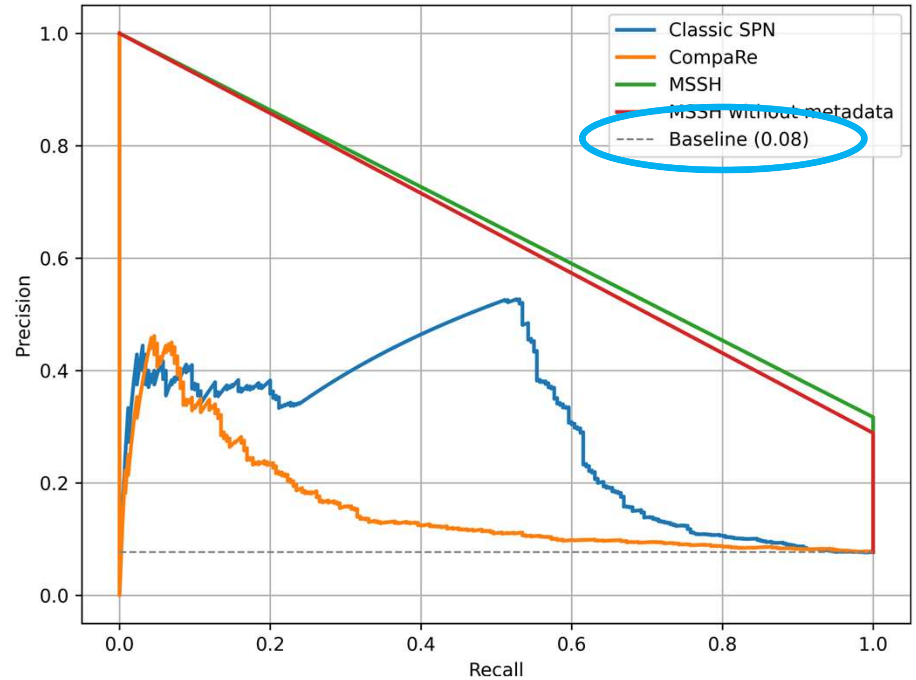
RESULTS | INDIVIDUAL DEVICE IDENTIFICATION

- **SPN:**
  shown to reliably differentiate sensors,
  modern cameras?
- **CompaRe:**
  may differentiate sensors, but not evaluated,
  modern cameras?
- **MSSH:**
  No hardware distinction possible, differentiable by software? → many non-unique *source SD's*

Brand	Model	Device
Apple	iPhone11 iPhone11	C20 C21
Apple	iPhone11ProMax	C22
Apple	iPhoneX	C19
Huawei	P30lite P20pro P20pro P20pro P10	C01 C02 C03 C04 C09
Huawei	PSmart2019	C05
Huawei	PSmart2019	C06
Huawei	P20lite P20lite	C07 C08
Samsung	GalaxyS6	C13
Samsung	GalaxyS9	C14
Samsung	GalaxyS9+	C15
Samsung	GalaxyA70	C16
OnePlus	6T 6	C17 C18
Xiaomi	MiNote10	C10
Xiaomi	RedmiNote8T MiA3	C11 C12

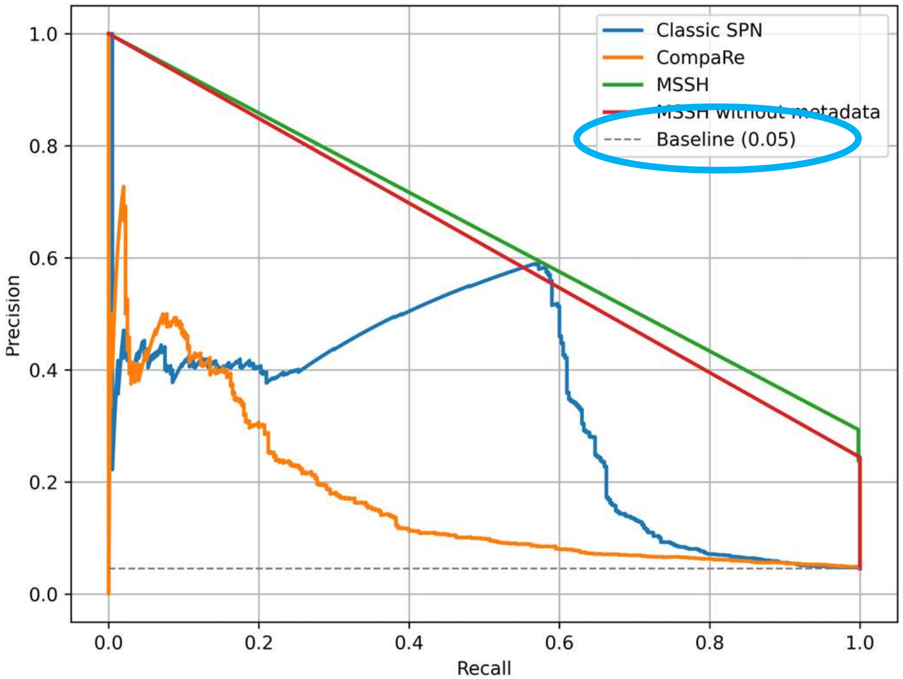
RESULTS | INDIVIDUAL DEVICE IDENTIFICATION

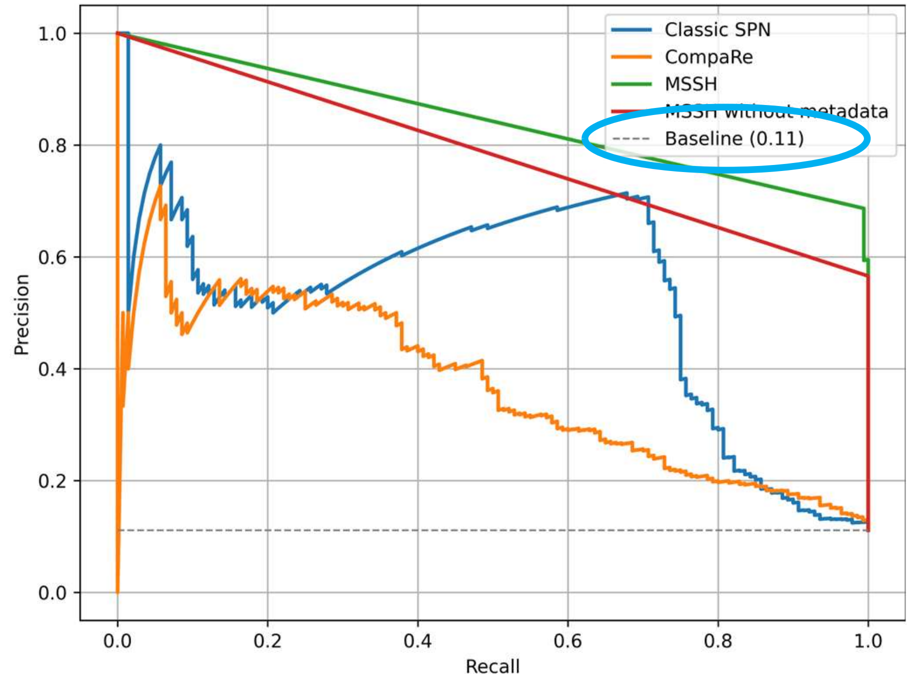




	SPN	CompaRe	MSSH	MSSH w/o MD
AUC comp. data set	0.79	0.69	0.94	0.93
AUC non-unique s-SD	0.75	0.62	0.91	0.90

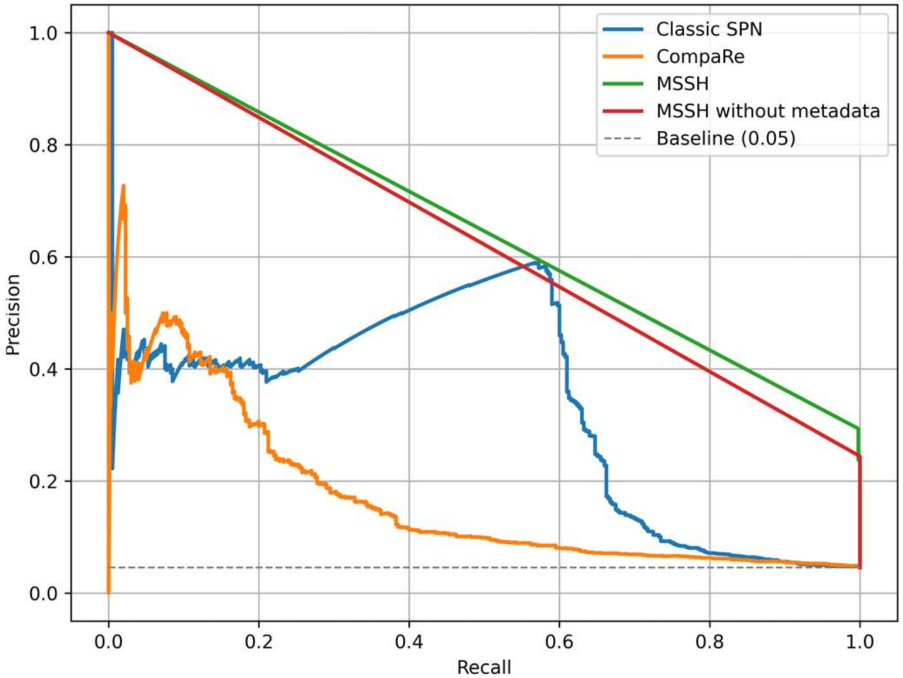
RESULTS | INDIVIDUAL DEVICE IDENTIFICATION

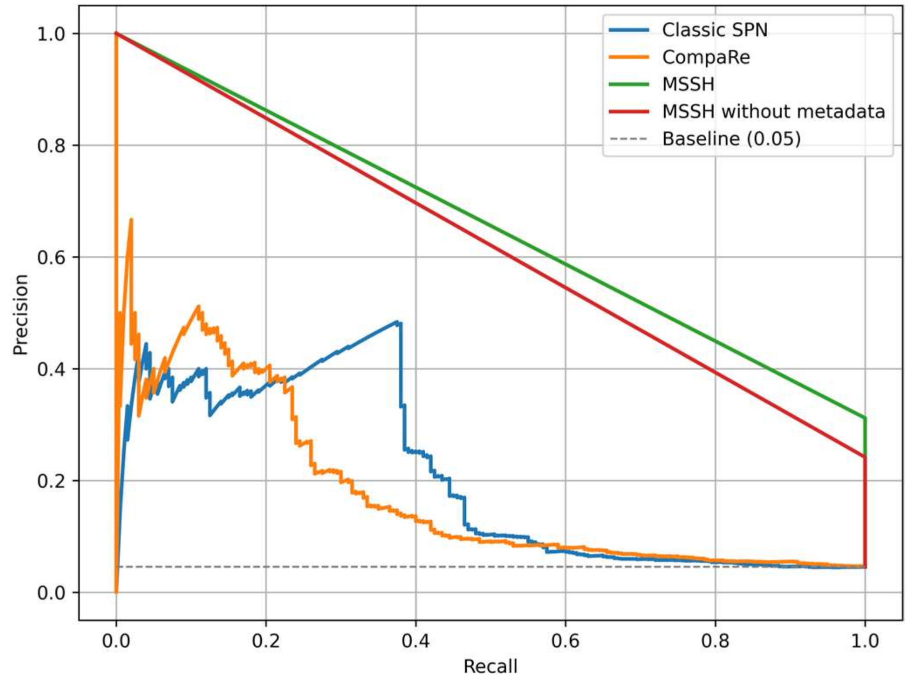




	SPN	CompaRe	MSSH	MSSH w/o MD
AUC comp. data set	0.79	0.69	0.94	0.93
AUC unique s-SD	0.83	0.77	0.97	0.95

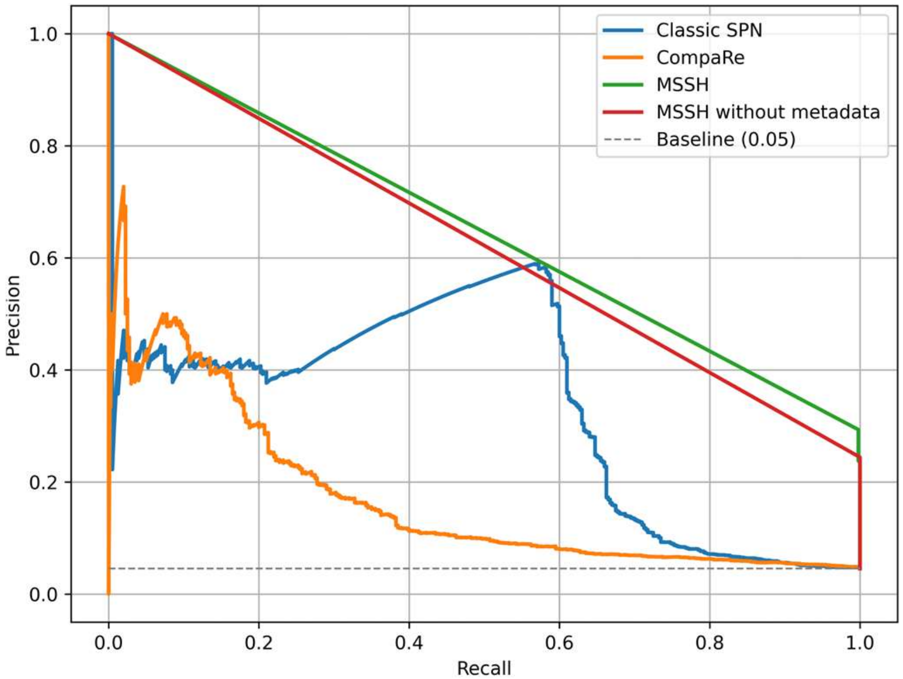
RESULTS | CAPTURING MODES - BOKEH

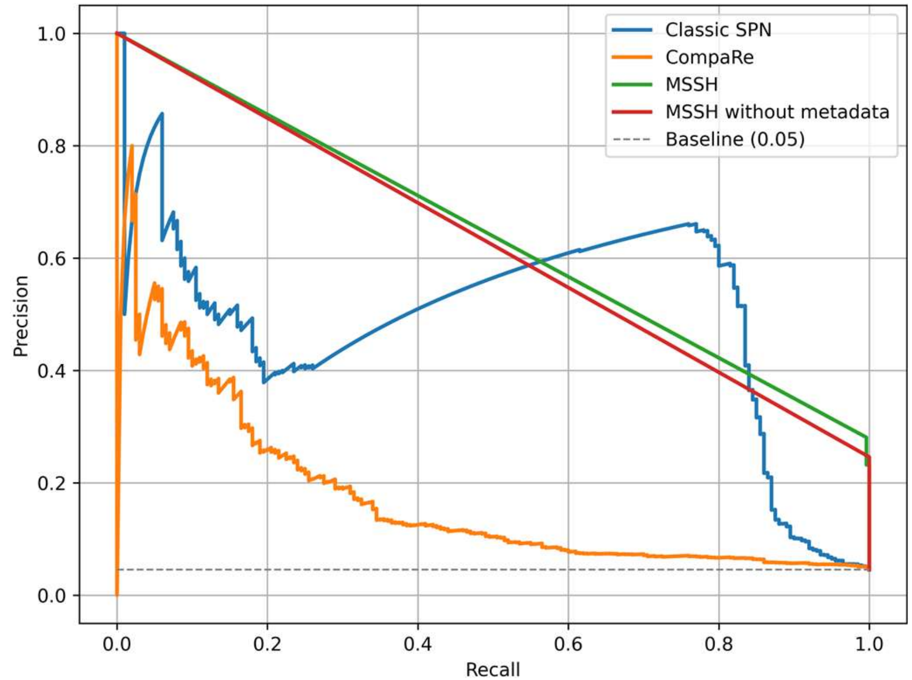




	SPN	CompaRe	MSSH	MSSH w/o MD
AUC comp. data set	0.79	0.69	0.94	0.93
AUC bokeh	0.68	0.69	0.95	0.95

RESULTS | CAPTURING MODES - STANDARD





	SPN	CompaRe	MSSH	MSSH w/o MD
AUC comp. data set	0.79	0.69	0.94	0.93
AUC standard	0.90	0.68	0.94	0.93

EVALUATION | EXTRACTION RUNTIME EFFICIENCY

- All: read images with b Bytes from storage
- MSSH:
  - extract features from byte stream: $O(b)$
  - unifying sets $O(1)$
- SPN | CompaRe:
  - decode image: $O(N)$ (N being the resolution)
  - Noise extraction/filtering: $O(N \log(N))$, e.g. for Wiener Filter
  - Additional signal processing operations
- In practice: reference generation for 22 devices…

15s MSSH, 23min. SPN | CompaRe

EVALUATION | COMPARISONS RUNTIME EFFICIENCY

- All: read image with b Bytes from storage
- MSSH:
  - set operations, e.g. $\cap$, $\setminus$: $O(1)$
- SPN:
  - 2D cross-correlation on original resolution, e.g. $O(N \log(N))$
  - PCE calculation
- CompaRe:
  - i.a.w. SPN, but with constant low resolution
- In practice: ~8000 comparisons in…

6min. MSSH, 51h CompaRe, 63h SPN

Conclusion

- Current SCF research focuses on camera/model verification, with low FPR, overlooking investigative phase needs
- Evaluated: SPN, CompaRe and MSSH
- Critical for investigation pre-processing:
  Only MSSH achieves perfect Recall
- Less relevant for investigations:
  SPN is superior in low-Recall/FPR regions
- Runtime performance: MSSH significantly faster than SPN|CompaRe

FUTURE WORK

- Enhance Precision of MSSH while sustaining Recall ≈ 1.0
- Explore combinations of MSSH with more accurate methods, BUT SPN has problems with same models
- Need for improved or alternative approaches for modern devices
- Validate MSSH robustness on larger, more realistic datasets

THANK YOU!
QUESTIONS?

Samantha Klier
Research Institute CODE
University of the Bundeswehr Munich

Samantha.Klier@unibw.de
https://www.unibw.de/digfor/

Research Institute
Cyber Defence
Universität der Bundeswehr München