Real **W**orld **C**rypto Symposium, 20**24**

Do Not Trust Anybody: ZK Proofs for Image Transformations Tile by Tile on Your Laptop

Pierpaolo Della Monica
Sapienza University of Rome

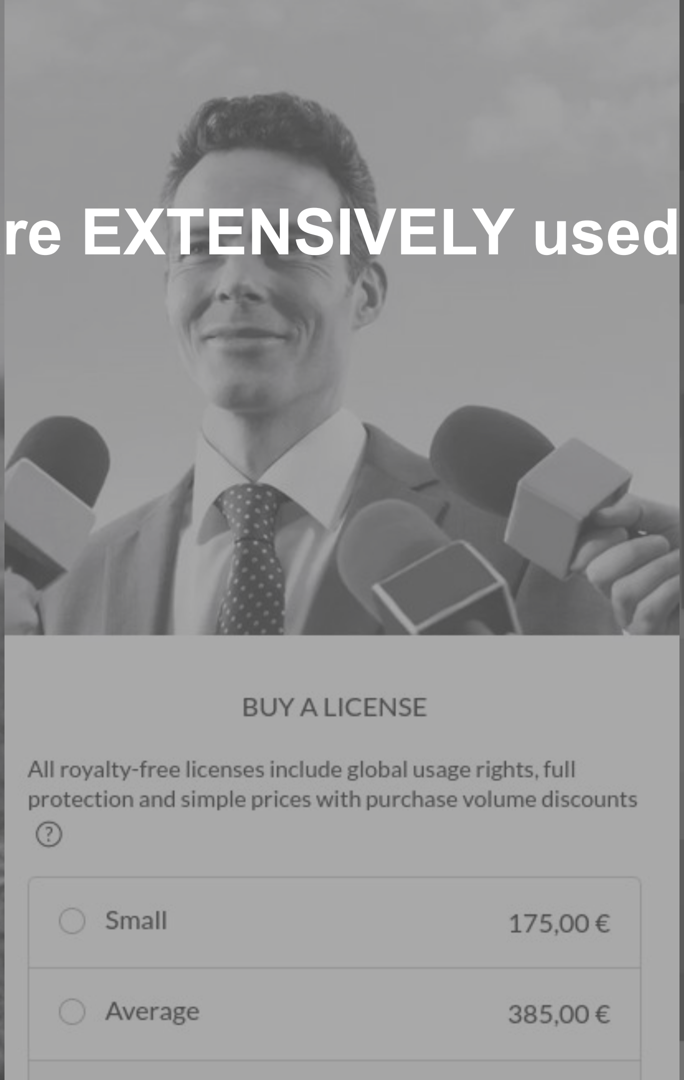
Ivan Visconti
University of Salerno

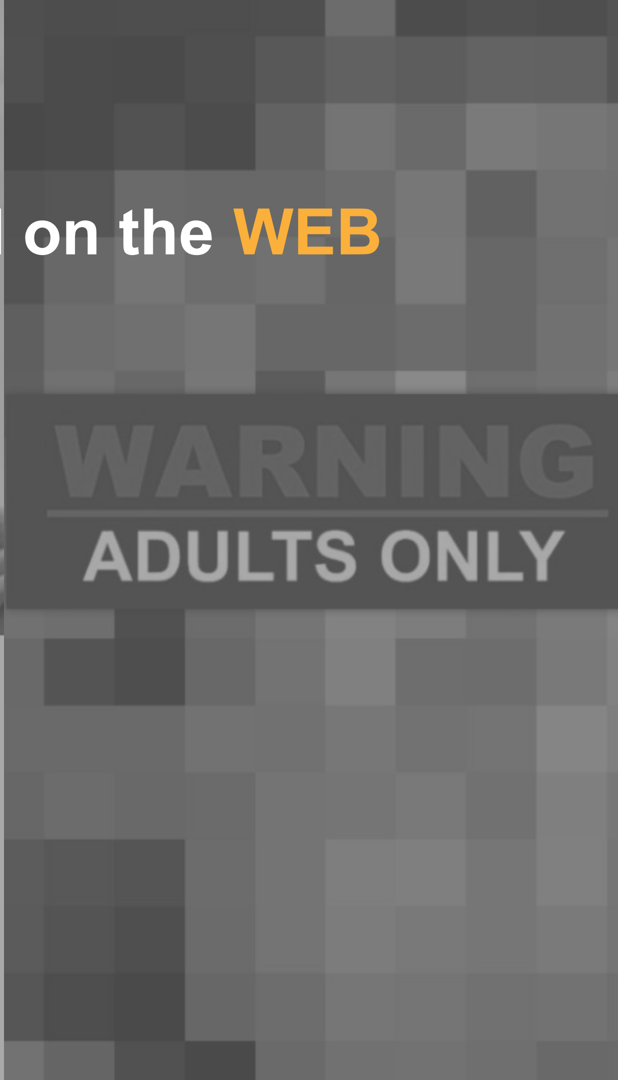
Andrea Vitaletti
Sapienza University of Rome

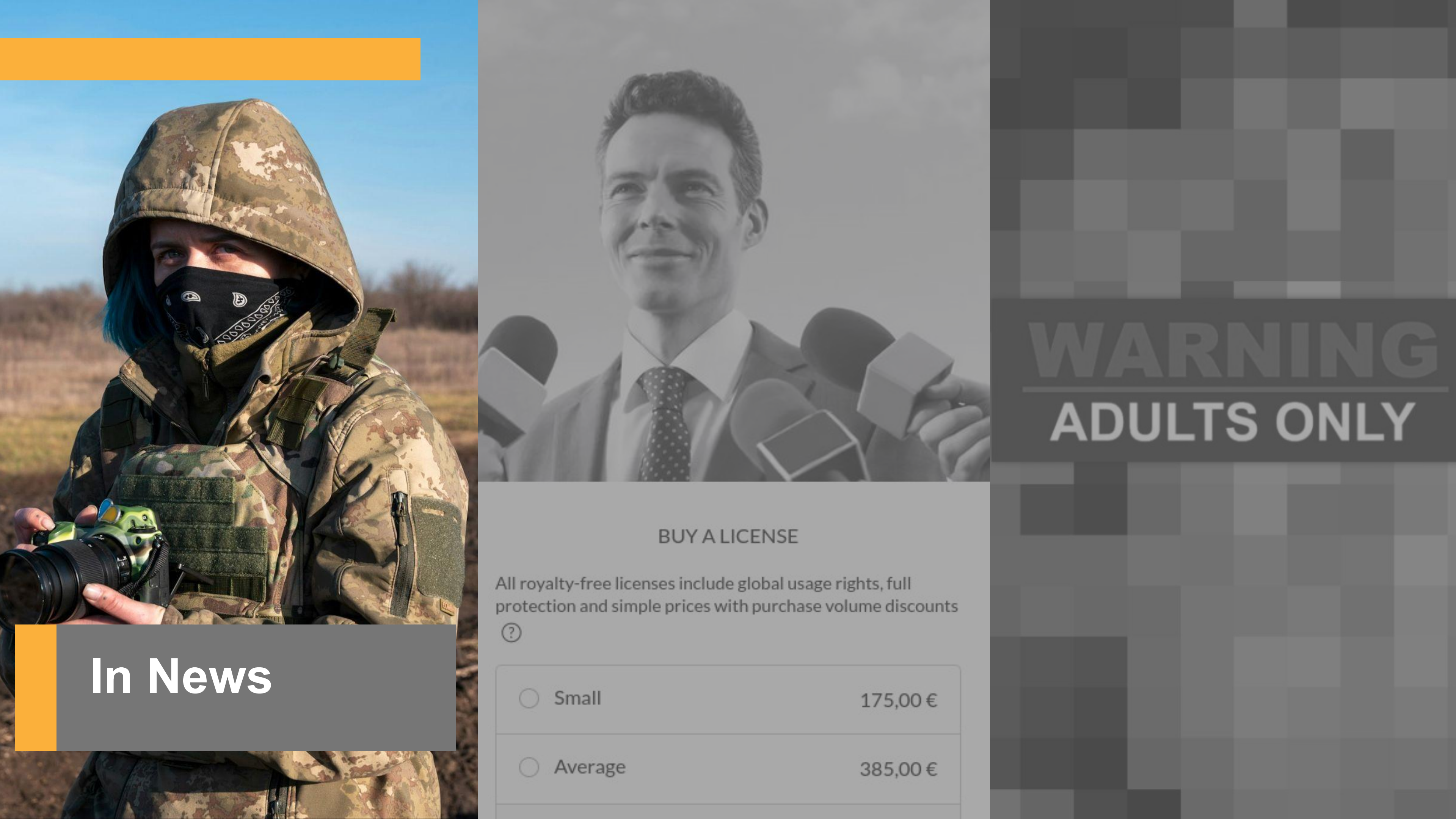
Marco Zecchini
University of Salerno

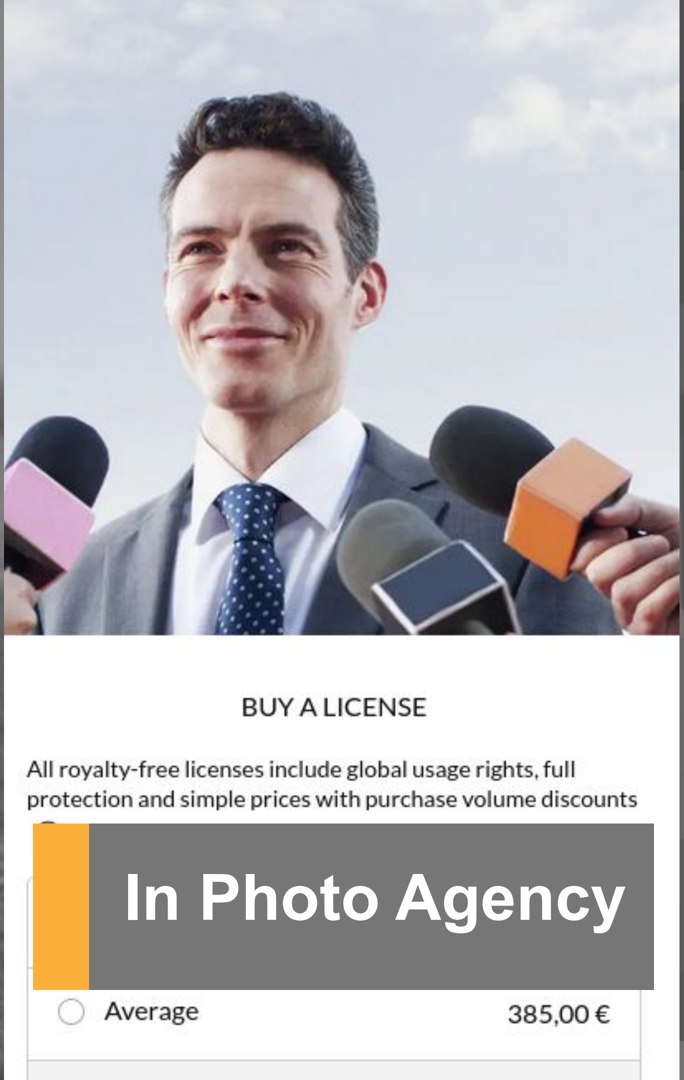
March 27, 2024 Toronto,Canada

IMAGES are **EXTENSIVELY** used on the **WEB**

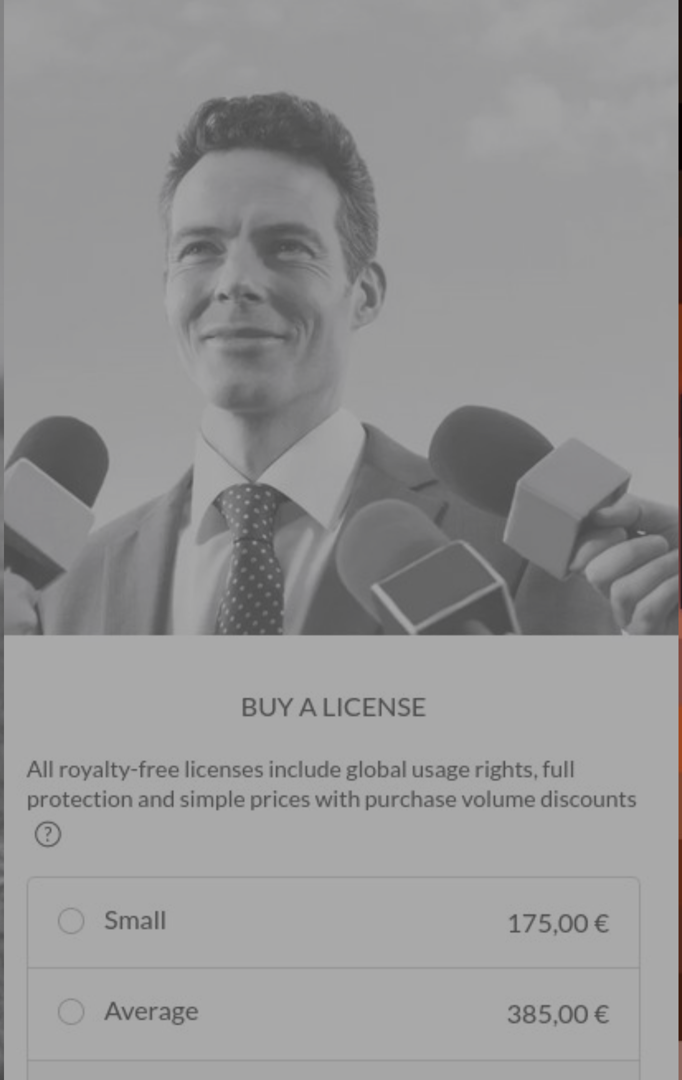





In News
BUY A LICENSE
All royalty-free licenses include global usage rights, full protection and simple prices with purchase volume discounts
Small
175,00 €
Average
385,00 €
WARNING
ADULTS ONLY

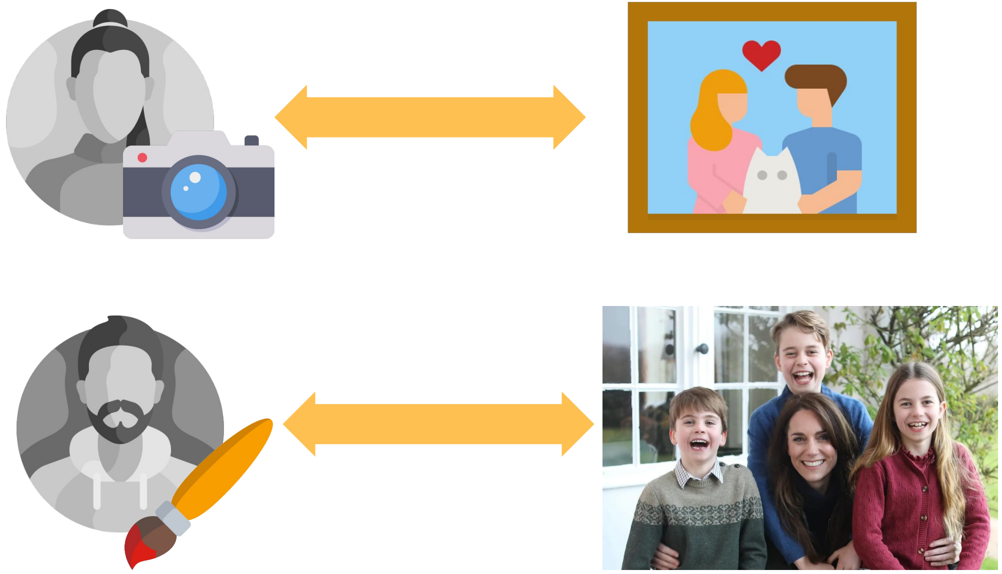


In Photo Agency

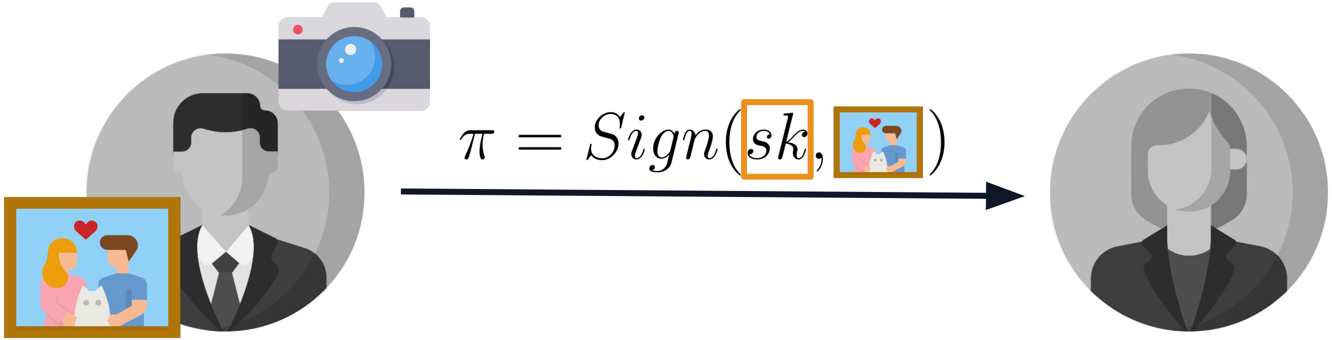


In Adult Sites

AUTHENTICITY
of **images** is important



Assuming that a photo is supposed to be published as it is



C2PA

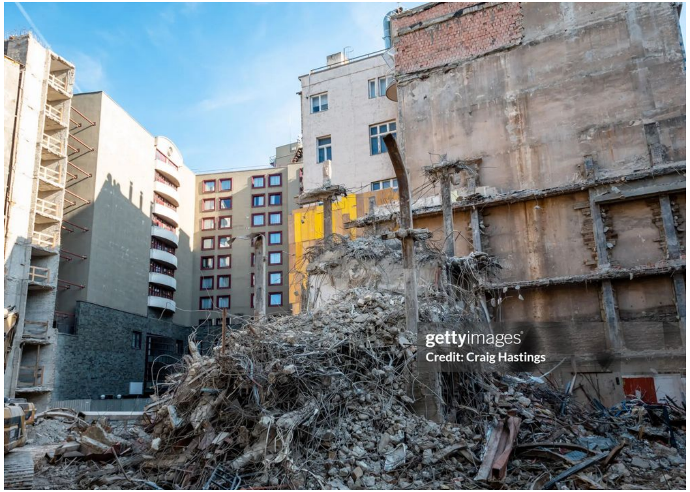
AUTHENTICITY:

Signature schemes suffice

Adobe BBC Google intel Microsoft PUBLICIS GROUPE SONY truepic

But online **images** are edited…

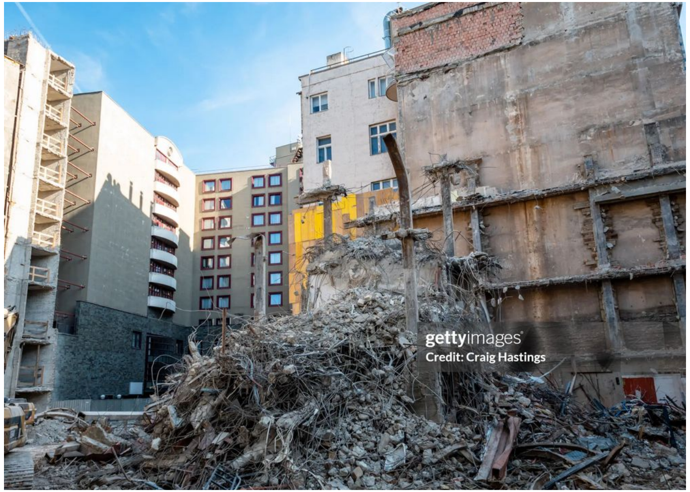
gettyimages®



Show preview of original images with watermarks and smaller dimension

But online **images** are edited…

gettyimages®



preview of original resized images with watermarks

BBC NEWS

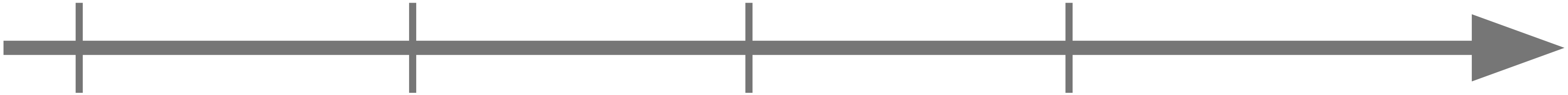
censored images for privacy

But online **images** are edited…

AUTHENTICITY

It is **not** possible to **verify** the C2PA **signature** of an edited **image**

[NT S&P2016]
Image authenticity through cryptography.
Extremely computational intensive
(e.g., tests on 128x128 images)



[NT S&P2016] *A. Naveh and E. Tromer, "**PhotoProof: Cryptographic Image Authentication for Any Set of Permissible Transformations**" - S&P - 2016*

[NT S&P2016]
Image authenticity through cryptography.
Extremely computational intensive
(e.g., tests on 128x128 images)

[KHSS 2022]
Image authenticity adopting digital signatures from cameras but **deviating from C2PA (2021) standard**.

Tests on HD image either missing **confidentiality** (computing on AWS) or relying on **HPC**.

[NT S&P2016] *A. Naveh and E. Tromer, "**PhotoProof: Cryptographic Image Authentication for Any Set of Permissible Transformations**" - S&P - 2016*
[KHSS 2022] *D. Kang, T. Hashimoto, I. Stoica, and Y. Sun, "**ZK-IMG: Attested Images via Zero-Knowledge Proofs to Fight Disinformation.**" - arXiv.org - 2022*

[NT S&P2016]
Image authenticity through cryptography.
Extremely computational intensive
(e.g., tests on 128x128 images)

[KHSS 2022]
Image authenticity adopting digital signatures from cameras but **deviating from C2PA (2021) standard**.

Tests on HD image either missing **confidentiality** (computing on AWS) or relying on **HPC**.

[DB RWC2023]
Image authenticity adopting Lattice Hash and Poseidon Hash for digital signatures.

Test on 30 MP image but **significant requirements on the computing platform.**

[NT S&P2016] *A. Naveh and E. Tromer, "**PhotoProof: Cryptographic Image Authentication for Any Set of Permissible Transformations**" - S&P - 2016*
[KHSS 2022] *D. Kang, T. Hashimoto, I. Stoica, and Y. Sun, "**ZK-IMG: Attested Images via Zero-Knowledge Proofs to Fight Disinformation.**" - arXiv.org - 2022*
[DB RWC2023] *T. Datta and D. Boneh, "**Using zk-proofs to fight disinformation**" - RWC - 2023*

[NT S&P2016]
Image authenticity through cryptography.
Extremely computational intensive
(e.g., tests on 128x128 images)

[KHSS 2022]
Image authenticity adopting digital signatures from cameras but **deviating from C2PA (2021) standard**.

Tests on HD image either missing **confidentiality** (computing on AWS) or relying on **HPC**.

[DB RWC2023]
Image authenticity adopting Lattice Hash and Poseidon Hash for digital signatures.

Test on 30 MP image but **significant requirements on the computing platform.**

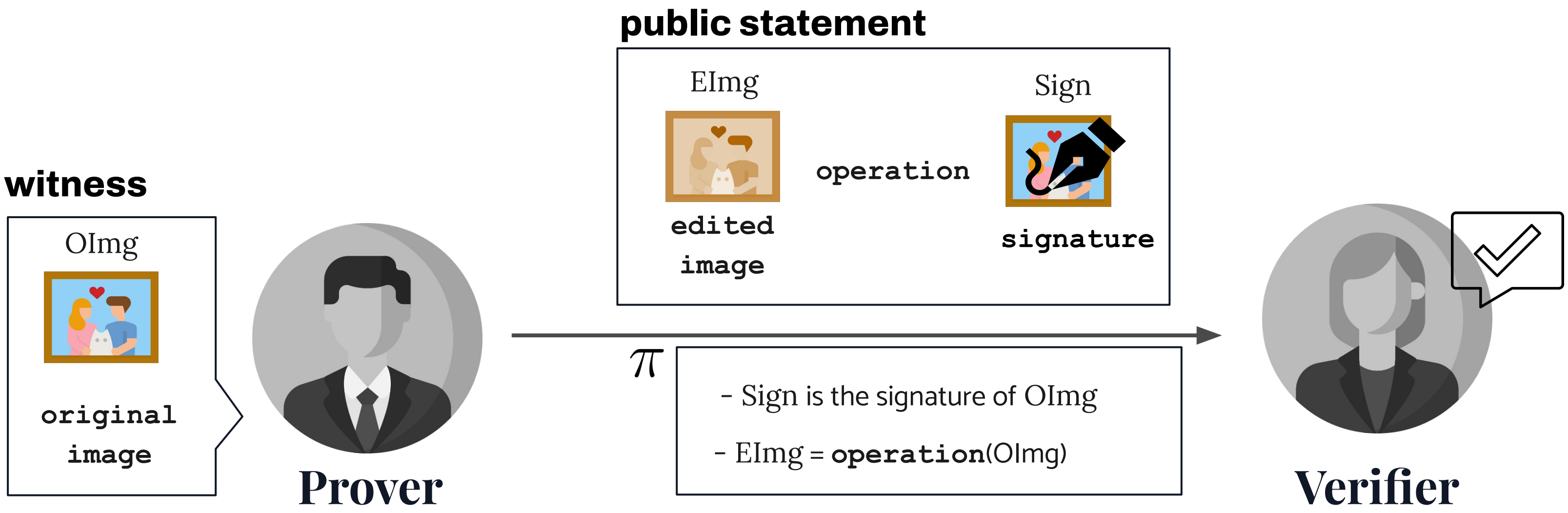
[LHCLCC MIPR2023]
Image authenticity proves correctness of a transformation considering only a small portion of an image.

Experimental results **similar** to ***[KHSS 2022]*** when the entire image is involved.

[NT S&P2016] *A. Naveh and E. Tromer, "**PhotoProof: Cryptographic Image Authentication for Any Set of Permissible Transformations**" - S&P - 2016*
[KHSS 2022] *D. Kang, T. Hashimoto, I. Stoica, and Y. Sun, "**ZK-IMG: Attested Images via Zero-Knowledge Proofs to Fight Disinformation.**" - arXiv.org - 2022*
[DB RWC2023] *T. Datta and D. Boneh, "**Using zk-proofs to fight disinformation**" - RWC - 2023*
[LHCLCC MIPR2023] *K. Li, C. Hsu, M. Chang, F. Liu, S. Chien, and W. Chen, "**Region-aware photo assurance system for image authentication**" - MIPR - 2023*

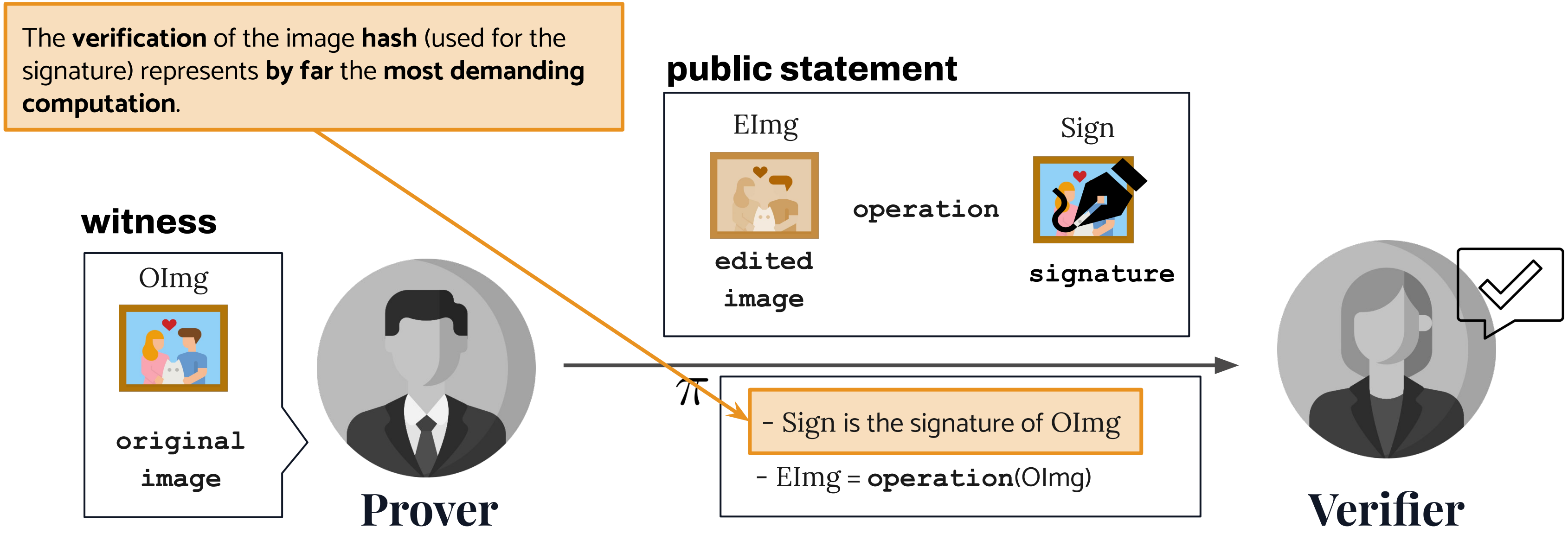
AUTHENTICITY:

a **ZK-SNARK** to **link** two images, an **original (and secret) image** and the corresponding **edited (and known) image**

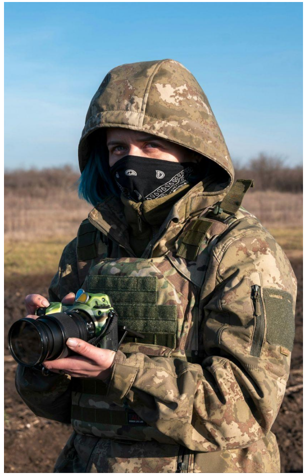


AUTHENTICITY:

a **ZK-SNARK** to **link** two images, an **original (and secret) image** and the corresponding **edited (and known) image**.



DO WE ALWAYS NEED **SUCCINCTNESS** ?



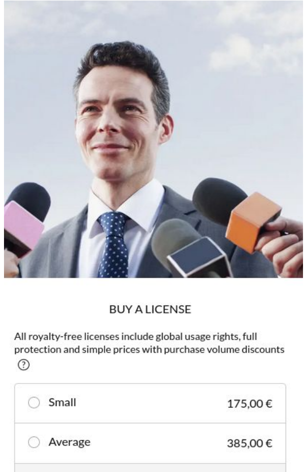
In news websites

DO WE ALWAYS NEED **SUCCINCTNESS** ?

In news websites

DO WE ALWAYS NEED **SUCCINCTNESS** ?

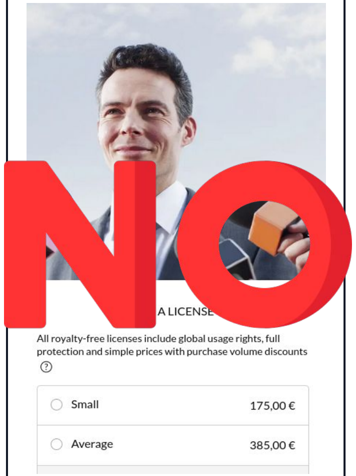
In news websites



In photo agency

DO WE ALWAYS NEED **SUCCINCTNESS** ?

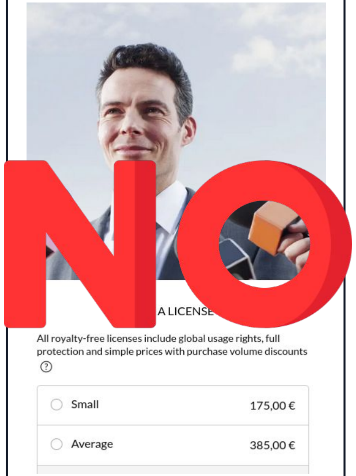
In news websites



In photo agency

DO WE ALWAYS NEED **SUCCINCTNESS** ?

In news websites

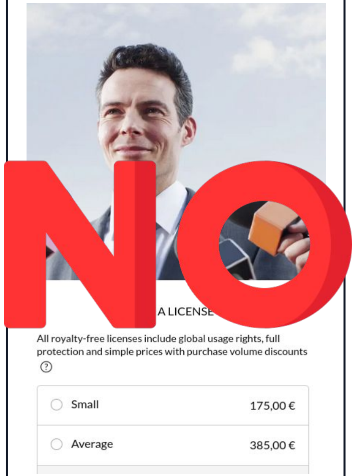


In photo agency

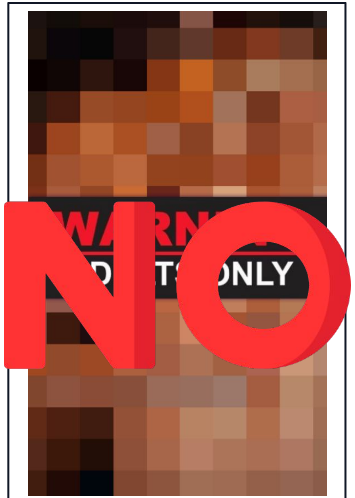
In adult sites

DO WE ALWAYS NEED **SUCCINCTNESS** ?

In news websites



In photo agency



In adult sites

SUCCINCTNESS OF THE PROOF IS OFTEN AN **OVERKILL** IN SEVERAL SCENARIOS AND A SUCCINCT FRAUD PROOF CAN BE GOOD ENOUGH

Our Results

We propose a system to prove **image authenticity guaranteeing**:

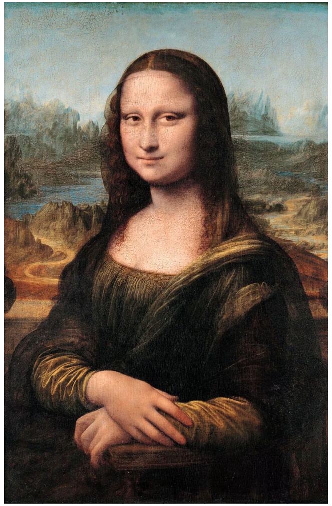
1. **Low memory consumption** for the **prover** (no HPC, your laptop is just fine)

2. **Succinct Fraud Proofs** fast verification for usability (e.g., browsers) and compactness for blockchains

3. **Confidentiality** of the original **image** (no cloud infr.) and **authenticity** of the transformed image defined and proved (starting with [NT S&P2016])

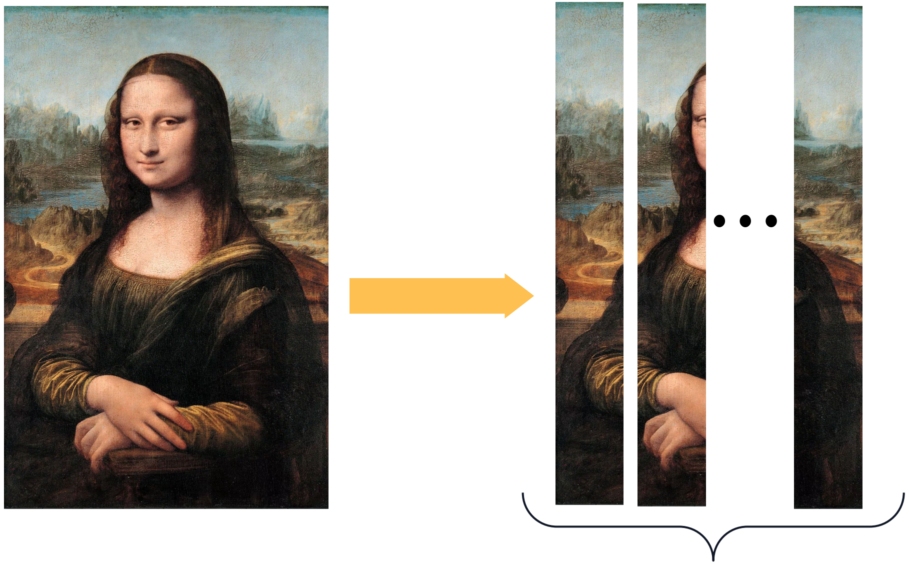
4. **Compliancy** with **C2PA** standard (at an additional, still affordable, cost for proof computation and size)

Image **Tiling**



*For large images proving knowledge of a pre-image of the hash is the real **bottleneck***

Image **Tiling**

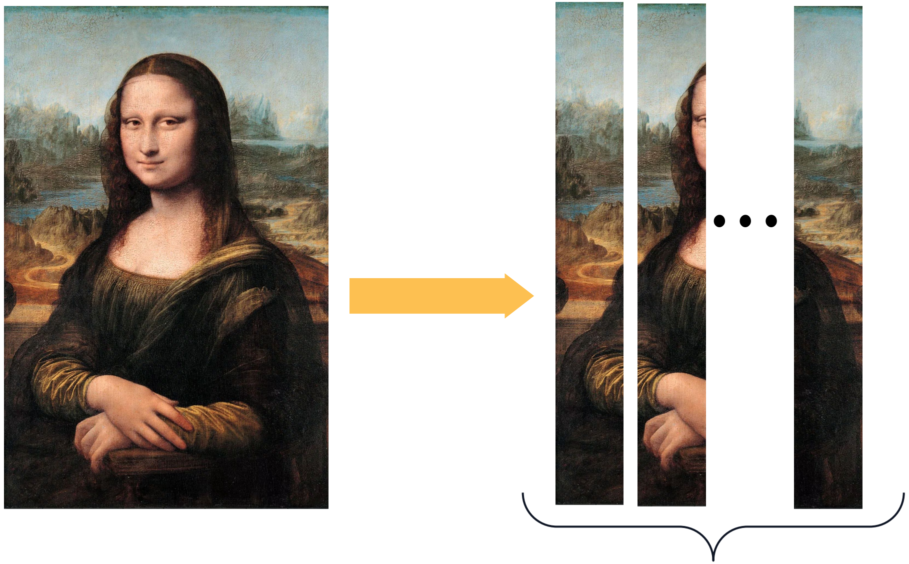


For large images proving knowledge of a pre-image of the hash is the real ***bottleneck***

Each tile has a reduced dimension and it is possible to split the computational effort

This methodology consists of **splitting** the image into several smaller **tiles.**
For each tile, a **ZKP** can be defined, enabling **hashing** for a **shorter witness** and producing multiple hashes that represent different subimages.

Image **Tiling**



For large images proving knowledge of a pre-image of the hash is the real ***bottleneck***

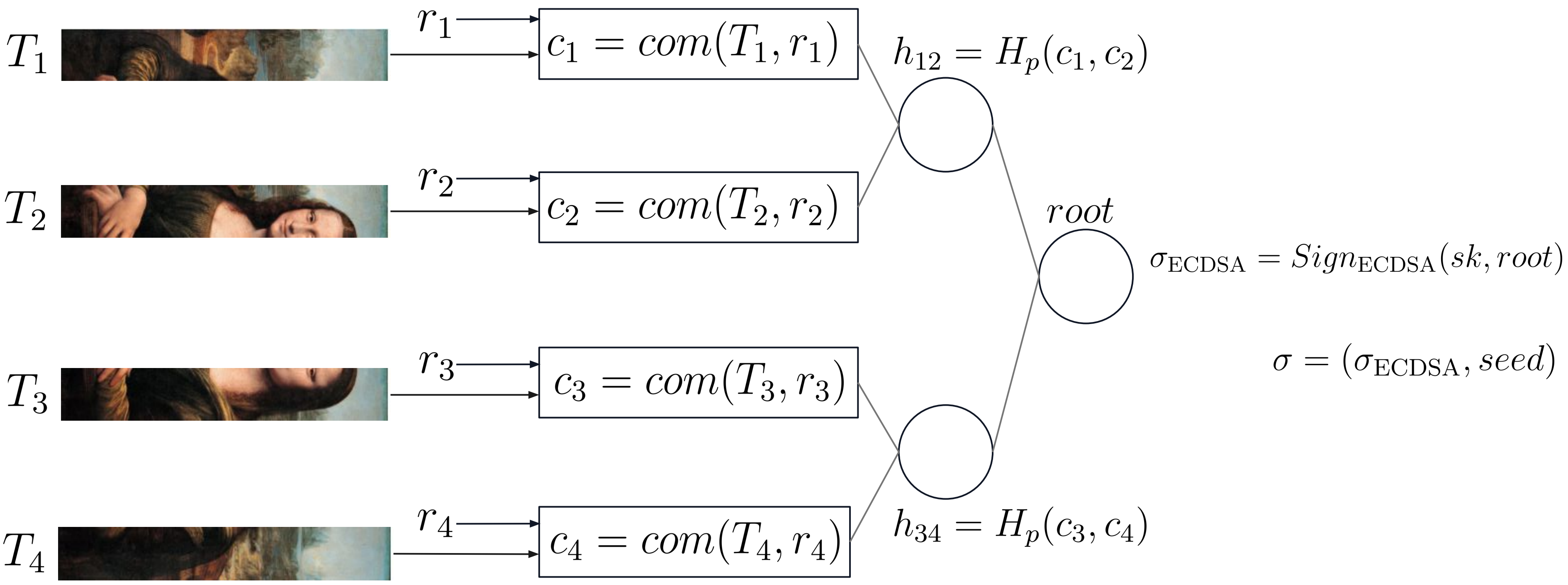
Each tile has a reduced dimension and it is possible to split the computational effort

This methodology consists of **splitting** the image into several smaller **tiles.**
For each tile, a **ZKP** can be defined, enabling **hashing** for a **shorter witness** and producing multiple hashes that represent different subimages.

It is important that the transformation of the full image can be computed working locally tile by tile. Many natural transformations follow this approach.

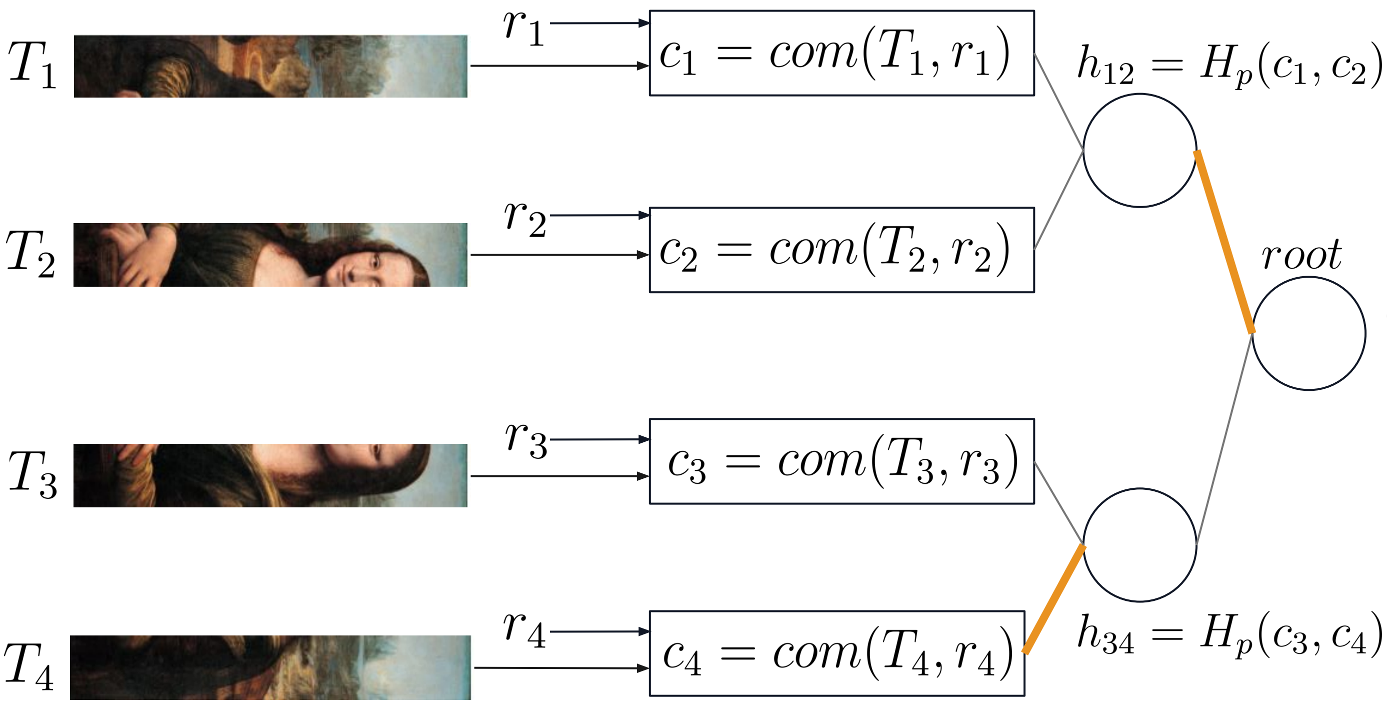
The **Signature** Scheme

$$r_i = PRF(seed, i)_{i \in \{1,2,3,4\}}$$



The **Signature** Scheme

$r_i = PRF(seed, i)_{i \in \{1,2,3,4\}}$



$\sigma_{\text{ECDSA}} = Sign_{\text{ECDSA}}(sk, root)$

$\sigma = (\sigma_{\text{ECDSA}}, seed)$

$B_3 := [c_4, h_{12}]$

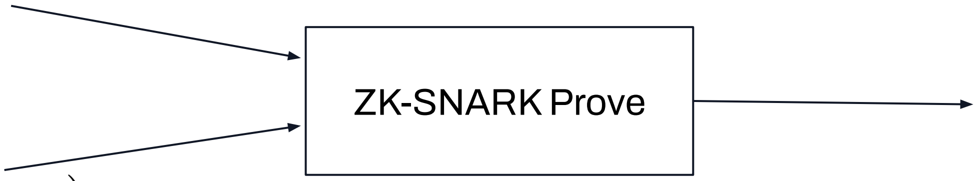
Represents the Merkle Branch to verify c_3

Local **Transformation**

Proof generation

$\textbf{For } i \in 1, ..., 4 \textbf{ then}$

$$x_i := (\text{Resize}(\cdot), \hat{T}_i, c_i)$$



$$\pi_i$$

$$w_i := (T_i, r_i)$$

$$\Big(c_i = com(T_i, r_i) \wedge \hat{T}_i = \text{Resize}(T_i)\Big)$$

Proof generation

For $i \in 1, ..., 4$ **then**

$$x_i := (\text{Resize}(\cdot), \hat{T}_i, c_i)$$

$$w_i := (T_i, r_i)$$

ZK-SNARK Prove → π_i

$$\left(c_i = com(T_i, r_i) \wedge \hat{T}_i = \text{Resize}(T_i)\right)$$

$$\Pi = \boxed{root} + \boxed{B_1, \ldots, B_4} + \boxed{\pi_1, \ldots, \pi_4}$$

Proof verification and Fraud proof

$$\Pi = \boxed{root} + \boxed{B_1, \ldots, B_4} + \boxed{\pi_1, \ldots, \pi_4}$$

For $i \in 1, ..., 4$ **then**

$$VerifyProof(vk_i, x_i, \pi_i)$$

If not correct, provide π_i as a **FRAUD PROOF**

Proof verification and Fraud proof

$$\Pi = \boxed{root} + \boxed{B_1, \ldots, B_4} + \boxed{\pi_1, \ldots, \pi_4}$$

For $i \in 1, ..., 4$ **then**

$$VerifyLeaf(c_i, root, B_i)$$

If not correct, provide B_i as a **FRAUD PROOF**

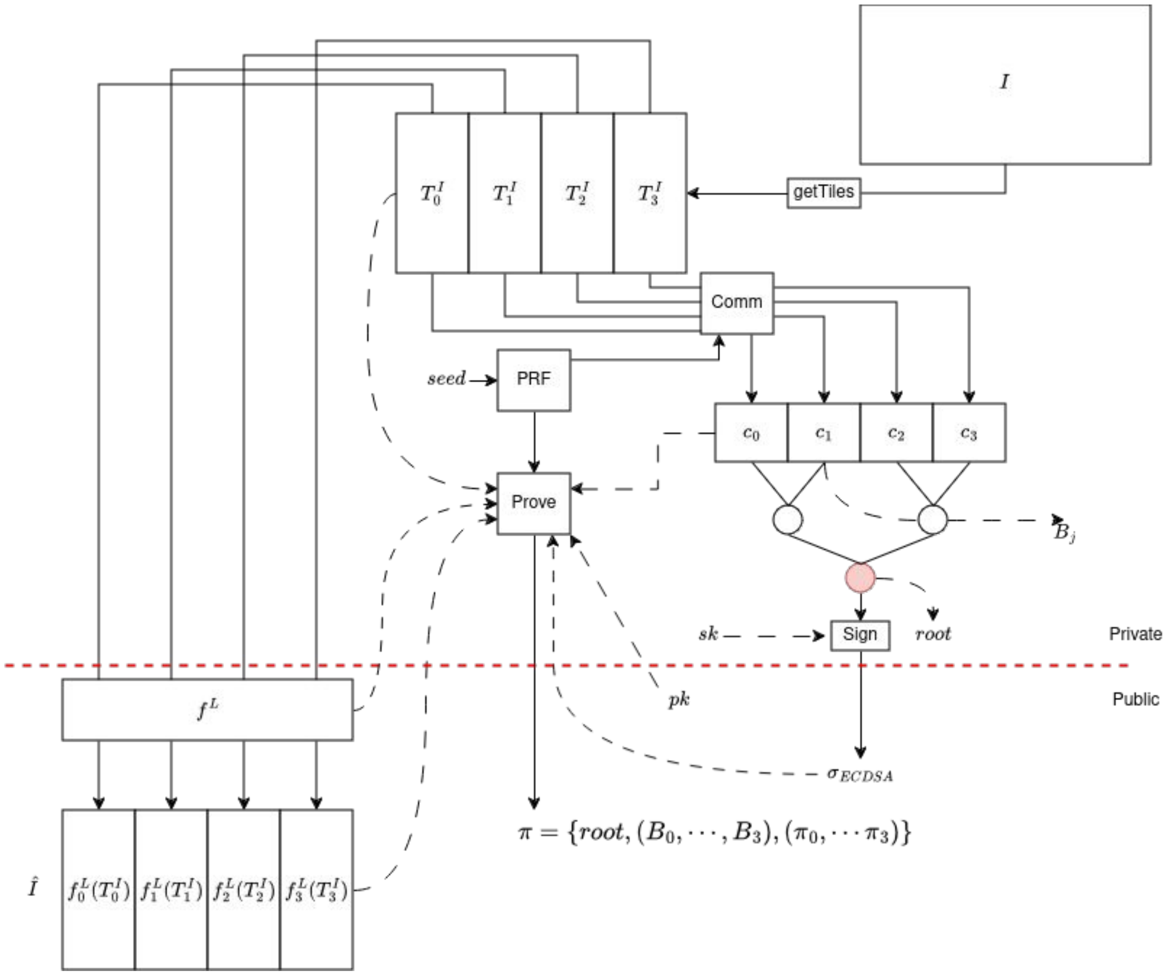
Proof verification and Fraud proof

$$\Pi = \boxed{root} + \boxed{B_1, \ldots, B_4} + \boxed{\pi_1, \ldots, \pi_4}$$

$$VerifySign_{\mathrm{ECDSA}}(pk, \sigma_{ECDSA}, root)$$

If not correct provide $\sigma_{ECDSA}, root$ as a **FRAUD PROOF**

Architecture overview for an **image** divided into **4 tiles**



Experiments

Our approach is **generic** and can be instantiated with different ZK proof techniques.

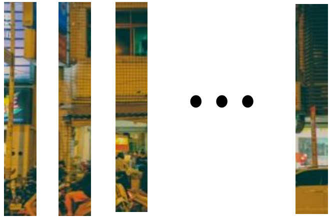
The following experiments were conducted using **Groth16** as ZK-SNARK instantiation, facilitating a **comparison** with the contemporary **state-of-the-art performance** and outcomes.

Experiments

FEASIBILITY ON **30MP IMAGE** (6000×4000 PIXELS)

We run the test on Intel i7@1.8 GHz, 8 cores and 16 GB of RAM

We divided the image in **131 tiles** of **513x361 pixels**



- Tile Proof generation:
  17.25 sec and 4.2 GB of RAM.
- Image Proof generation:
  2260 sec (~**38 min**) and **4.2 GB** of RAM.
- Verification time:
  65 sec (**0.5 sec per Tile**) and **<150MB** of RAM
- Proof size:
  800 bytes per tile (104.8 KB in total)

* **Setup** operations must be performed **only once** for each fixed **dimension** and required ~90 min
Fraud Proof requires a maximum of 0.5 sec and has a maximum size of 800 bytes.

Experiments

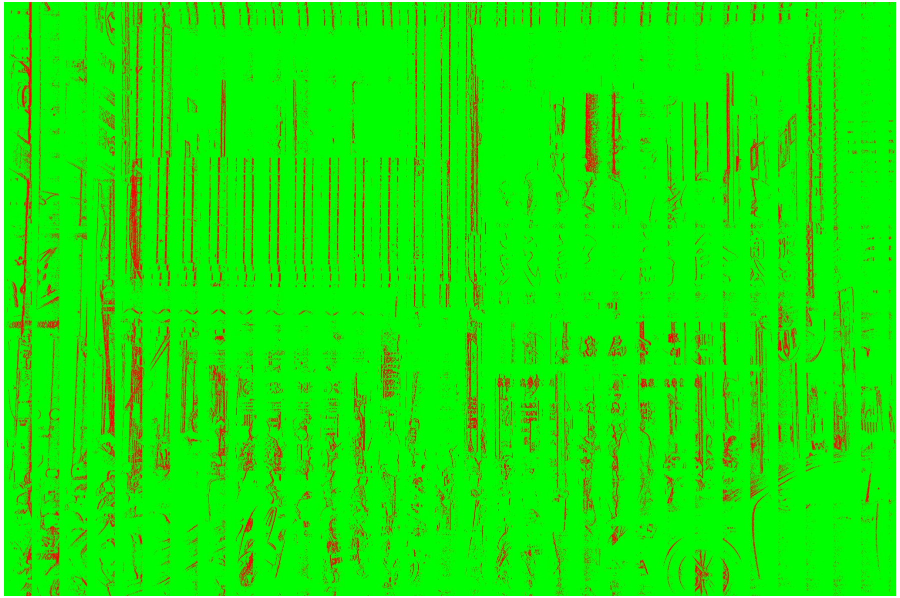
PERFORMANCE COMPARISON with *[KHSS 2022]*

Resize from HD to SD image

	Prov	Ver (FPVer)	Proof Size (FP Size)	Resources	
ZK-IMG [KHSS22]	328.2s	5.3 ms (N.A.)	3.04 KB (N.A)	70.7 GB on Intel Xeon 8375C with 64 vCPU	☹
This paper	86.25 s	2.5 s (0.5 s)	4.4 KB (800 bytes)	4.2 GB on Intel Core i7-8565U with 16 vCPU	☺

Experiments

ON THE **QUALITY** OF LOCAL RESIZING



A filter that **highlights pixels** with a **variance** of at least **5** in any of the **RGB** channels.

7% of pixels in total

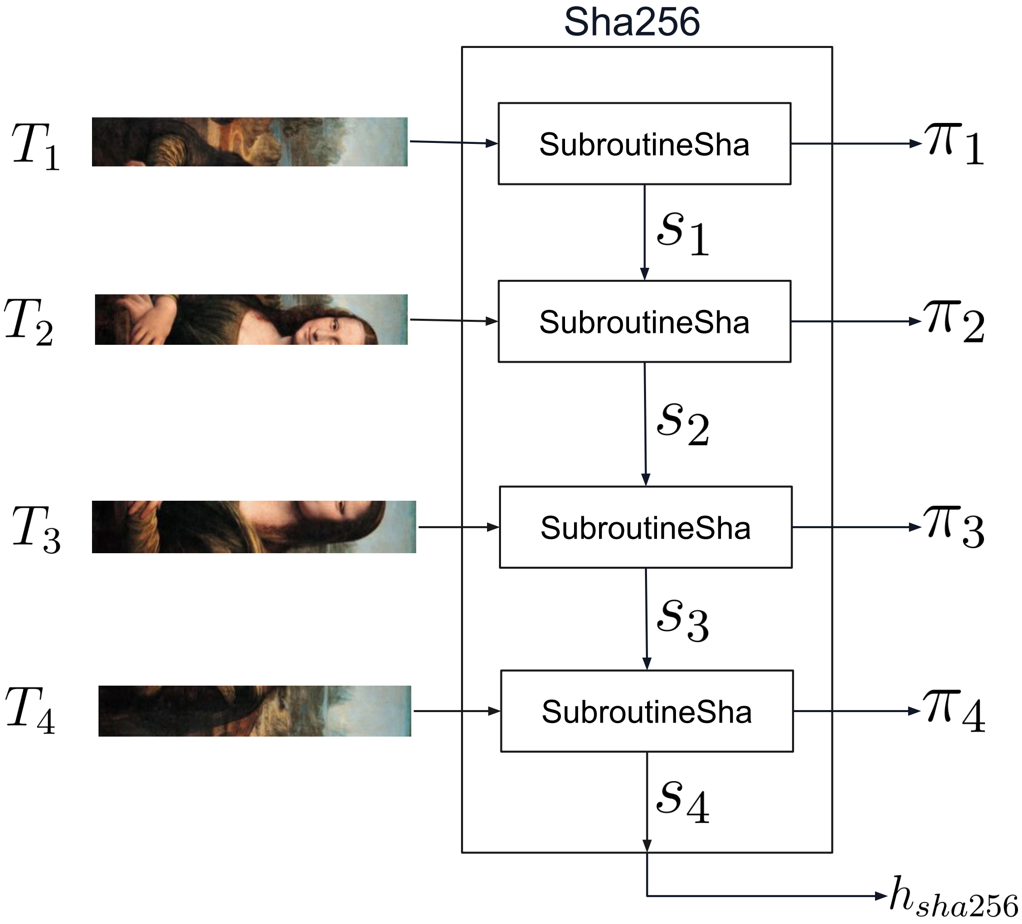
Experiments

ON THE **QUALITY** OF **LOCAL** RESIZING

Resize on the **Full Image**

Resize and merge on the **Tiles**

Compliance with C2PA



$$\Big\{(\mathrm{Resize}(\cdot), \hat{T}_2, c_2, z_2, z_1), (T_2, s_2, s_1, r_1, r_2, r_3) : \\ s_2 = \mathrm{subroutineSha256}(T_2, s_1) \wedge \\ c_2 = \mathrm{Comm}(T_2, r_3) \wedge \\ z_1 = \mathrm{Comm}(s_1, r_1) \wedge \\ z_2 = \mathrm{Comm}(s_2, r_2) \wedge \\ \hat{T}_2 = \mathrm{Resize}(T_2) \Big\}$$

For an HD image using only 4 GB,
the **proof generation time** is
3088 sec (51 min),
with a proof **size** of **280 KB**.

The **verification time** is 178.5 sec,
the **fraud proof** verification time is **0.5 sec,**
with a fraud proof **size** of **800 B**.

THANKS!

This presentation includes icons from Flaticon