

PawPrint v2024b

A Production-Grade Traditional Image Watermark

Bill Huang, Jiaxin Guan, and Cindy Han

16.09.2026, IEEE ICIP x JPEG Trust



Presentation Contents

- PawPrint v2024b
 - Image Watermarks & Modern ML
 - PawPrint v2024b
 - Six Innovative Ingredients
 - Future Outlook
- Watermarking for Content Provenance
 - JPEG Trust and C2PA
 - Problem Parameters
 - JPEG Trust Level 3 Proof-of-Concept
 - New Technologies
- Q&A



PawPrint v2024b

Image Watermarks & Modern ML

- first golden age in the late 1990s
 - DCT / DFT / DWT
 - spread-spectrum
 - key-based techniques
- new golden age in the 2020s
 - NN watermarks (neural networks + deep learning)
 - HiDDeN (Li Fei-Fei et al.) in 2018
 - Big Tech: SynthID (Google), TrustMark (Adobe), PixelSeal (Meta)
 - huge jump in capability



PawPrint v2024b

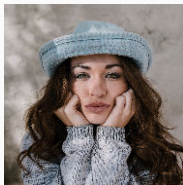
- does not use any ML
- developed by our team at Trufo
- very good at what it is designed to do
 - imperceptible (45+ PSNR, etc.)
 - durable against practical (benign) attacks
 - SotA NN watermarks have only recently caught up to PawPrint v2024b



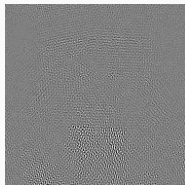
PawPrint v2024b



original



watermarked



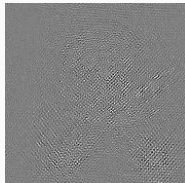
residual (20 $\times$)



original



watermarked



residual (20 $\times$)



PawPrint v2024b

Metric	PawPrint v2024b	PixelSeal+BCH	DWT-DCT-SVD
Payload	100 bits	100 bits	64 bits
PSNR	46.79 dB	46.12 dB	39.34 dB
wPSNR	42.08 dB	41.42 dB	34.62 dB
SSIM	0.9965	0.9967	0.9999
Encode time	1.15 s	0.02 s	0.33 s
Decode time	0.086 s	0.017 s	0.220 s
Clean BER	0.0000	0.0000	0.0000
Attack BER	0.0178	0.0102	0.1967

Aggregate statistics on the JPEG Trust benchmark (100 camera-captured + 100 synthetic images).



PawPrint v2024b

Attack	PawPrint v2024b	PixelSeal+BCH	DWT-DCT-SVD
Gaussian noise	0.0260	0.0703	0.1485
Speckle noise	0.0172	0.0434	0.0477
Blur	0.0000	0.0000	0.0000
Brightness	0.0100	0.0083	0.3292
Sharpness	0.0000	0.0000	0.0021
Gamma	0.0029	0.0008	0.3361
Histogram eq.	0.0878	0.0020	0.4554
Median filter	0.0011	0.0000	0.0043
Rotation	0.0356	0.0006	0.4905

Attack	PawPrint v2024b	PixelSeal+BCH	DWT-DCT-SVD
Crop 10%	0.0200	0.0010	0.4831
Scale 0.5	0.0000	0.0000	0.0000
Flip H	0.0000	0.0000	0.4774
Resize 50%	0.0000	0.0000	0.0000
JPEG q=50	0.0004	0.0000	0.2311
JPEG 2000	0.0037	0.0045	0.0409
JPEG XL	0.0031	0.0000	0.1237
JPEG AI	0.0353	0.0094	0.3531
AI remove	0.0780	0.0422	0.0161

BER by attack (lower is better; bold = best of the three).



Six Innovative Ingredients

- six techniques were developed to make PawPrint v2024b possible
 - signal separation transform
 - alignment pattern and message payload pattern
 - zero-ing out signals (vs. adding large signals)
 - symbols instead of binary 1/0
 - scaling based on local perceptive capacity
 - iterative descent for encoding



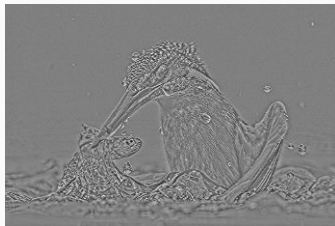
- function $s(I)$ is applied to reduce dynamic range
 - largely preserves watermark features
 - $s(I) = I - r(I)$, where $r(I)$ is a local-neighborhood mean weighted by inverse variance



kingfisher



Y channel



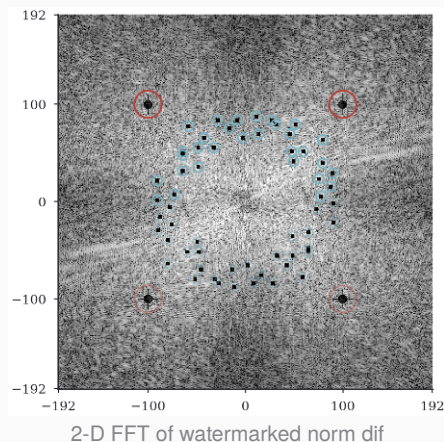
norm dif (20 $\times$)



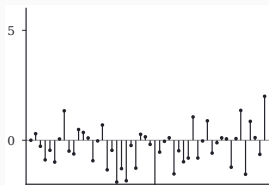
Six Innovative Ingredients

Pattern Design

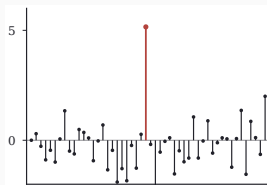
- two alignment points to determine geometric alignment
- 30 message points to carry watermark payload



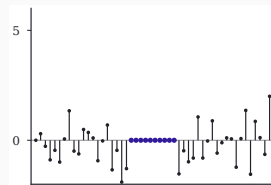
- pushing pattern coefficients to zero



$$x_i \sim \mathcal{N}(0, 1)$$



typical: add +5



ours: set to 0

- two nice properties:
 - computed t -stat is high since the stdev is lower
 - the watermark residual cannot be transplanted onto another image



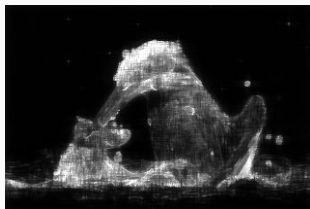
- each symbol carries 5 bits
- improved imperceptibility
 - each of the 32 possibilities zeroes out 1/32 of the symbol region
 - of the usable coefficients, only 1/32 are changed
- fits well with $RS_{2^5}[30, 20]$ code



Six Innovative Ingredients

Local Scaling

- PSNR is not the most accurate measure
 - same pattern is more visible in flat / smooth areas
 - same pattern is more visible on facial features
- we scale the pattern based on the “local perceptive capacity”
 - based on local variance / local gradient



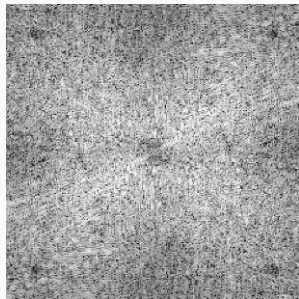
kingfisher $\rightarrow$ norm amp $a(I)$



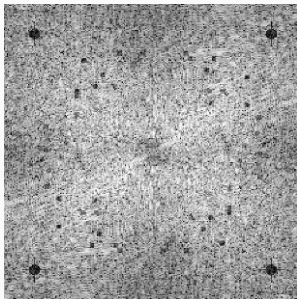
Six Innovative Ingredients

Iterative Descent

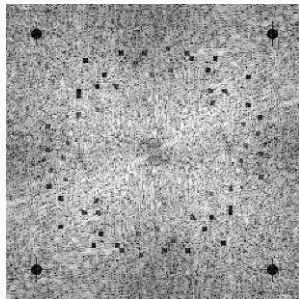
- setting the target coefficients to zero is not so easy
 - no clean analytical solution
- use iterative descent to improve the pattern



iteration 1



iteration 2



iteration 3



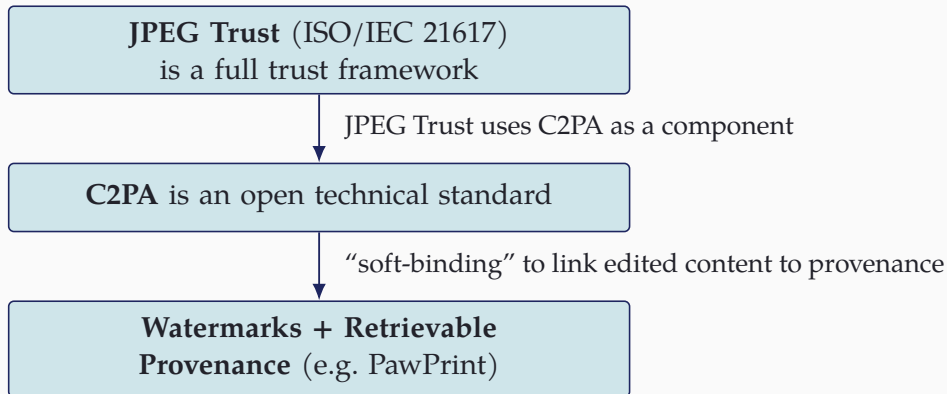
Future Outlook

- NN watermarks have a higher ceiling
 - faster if you have a GPU
- any use for “traditional” watermarks like PawPrint v2024b?
 - occupies a different feature space → can layer traditional + NN more safely
 - we use NN watermarks in production now



Watermarking for Content Provenance

JPEG Trust and C2PA



Problem Parameters

- watermark criteria in content provenance



- historically, watermarks for **private** usage
 - tracing leaks to defend against piracy & misuse
 - unknown / undesirable to user
 - want to prevent adversarial removal
- content provenance is for **public** usage
 - should be easy to decode
 - opt-in / valuable to user
 - want to survive benign distribution



- closed-source? **SynthID**
 - not suitable for public decoding
- open-source? **TrustMark, PixelSeal**
 - various issues in production...
- governed-usage (our model) **PawPrint**
 - ID space is managed across decentralized usage
 - single-purpose (provenance)
 - more teeth against bad actors



- durability
 - practical “benign attacks” are tough!
 - e.g. very few audio watermarks survive pitch modulation
 - e.g. very few video watermarks survive screenshotting
- invisibility
 - importance varies: low-value (everyday) vs. high-value (studio, cinema)
- speed
 - decode speed is particularly important
- support
 - complex file formats (e.g. streaming)
 - ease of deployment into workflows
- benchmarks
 - good benchmarks (like this one)
 - many benchmarks (for different use settings)



JPEG Trust Level 3 Proof-of-Concept

- sticker souvenirs
 - ...are watermarked!
 - (the watermark is not trained for print, so we were pleasantly surprised that it actually worked)



JPEG Trust Level 3 Proof-of-Concept

History

newest first

IMG_9621.jpeg

This file. It declares no previous versions.

No Content Credentials for this version.

SELECTED

This file
No credentials of its own
Details open

watermark

RECOVERED FROM WATERMARK

Signed original
AI-assisted
Trufo Inc.
August 21, 2026
signer trusted
hash mismatch
Show details >

Component
AI-generated
OpenAI OpCo, LLC
August 20, 2026
trusted
Show details >

Content Credentials
this laptop
Actions
L. Conrad, A. Altmann
Metadata
J. Hensley, E. JCP 2026
CAWG identity
B. Hensley (Trufo)
Generated on
13 September 2026
Recorded by
Trufo Provenance Platform

Preview from your device.

The dotted line is the watermark link, looked up by Trufo. Only this file was checked here; previous versions show what was recorded about them.

Green: trusted. Orange: not on a trust list, a failure that did not change the result, or a recovered original that no longer matches. Red: rejected.

verify.trufo.ai



JPEG Trust Level 3 Proof-of-Concept

History

newest first



This file
No credentials of its own

[Show details >](#)

watermark

RECOVERED FROM WATERMARK

SELECTED

Signed original
AI-assisted
Trufo Inc.
August 21, 2026
signer trusted
hash mismatch

[Details open](#)



Component
AI-generated
OpenAI OpCo, LLC
August 20, 2026
trusted

[Show details >](#)

The signed original

The signed original, built from 1 earlier version listed beneath it.



Content Credentials
this laptop

Actions
Opened Attended

Metadata
Timestamp ICP 2026

CAWG identity

Recorded on
13 September 2026

Recorded by
Trufo Provenance Platform

Thumbnail stored by the signer.

Declared origin

SOURCE TYPE

Edited using generative AI

C2PA signature

PRODUCT	Trufo Provenance Service (Distributed)
SIGNED BY	Trufo Inc.
COMMON NAME	Trufo Provenance Distributed Service
SIGNED ON	August 21, 2026 at 12:27 PM UTC
ISSUED BY	Trufo C2PA Claim Signing CA (2025)

Edits and activity (3)

Recorded order, which C2PA does not require to be chronological.

1. **Created** (c2pa.created): First creation of the asset.
2. **Watermarked bound** (c2pa.watermarked.bound): Action: c2pa.watermarked.bound
3. **Placed** (c2pa.placed): Content from another asset was placed in.

CAWG metadata

CREATOR	Cindy Han
---------	-----------

verify.trufo.ai



JPEG Trust Level 3 Proof-of-Concept

History

newest first



This file
No credentials of its own

Show details >

watermark

RECOVERED FROM WATERMARK



Signed original

AI-assisted
Trufo Inc.
August 21, 2026

signer trusted
hash mismatch

Show details >

SELECTED



Component

AI-generated
OpenAI OpCo, LLC
August 20, 2026

trusted

Details open

Component · openai-tampere-artwork.png

Trufo Provenance Service (Distributed) 1.1.1 used this as a part of it. It carries a signature of its own.



As recorded when signed.

Declared origin

SOURCE TYPE
Created using generative AI

C2PA signature

PRODUCT	OpenAI Media Service API
SIGNED BY	OpenAI OpCo, LLC
COMMON NAME	OpenAI Media Service
SIGNED ON	August 20, 2026 at 08:19 PM UTC
ISSUED BY	Trufo C2PA Claim Signing CA (2025)

Edits and activity (3)

Recorded order, which C2PA does not require to be chronological.

1. **Created** (c2pa.created): First creation of the asset. - By gpt-image 2.0 - August 20, 2026 (UTC)
2. **Converted** (c2pa.converted): Converted to another format. - August 20, 2026 (UTC)
3. **Watermarked unbound** (c2pa.watermarked.unbound): Action: c2pa.watermarked.unbound - August 20, 2026 (UTC)

Validation

verify.trufo.ai



New Technologies

- active areas of research at Trufo
 - watermark re-application (replace the watermark or extend the provenance?)
 - watermarking for 3D assets
 - watermarking with minimal re-encoding loss (ffmpeg)
 - fingerprinting for accurate asset recovery (within a large corpus)
 - fingerprinting to defend against public watermark forgery
- open to collaboration



Thank you for listening. Any questions?

