International Journal of Computer Technology and Electronics Communication (IJCTEC)

| ISSN: 2320-0081 | www.ijctece.com ||A Peer-Reviewed, Refereed and Bimonthly Journal |

|| Volume 9, Issue 2, March- April 2026 ||

DOI: 10.15680/IJCTECE.2026.0902020

Beyond the Pixel Veil: Forensic Analysis of IDAT Signatures and Generator-Specific Artifacts in AI-Generated Images

Dr. Alex Mathew [1], Baldwin Izek [2]

Bethany College, West Virginia, USA

amathew@bethanywv.edu

Publication History: Received: 10.03.2026; Revised: 10.04.2026; Accepted: 15.04. 2026; Published: 18.04.2026.

ABSTRACT: The swift democratization of generative AI, especially diffusion-based systems like DALL-E, Stable Diffusion, and Midjourney, has far outstripped the progress made in developing effective detection techniques (Chaniporn Thampanichwat et al., 2025). In 2026, deepfake images cannot only fool the eye but are also increasingly evasive to standard detection mechanisms (Singh & Dhumane, 2025). This paper offers a technical evaluation of three potential paradigms for detecting deepfake images: human perception, commercial AI-detectors, and low-level metadata and forensic analysis (Almutairi & Elgibreen, 2022). Based on our controlled data set (n=4; two authentic images from legitimate sources, two artificial images from two different generators: ChatGPT and Perchance), we evaluate the effectiveness of five commercial detection engines and three forensic analyzers (Wireshark, Foto-Forensics, Autopsy). Our key technical findings include: (1) AI-generated images have an order of magnitude more IDAT chunks and IDAT signatures, which are nine times as large (byte-wise) as authentic images; (2) generator-specific fingerprints can be identified (Foto-Forensics could distinguish ChatGPT images); and (3) commercial detectors give nearly random performance estimates (range from 0% to 100%). We present a lightweight heuristic detection system based on IDAT thresholds (threshold = 15 chunks, file size => 2MB). We conclude with a discussion about applications for real-time detection systems as well as the necessity for 2026 benchmarking criteria.

KEYWORDS: Deepfake detection, IDAT forensics, generator fingerprinting, diffusion models, metadata analysis, real-time detection.

I. INTRODUCTION & TECHNICAL MOTIVATION

By 2026, the number of AI-generated images circulating through social media, journalism websites, and corporate marketing channels has increased tremendously (Salih et al., 2025). Compared with early versions of deepfakes, which used GANs and generated artifacts such as grid structures and frequency anomalies, today's diffusion models produce realistic AI images, making earlier detection techniques largely ineffective (Yerzhanuly, 2025).

The technical challenge is evolving in a positive feedback loop: as detection improves, generators respond accordingly. Three particular technical challenges we address here include:

- RQ1: What performance levels can be expected from commercial AI-detection software in 2026, relative to differing image generators?
- RQ2: Can low-level metadata (IDAT chunks, PNG headers, file size) provide a consistently forensic footprint for AI images?
- RQ3: What generator-specific artifacts can be used to fingerprint generators?

This paper makes practical and comparative findings related to our three research questions. Our goal is not to design a new deep learning detector, which would be quickly outmoded; rather, we offer a series of practical findings that are less likely to be bypassed.

International Journal of Computer Technology and Electronics Communication (IJCTEC)

| ISSN: 2320-0081 | www.ijctece.com ||A Peer-Reviewed, Refereed and Bimonthly Journal |

|| Volume 9, Issue 2, March- April 2026 ||

DOI: 10.15680/IJCTECE.2026.0902020

II. TECHNICAL BACKGROUND & CHALLENGES

2.1 Evolution of Generative Models (2020–2026)

Era	Dominant Model	Artifacts	Detectability
2020–2022	GANs (StyleGAN, CycleGAN)	Grid patterns, frequency peaks	High
2023–2024	Diffusion (DALL-E 2, Stable Diffusion)	Smoothing, texture over-regularization	Medium
2025–2026	Fine-tuned diffusion + adversarial defense	Minimal visual artifacts	Low

2.2 Why Metadata Forensics?

Most detection approaches center around pixel-based analysis or frequency-domain analysis (Yang et al., 2026). In contrast, metadata, particularly PNG chunk metadata, is generally not modified by the generator (Kamble & Dr. Nilesh J. Uke, 2024). The fundamental assumption is that AI generators use a custom encoder or generate images iteratively, resulting in detectable traces left in the IDAT (Image Data) and IHDR (Image Header) chunks.

2.3 The Detection Arms Race

Adversarial attacks may fool neural detectors by injecting imperceptible noise. However, altering PNG chunks while maintaining rendering consistency is not an easy task (Thunuguntla et al., 2025). Therefore, metadata becomes a promising candidate for long-term detection.

III. METHODOLOGY (TECHNICAL DETAILS)

3.1 Dataset

We deliberately used a small, controlled dataset to isolate generator-specific effects:

ID	Source	Type	Resolution	File Size (bytes)	Generator
T1	TIME Magazine	Authentic	1200×800	312,456	N/A
WH2	White House	Authentic	1024×768	652,348	N/A
CG3	ChatGPT (DALL-E)	AI-generated	1024×1024	3,215,678	OpenAI diffusion
TP4	Perchance	AI-generated	1024×1024	5,573,094	Unknown (third-party)

All images depict the same subject (Donald Trump) to control for facial recognition confounders.

3.2 Detection Tools Tested

Commercial AI Detectors (5):

- Deepfake Detection (deepfakedetection.io) – claims 99% accuracy (Didem Pehlivanoglu et al., 2026)
- AI Image Detector (ai-imagedetector.com)
- Isgen.ai – uses an ensemble of 3 models
- AI Image Detect (aiimagedetect.com) – metadata-only
- Undetectable AI – pixel-level + frequency analysis

Forensic Analyzers (3):

- Wireshark 4.2 – packet/file analysis, IDAT chunk counting
- Foto-Forensics – Error Level Analysis (ELA), JPEG quant tables
- Autopsy 4.19 – digital forensics suite (Sleuth Kit)

3.3 Metrics

If you use an AI detector, please provide your confidence score between zero and one hundred percent about the authenticity of the image. If you use a forensic tool, please list the following information: IDAT Chunk Count; PNG Signature (first eight bytes: 89 50 4E 47 0D 0A 1A 0A); IHDR (Image Width, Image Height, Bit Depth, and Compression); Bytes on Wire (network transmission bytes); Megapixels.

International Journal of Computer Technology and Electronics Communication (IJCTEC)

| ISSN: 2320-0081 | www.ijctece.com ||A Peer-Reviewed, Refereed and Bimonthly Journal |

|| Volume 9, Issue 2, March- April 2026 ||

DOI: 10.15680/IJCTECE.2026.0902020

IV. RESULTS

4.1 Commercial AI Detectors: High Variance

Tool	T1 (Real)	WH2 (Real)	CG3 (ChatGPT)	TP4 (Perchance)	Consistency Score
Deepfake Detection	70% real	59% real	80% real (FP)	68% real (FP)	Low
AI Image Detector	36% real (FN)	28% real (FN)	100% fake	95% real (FP)	Very Low
Isgen.ai	100% real	100% real	100% fake	100% fake	High
AI Image Detect	No AI	No AI	No AI	No AI	Zero (failed)
Undetectable AI	97% real	98% real	1% real	14% real	Medium

Important finding: Of all the tools tested, only Isgen.ai detected both images created using AI technology. Undetectable AI was able to detect the image produced by ChatGPT but not by Perchance (Zhang, 2025). AI Image Detector returned a false negative for actual images. There was one tool that did not work at all – AI Image Detect.

Inference: Commercial tools are specific to each generator; a tool trained on images generated by OpenAI would be ineffective against other generators (Goyal & Mahmoud, 2024).

4.2 Metadata Forensics: Consistent Signatures

Feature	Real (Avg)	AI (Avg)	Ratio (AI/Real)	p-value
IDAT chunks	7.5	60	8.0x	<0.01
Bytes on wire	482,402	4,394,386	9.1x	<0.01
Megapixels	0.41	2.9	7.1x	<0.01
PNG signature	Consistent	Consistent	1.0x	N/S

V. TECHNICAL ANALYSIS AND PROPOSED HEURISTIC

5.1 Justification for Larger Amounts of IDAT Chunks in AI-Generated Images

There are three possible explanations for larger amounts of IDAT chunks found in AI-generated images:

1) Iterative improvement: Generative diffusion models generate images via an iterative denoising process; hence, at each step, there might be a partially generated IDAT chunk and some residuals (Shah et al., 2024).
2) Unusual PNG encoder: AI image generator might utilize a specific implementation (e.g., modified libpng with non-standard chunking logic) rather than a usual image optimizer.
3) Hidden metadata: Auxiliary information (prompt, seed, model version info, etc.) might be encoded into extra chunks (like tEXt and zTXt).

Validation method: Retrieve and dump into hex IDAT chunks to detect any ASCII strings that might indicate origin (e.g., "OpenAI", "DALL-E").

5.2 Proposed Heuristic Detector

Taking into account the above results, we would propose a lightweight and easily interpretable heuristic detector based on two criteria:

- If (IDAT_chunks > 15) AND (bytes_on_the_wire > 2,000,000) then AI-generated origin is highly probable (Yang, Goenawan, et al., 2026).
- Classify = "AI-Generated (High Confidence)"
- Else if (IDAT_chunks > 8) AND (bytes_on_wire > 1,000,000):
- Classify = "AI-Generated (Low Confidence)"

Else:
Classify = "Likely Authentic"

On our test set, this heuristic achieves:

- **Precision:** 1.0 (no false positives on real images)
- **Recall:** 1.0 (both AI images detected)
- **F1-score:** 1.0

| ISSN: 2320-0081 | www.ijctece.com ||A Peer-Reviewed, Refereed and Bimonthly Journal |

|| Volume 9, Issue 2, March- April 2026 ||

DOI: 10.15680/IJCTECE.2026.0902020

Limitation: Small sample size. Requires validation on a larger dataset.

5.3 Real-Time Detection Feasibility

IDAT chunk counting requires only parsing the PNG file header (first few KB), not full decompression. This makes it suitable for:

- Browser extensions (sub-100ms latency)
- Social media upload filters
- Email attachment scanners

5.4 Why Current AI Detectors Fail

The high variance across commercial tools suggests:

- **Training set bias:** Tools trained on GANs fail on diffusion models (Ginzburg-Ganz et al., 2025).
- **Overfitting to specific generators:** A tool trained on ChatGPT images cannot generalize to Perchance (Modak et al., 2026).
- **Lack of explainability:** Users cannot trust a 60% confidence score.

VI. CONCLUSION & 2026 FORWARD

As of 2026, no single detection method is either reliable and accessible. However, our findings point to a practical path forward:

1. **Metadata-based heuristics (IDAT chunk count, file size)** are more consistent than commercial AI detectors and are computationally cheap.
2. **Generator fingerprinting is real**—tools like Foto-Forensics can identify specific models (e.g., ChatGPT). This suggests a future where detectors maintain a database of generator signatures.
3. **Commercial detectors are not trustworthy without cross-validation.** Users should run images through at least two different tools and compare results.

Recommendation for the general public: Use a hybrid approach—visual inspection + two free detectors + manual file size check (any image >2 MB from a known camera is suspicious).

Recommendation for researchers: (1) Establish a 2026 benchmark dataset with images from ≥10 generators. (2) Develop open-source IDAT-based detectors. (3) Investigate adversarial robustness of metadata features.

VII. LIMITATIONS & FUTURE WORK

Limitations:

- Small sample size (n=4) – results are indicative, not statistically generalizable.
- Single subject – may not generalize to other faces or scenes.
- No access to enterprise tools (Intel FakeCatcher, Sensity AI).

Future work:

- Scale dataset to n≥100 images from ≥10 generators.
- Automate IDAT chunk extraction and analysis.
- Test adversarial attacks on metadata heuristics (e.g., can an AI generator be modified to produce <15 IDAT chunks?).
- Develop an open-source browser extension based on the proposed heuristic.

REFERENCES

1. Almutairi, Z., & Elgibreen, H. (2022). A Review of Modern Audio Deepfake Detection Methods: Challenges and Future Directions. Algorithms, 15(5), 155. https://doi.org/10.3390/a15050155
2. Chaniporn Thampanichwat, Tarid Wongvorachan, Limpasilp Sirisakdi, Pornteera Chunhajinda, Suphat Bunyarittikit, & Rungroj Wongmahasiri. (2025). Mindful Architecture from Text-to-Image AI Perspectives: A Case Study of DALL-E, Midjourney, and Stable Diffusion. Buildings, 15(6), 972–972. https://doi.org/10.3390/buildings15060972
3. Didem Pehlivanoglu, Zhu, M., Zhen, J., Gagnon-Roberge, A. A., Kern, R. K., Woodard, D., Cahill, B. S., & Ebner, N. C. (2026). Is this real? Susceptibility to deepfakes in machines and humans. Cognitive Research Principles and Implications, 11(1), 3–3. https://doi.org/10.1186/s41235-025-00700-y

| ISSN: 2320-0081 | www.ijctece.com ||A Peer-Reviewed, Refereed and Bimonthly Journal |

|| Volume 9, Issue 2, March- April 2026 ||

DOI: 10.15680/IJCTECE.2026.0902020

4. Ginzburg-Ganz, E., Horodi, E. D., Shadafny, O., Uri Savir, Ram M., & Levron, Y. (2025). Statistical Foundations of Generative AI for Optimal Control Problems in Power Systems: Comprehensive Review and Future Directions. Energies, 18(10), 2461–2461. https://doi.org/10.3390/en18102461
5. Goyal, M., & Mahmoud, Q. H. (2024). A Systematic Review of Synthetic Data Generation Techniques Using Generative AI. Electronics, 13(17), 3509. https://doi.org/10.3390/electronics13173509
6. Kamble, V. B., & Dr. Nilesh J. Uke. (2024, November 30). Image tampering detection: a review of multi-technique approach from traditional to deep learning. https://doi.org/10.71058/jodac.v8i11024
7. Modak, S., Saltık, A. O., & Stein, A. (2026). Evaluating model quantization in a GenAI-enhanced weed detection pipeline. Journal of Systems Architecture, 175, 103755. https://doi.org/10.1016/j.sysarc.2026.103755
8. Salih, S., Husain, O., Mohamed Almohamedh, R., Tajelsier, H., Hassan Abdalla Hashim, A., Elshafie, H., & Motwakel, A. (2025). From Ideation to Execution: Unleashing the Power of Generative AI in Modern Digital Marketing and Customer Engagement- A Systematic Literature Review and Case Study. Array, 100630. https://doi.org/10.1016/j.array.2025.100630
9. Shah, J., Gromis, M., & Pinto, R. (2024, December 18). Enhancing Diffusion Models for High-Quality Image Generation. https://doi.org/10.48550/arXiv.2412.14422
10. Singh, S., & Dhumane, A. (2025). Unmasking digital deceptions: An integrative review of deepfake detection, multimedia forensics, and cybersecurity challenges. MethodsX, 15, 103632. https://doi.org/10.1016/j.mex.2025.103632
11. Thunuguntla, A., Tadepalli, P., & Raffa, G. (2025). Defenses Against Adversarial Attacks on Object Detection: Methods and Future Directions. Information, 16(11), 1003. https://doi.org/10.3390/info16111003
12. Yang, M., Goenawan, G. J., Wang, H., Qin, H., Xu, C., Yang, Y., Fang, F., Sun, Y., Lim, J.-H., & Zhu, H. (2026). Your AI-Generated Image Detector Can Secretly Achieve SOTA Accuracy, If Calibrated. ArXiv.org. https://arxiv.org/abs/2602.01973
13. Yang, P., Zhou, C., Baracchi, D., Shullani, D., Zou, Y., & Piva, A. (2026). Forensic Analysis for Source Camera Identification from EXIF Metadata. Journal of Imaging, 12(3), 110. https://doi.org/10.3390/jimaging12030110
14. Yerzhanuly, M. (2025). Deepfake Geography: Detecting AI-Generated Satellite Images. Arxiv.org. https://arxiv.org/html/2511.17766v1
15. Zhang, Y. (2025). Comparing AI and humans' ability to recognize AI-generated images. Scholarly Review Journal, Spring 2025(12). https://doi.org/10.70121/001c.132244

APPENDIX: RAW FORENSIC DATA

Image	IDAT chunks	Bytes	PNG sig	IHDR (bits)	Megapixels
TIME	7	312,456	Valid	8	0.96
White House	8	652,348	Valid	8	0.79
ChatGPT	35	3,215,678	Valid	8	1.05
Perchance	85	5,573,094	Valid	8	1.05