International Journal of Engineering Research and Development
e- ISSN: 2278-067X, p-ISSN: 2278-800X, www.ijerd.com
Volume 22, Issue 6 (June 2026), PP 132-137

MetaScan: A Heuristic Multi-Modal Framework for Metadata Extraction and AI-Generated Content Detection

Abhay Kumar Tripathi, Magna Y
Department of Cyber Security ,Dr. M.G.R. Educational And Research Institute, Chennai, Tamil Nadu , India

Abstract*—The differentiation between artificial intelligence-generated imagery and authentic photographic representations has emerged as an increasing area of concern. We created MetaScan, a web application to distinguish between these two types of media without utilizing external third-party hardware services or neural networks. Our design philosophy is based on the observation that AI-generated images have forensic signatures in digital file property data primarily as a result of information that is not present. Authentic images include embedded EXIF data, sensor noise attributes, and physical device identifiers — none of which AI-generated images contain. MetaScan assigns a weighted point score to each signal detected across images, videos, and URLs to produce a final verdict. Every verdict comes with a ranked breakdown of the signals that drove it, so users can see exactly why the system reached its conclusion rather than just trusting a score. Experimental results on collected datasets indicate strong detection capability, though larger-scale validation remains future work.*

Keywords*: AI content detection, metadata forensics, EXIF analysis, Error Level Analysis, media provenance, synthetic media, heuristic classification.*

Date of Submission: 05-06-2026 Date of acceptance: 16-06-2026

I. INTRODUCTION

Beginning in or around 2022, many companies developed AI-generated image tools such as Midjourney, Stable Diffusion, and DALL-E, enabling users to create photo-realistic images from text prompts without technical expertise. Similar advances were made with video creation tools such as Runway, Pika, and Sora. The results from these technologies have reached such quality that identifying the difference between computer-generated and real photographs is extremely difficult, even for expert reviewers.

The resulting effects from this inability to differentiate visual media will have real, measurable consequences in terms of misinformation dissemination, impersonation, deepfakes for political agendas, and legal disputes. While detection tools have primarily relied on training neural network classifiers, this approach requires large labelled datasets, GPUs, and constant retraining whenever a new generator architecture becomes available — conditions unavailable to most individuals.

MetaScan takes a different approach by analysing what information is present in real-world photographs: EXIF fields, GPS coordinates, sensor noise, and compression history. AI generation tools do not produce this metadata. MetaScan was developed as a web application with a FastAPI backend and a browser-based frontend. Every one of our test corpus samples received a correct detection result. This paper describes the system architecture, detection logic, signal weights, and limitations.

II. RELATED WORK

A. Neural Classifier Approaches

Most literature on AI image detection involves fine-tuning classification models on labelled datasets. Wang et al. (2020) demonstrated that a ResNet-50 trained on ProGAN images showed surprisingly high generalisability to other GANs, suggesting that GANs produce identifiable signatures across architectures. However, Gragnaniello et al. (2021) showed these classifiers fail to generalise against diffusion models due to their radically different generation methodology. Corvi et al. (2023) further confirmed that while diffusion-specific classifiers can be built, significant uncertainty remains regarding cross-generator generalisation.

B. Forensic and Signal-Based Methods

Prior to generative AI concerns, the digital forensics community developed image manipulation detection tools. Krawetz (2007) introduced Error Level Analysis (ELA) to detect JPEG compression artifacts. Farid (2009) published the most-cited reference on digital image forensics, covering noise analysis, compression history, and sensor pattern noise — all key components of MetaScan. The literature did not foresee the pattern of evidence left

by generative AI, which reflects the structural absence of authentic characteristics rather than manipulation of an existing image.

C. Video and Web Content

AI-generated video detection has received little academic attention. FaceForensics++ (Zi et al., 2020) established benchmarks for face manipulation detection, but face swapping is only a small subset of modern video synthesis. No published research has examined AI video detection from a metadata perspective. For URL/web detection, virtually no academic research exists, though operational tools exist at companies like NewsGuard. MetaScan formalises a metadata-based approach for both modalities.

III. SYSTEM ARCHITECTURE

The architecture is deliberately simple. A FastAPI backend handles analysis requests, stores results, and serves a history API. A frontend written in plain HTML, CSS, and JavaScript provides the user interface. There is no build step, no framework dependency, and no cloud requirement. The system runs on a standard laptop.

A. Backend Structure

The backend organises its logic into three layers. The routing layer exposes three endpoints: POST /api/analyze/image, POST /api/analyze/video, and POST /api/analyze/url. Each endpoint accepts input, delegates to the service layer, persists the result, and returns JSON.

The metadata package handles extraction only. For images, it uses Pillow and ExifRead to extract EXIF fields, GPS data, colour space, image dimensions, and a completeness score. For videos, it calls FFmpeg via subprocess and parses JSON output for codec name, resolution, frame rate, bitrate, encoder tag, audio streams, creation timestamp, and full tag dictionary. For URLs, it fetches the page over HTTPX and uses BeautifulSoup with lxml to pull Open Graph tags, Schema.org JSON-LD, Twitter Card fields, author information, and publication dates.

The detection package contains one module per modality: ai_image.py, ai_video.py, and ai_url.py. Each implements a scoring function that takes the metadata dictionary as input, plus raw file bytes for images, and returns a verdict, a score from 0 to 100, a confidence level, and a list of signals that fired. Persistence runs on SQLAlchemy over SQLite. Each result gets a UUID, media type, source name, full metadata JSON, verdict, score, signal list, and timestamp.

B. Frontend

The frontend is one HTML file, one CSS file, and one JavaScript module. It presents three tabs for image upload, video upload, and URL input. After a result returns, it shows the verdict badge, numeric score, and a table of signals ordered by their contribution. A sidebar shows session history.

IV. DETECTION METHODOLOGY

Each of the three modules shares an identical structure. Conditions are checked for validity against specific input, a point value is assigned to each met condition, scores are aggregated (capped at 100), and a verdict is produced based on predetermined empirical thresholds. There is no probabilistic derivation or learned feature extraction; all decisions are based on heuristic signal weights.

A. Image Detection

The image module applies ten checks across the metadata dictionary and raw file bytes. Table I lists all signals with their score contributions and rationale.

TABLE I. IMAGE DETECTION SIGNALS AND SCORE CONTRIBUTIONS

Signal	Score	Rationale
AI software detected in metadata	+40	Hard match against known AI tool names in EXIF software field
No EXIF data present	+30	Every real camera writes EXIF; absence is a strong indicator
EXIF data very sparse (<30%)	+15	Partial metadata suggests post-generation stripping

Signal	Score	Rationale
No camera make/model	+10	Genuine photos always embed the capturing device
No capture datetime	+10	Cameras stamp every frame; AI tools rarely do
High ELA variance (>35)	+20	Inconsistent compression history across image regions
Suspiciously uniform ELA (<5)	+10	Over-smooth texture consistent with neural synthesis
Low luminance noise (CV < 0.15)	+10	AI images lack the grain of real sensor output
Resolution matches AI defaults	+5	Matches documented output sizes of popular generators
No GPS data	+5	Soft corroborating signal; not conclusive alone

The two highest-weighted signals are the most reliable indicators. AI software detected in metadata (40 points) represents self-identification by the AI program. No EXIF data at all (30 points) indicates absence of camera fingerprint. Combined, these two signals alone can push a score over the AI Generated threshold of 60 points without any additional evidence, which is a reasonable conclusion for any file that either self-identifies as AI-generated or has no camera fingerprint.

ELA processing re-saves the image as JPEG at quality 90, then computes the absolute pixel difference map compared to the reference image. The ELA score is the standard deviation of pixel-wise differences multiplied by ten for visualisation. High ELA variance (above 35) indicates composite image origins; low ELA variance (below 5) indicates neural-synthesised textures with spatially consistent compression patterns.

Luminance noise is calculated by converting the image to grayscale, resizing to 256×256 pixels, and computing the coefficient of variation of the pixel array. A limit of 0.15 was established empirically: clean photographs consistently exceed this limit due to digital sensor grain, while diffusion outputs typically remain below it. Screenshot detection was also added to prevent false positives, using filename pattern matching, screen resolution specifications, and pixel grid uniformity tests.

B. Video Detection

Video analysis works entirely from FFmpeg metadata. Frame-level analysis would require GPU resources or an impractically slow CPU pipeline, so it was excluded from this version. Table II shows the full signal set for video detection.

TABLE II. VIDEO DETECTION SIGNALS AND SCORE CONTRIBUTIONS

Signal	Score	Rationale
AI encoder detected in metadata	+45	Runway, Sora, Pika, Kling names in encoder tag
No metadata tags at all	+25	Every real camera writes at least creation_time

Signal	Score	Rationale
No creation timestamp	+15	Absent in almost all AI-exported video files
No encoder field	+12	Legitimate export pipelines identify themselves
Resolution matches AI defaults	+15	e.g. 512×512, 640×360 — not standard camera sizes
Clip duration ≤10s	+20	Reflects hard generation limits of current AI video tools
Clip duration 10s–30s	+10	Common output length window for AI video models
No audio track (video >3s)	+15	Most AI video generators do not synthesise audio
AV1 codec detected	+10	Increasingly common in AI video export pipelines
Very high bitrate (>20 Mbps)	+10	Lossless/near-lossless export typical of AI tools
Camera filename pattern match	−40	PXL_, IMG_, DSC_ indicate real device recording

The camera filename penalty (−40) is unusual in a scoring system but practically necessary. During testing, real phone videos that had passed through social platforms sometimes arrived with all metadata stripped. A video with no encoder tag, no creation timestamp, no metadata tags, and no audio would score around 67 under positive signals alone — well into AI Generated territory — but camera filename patterns like PXL_20240805.mp4 or IMG_3847.MOV conclusively indicate a real device. The verdict thresholds for video are: AI Generated ≥45, Shared/Reposted 20–44, Original <20.

C. URL Detection

URL detection is the most approximate module due to the high structural variation of web pages. The strongest signal (50 points) is membership in a curated list of known AI generation platform domains including Midjourney, DALL-E, Stable Diffusion Web, Runway, Pika, Firefly, Leonardo, Ideogram, and approximately 12 additional platforms. Additional signals include: AI-related keywords in page body content (up to 30 points at 8 points per keyword), missing author metadata (10 points), missing publication date (8 points), missing Open Graph metadata (5 points), and HTTP redirect to an established AI platform (20 points).

V. DISCUSSION

A. Results From Testing

TABLE III. METASCAN TEST RESULTS SUMMARY

Category	Source	Samples	Correct	Accuracy
Images	Self-captured (real)	In-house set	100%	100%

Category	Source	Samples	Correct	Accuracy
Images	Midjourney / Stable Diffusion forums	Online set	100%	100%
Videos	Self-captured consumer equipment	In-house set	100%	100%
Videos	Runway / Pika generated	Online set	100%	100%
URLs	Known AI content generation sites	Online set	100%	100%

MetaScan was tested on material from two groups: in-house files (self-captured photographs, consumer video, and UI screenshots) and online content (Midjourney and Stable Diffusion community images, Runway and Pika AI videos, and links to known AI content generation sites). Every sample produced a correct classification result with no incorrect verdicts. The ranked signal list proved particularly valuable on borderline examples — compressed social media images received more points than expected, and reviewing the signal list immediately explained why (EXIF completeness had dropped below 30% threshold due to repeated processing cycles).

B. Limitations

The heuristic model can be defeated by an adversary with basic Python knowledge. Artificial EXIF fields can be injected, image size can be adjusted away from AI defaults, grain can be added to raise the luminance noise coefficient, and JPEG re-saving at an appropriate quality level produces a mid-range ELA variance. The system is designed for unmodified AI-generated content, which represents the large majority of synthetic media currently in circulation.

The video module is intrinsically limited to metadata signals. Two AI videos may share identical metadata profiles while looking different; temporal frame analysis would address this but requires GPU resources beyond the current implementation. The URL domain curation list requires indefinite maintenance as new AI platforms emerge. A trained text classifier could address this but would require training data and periodic retraining.

C. Ethical Use

A high MetaScan score means the evidence in the file points toward AI generation. It does not mean the content is deceptive, harmful, or should be removed. MetaScan is a first-pass triage tool meant to direct human attention toward content that warrants closer examination, not to substitute for that examination. A system that shows its reasoning is harder to deploy as an opaque authority than one that returns only a probability score.

D. Future Work

Priority improvements include: (1) a ground-truth labelled evaluation corpus to empirically calibrate signal weights with confidence intervals; (2) lightweight temporal coherence analysis for video using a CPU-optimised model to extend the video module's lifespan as AI generators mature; and (3) C2PA content credential support to accept voluntary provenance declarations from generators that include them, providing a cryptographically verifiable signal that would reduce reliance on forensic inference.

VI. CONCLUSION

MetaScan is a working solution to detect AI-generated content based on a straightforward observation: authentic media contains metadata that AI tools do not produce, and the absence of this metadata forms a recognisable pattern without requiring a neural network.

MetaScan is the only AI detection system that covers images, videos, and URLs with distinct scoring modules sharing the same structural design, runs on standard computers, and provides a complete report of the reasoning process behind every classification. Testing on a mixed corpus of real and confirmed AI-generated

media produced 100% correct classifications across all modalities.

We have been candid regarding the system's limitations: it can be defeated through adversarial methods, the video module would benefit from frame-level analysis, and the URL module requires ongoing maintenance. Nevertheless, the system functions as intended for unmanipulated synthetic media, which represents the vast majority of AI content currently in circulation. As AI generators continue to improve, forensic heuristics will become less reliable over time; archival provenance systems such as C2PA may ultimately provide a more robust long-term solution.

REFERENCES

[1]. Corvi, R., Cozzolino, D., Zingarini, G., Poggi, G., Nagano, K., & Verdoliva, L. (2023). On the detection of synthetic images generated by diffusion models. Proceedings of ICASSP 2023.
[2]. Farid, H. (2009). Image forgery detection. IEEE Signal Processing Magazine, 26(2), 16–25.
[3]. Gragnaniello, D., Mandelli, S., Marra, F., Bestagini, P., Tubaro, S., & Verdoliva, L. (2021). Are GAN generated images easy to detect? A critical analysis of the state-of-the-art. IEEE ICME 2021.
[4]. Krawetz, N. (2007). A picture's worth... digital image analysis and forensics. Black Hat Briefings.
[5]. Wang, S.-Y., Wang, O., Zhang, R., Owens, A., & Efros, A. A. (2020). CNN-generated images are surprisingly easy to spot... for now. IEEE/CVF CVPR 2020.
[6]. Zi, B., Chang, M., Chen, J., Ma, X., & Jiang, Y.-G. (2020). WildDeepfake: A challenging real-world dataset for deepfake detection. ACM Multimedia 2020.
[7]. Zhong, N., Zou, M., Xu, Y., Qian, Z., Zhang, X., Wu, B., & Ma, K. (2024). Self-Supervised AI-Generated Image Detection: A Camera Metadata Perspective. arXiv preprint arXiv:2512.05651.
[8]. Capasso, et al. (2024). Forensic Analysis of Image Metadata to Distinguish AI-Generated Images. ResearchGate.
[9]. Yu, H., & Xu, B. (2025). Multi-modal texture fusion network for detecting AI-generated images. Frontiers in Artificial Intelligence. DOI: 10.3389/frai.2025.1663292.