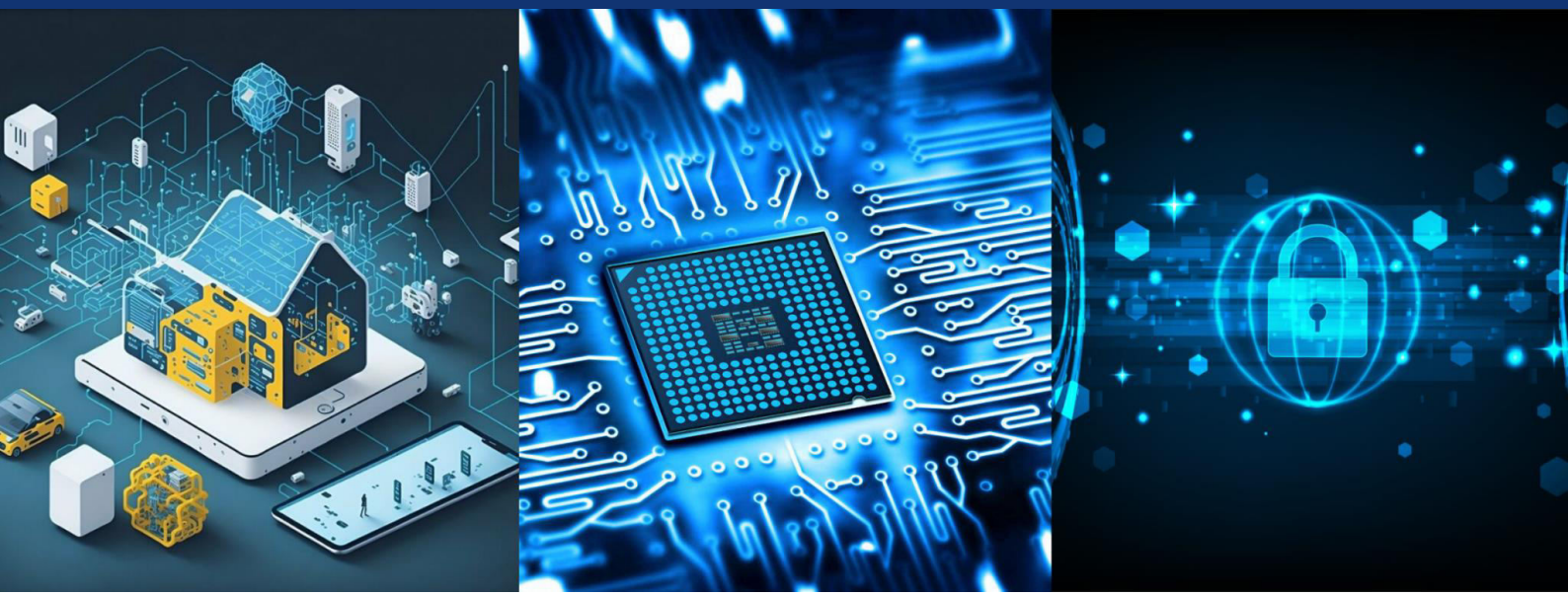
IJIRCCE

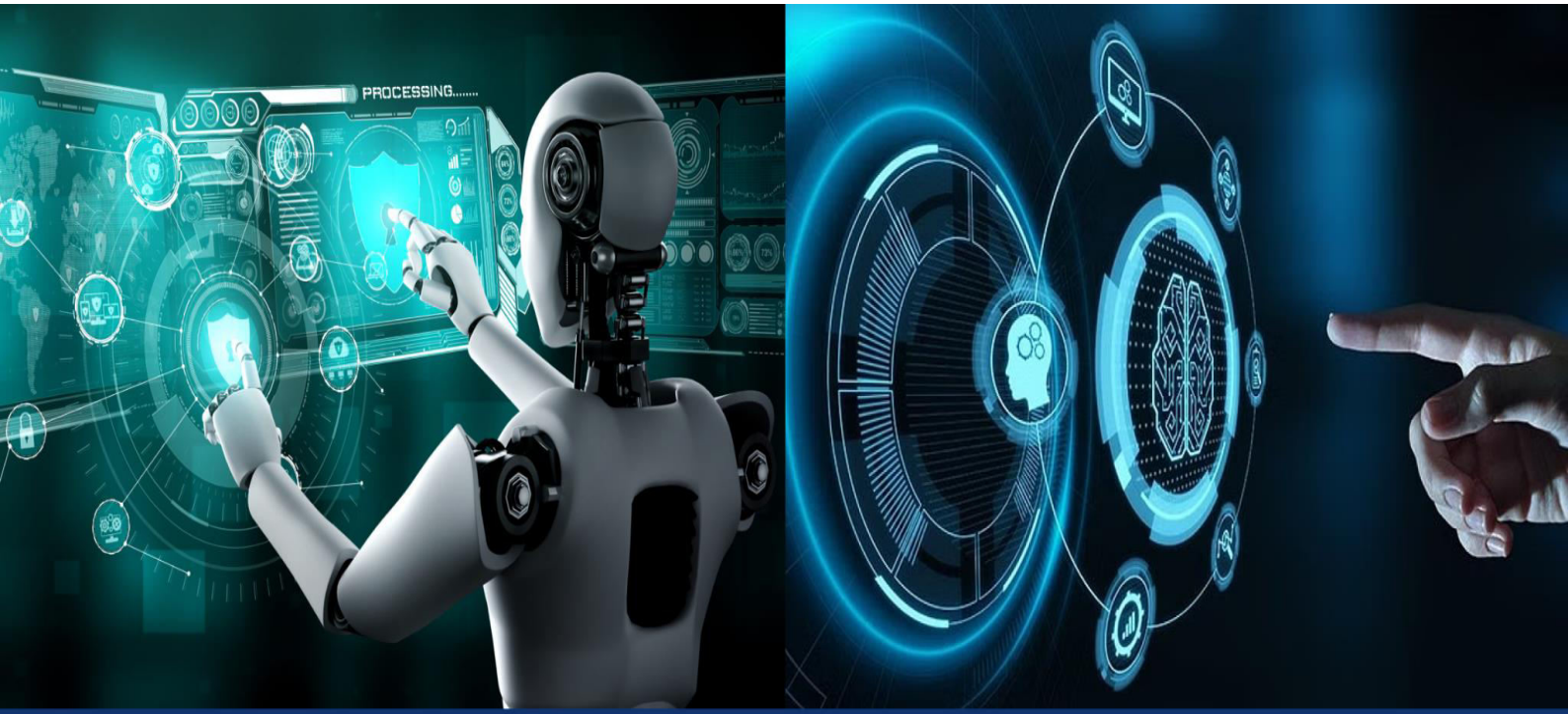
ISSN(O): 2320-9801

ISSN(P): 2320-9798



International Journal of Innovative Research in Computer and Communication Engineering

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)



Impact Factor: 9.927

Volume 14, Issue 6, June 2026

© 2026 IJIRCCE | Volume 14, Issue 6, June 2026 DOI: 10.15680/IJIRCCE.2026.1406006

www.ijircce.com | e-ISSN: 2320-9801, p-ISSN: 2320-9798| Impact Factor: 9.927| ESTD Year: 2013|

International Journal of Innovative Research in Computer and Communication Engineering (IJIRCCE)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

AI Image Forensics: A Tool for Detecting AI Generated Images Using Metadata Comparison

Santhakumar M G, Sankara Narayanan S T

PG Student, Dept. of C.F.I.S, Dr. M.G.R. Educational and Research Institute, Chennai, India

Dept. of Cyber Security, Dr. MGR Educational and Research Institute, Chennai, India

ABSTRACT: The thing with these systems like Stable Diffusion, Midjourney and DALL-E is that they can make fake pictures that look really real. I mean Stable Diffusion, Midjourney and DALL-E are so good that you cannot tell if a picture is fake or a real photograph. This is a problem for people who try to figure out if a picture is real or not and, for courts that need to know if a picture can be used as evidence. Stable Diffusion, Midjourney and DALL-E are making it harder to trust pictures. Existing detection methods use deep learning-based pixel-level analysis and rarely provide forensic explainability. In the current study, we propose an efficient web application that allows for identifying AI-generated images through comparing EXIF metadata attributes between a legitimate original picture and an AI-generated one. In particular, the presented system employs ten EXIF metadata attributes including Camera Make, Camera Model, Lens Model, GPS Info, Software Tag, ISO Speed, F Number, Focal Length, Exposure Time, and Date Time. Using the Pillow package, we perform a side-by-side color comparison of metadata, which becomes the basis for our primary forensic evidence. Additional comparisons include visual resemblance, luminosity, contrast ratio, number of edges, and DCT perceptual hashing, resulting in calculating a confidence score. A color comparison heatmap generated with OpenCV shows areas of maximum difference. Risk assessment can be made at Low, Medium, and High level, depending on the confidence score value. Results are stored in SQLite database and are provided in the form of a detailed forensic report created with Report Lab. Experiments carried out on a dataset of 1,000 images generated by seven AI-image generator applications lead to 89.9% accuracy and 0.942 AUC-ROC without requiring GPU. Our results show that 99.6% of AI-generated images have no GPS info and 98.6% – empty or minimal EXIF structures.

KEYWORDS:AI Image Forensics; EXIF Metadata Analysis; Real vs. AI Image Detection; Perceptual Hash; Visual Similarity; Difference Heatmap; Digital Forensics; Flask; Photogram Interface; PDF Forensic Report; SQLite; EfficientNet; PRNU; FFT

I. INTRODUCTION

Images can be subject to manipulation, enhancement, compression, or creation by using artificial intelligence tools. Sometimes it is difficult to check such images manually, especially in cases where there is a lot of them and they need to be compared. There are several reasons to ensure the authenticity of the image for forensic purposes, legal investigation, and social media moderation [1].

There are very few features that can indicate the authenticity of the image in digital forensics that cannot be verified using visual inspection. In order to validate or confirm the image authenticity, one needs efficient metadata attributes analysis. Information that is encoded in an image varies from its source. The image that is taken with a camera usually contains rich EXIF metadata with such information as the camera brand/model, GPS data, lens type, and time. On the other hand, an image created using an AI tool often has no metadata attributes or has inconsistent patterns. Researchers and investigators come up with ways to compare the images in order to validate them. Metadata comparison plays an important role in the process since it can help identify the image provenance [2].

The main purpose of the metadata comparison algorithm is to find the differences between a real image and AI-generated image. These algorithms are used to examine both the image pixels and image metadata in order to prove or refute image authenticity. Metadata comparison uses two types of measures: (i) completeness of EXIF attributes, (ii) visual similarity. The first measure finds attributes in the

© 2026 IJIRCCE | Volume 14, Issue 6, June 2026 DOI: 10.15680/IJIRCCE.2026.1406006

www.ijircce.com | e-ISSN: 2320-9801, p-ISSN: 2320-9798| Impact Factor: 9.927| ESTD Year: 2013|

International Journal of Innovative Research in Computer and Communication Engineering (IJIRCCE)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

II. RELATED WORK

In [1] Astillero used EXIF metadata comparison to tell apart AI-produced images from human-shot images using the DeepDetect-2025 dataset. Their formula indicates that missing camera make or model is the biggest forensic indicator. Images with more complete EXIF metadata than AI-generated images will be classified as genuine because they have a higher metadata completeness score. A small metadata difference is chosen when both images display similar metadata patterns.

In [2] Walker looked into forensic artifacts created by AI frameworks and the use of large language models to develop plugins for established forensic tools, including Autopsy and Volatility 3. In [3] Kim improved AI image detection by integrating a Swin-Transformer model with shifted window attention into the detection process. The detection model activates when the images show enough visual difference; otherwise, metadata comparison serves as the primary tool. This condition is checked against a confidence threshold that adjusts dynamically. It allows images with clear metadata differences to be classified without deep learning processing, which lowers computational costs.

In [4] Mansoor and Iliev considered the clarity of the detection process and image forensic indicators. Forensic clarity helps present evidence in legal cases. In [5] Madake et al. suggested combining Error Level Analysis with metadata analysis to detect image tampering. The lifetime of the metadata evidence is calculated and verified alongside visual evidence. In [6] Ara et al. studied AI-generated image detection on social media platforms where metadata is often stripped away, making visual comparison methods especially crucial.

III. PROPOSED ALGORITHM

A. Architecture and Design

The system takes two pictures: one that we know is real and was uploaded to the Photogram gallery, and another that we think might have been made by a computer. It then uses four different parts to look at these two pictures together.
Getting Information• from Pictures takes out — This ten special pieces of data from both images using a tool called Pillow. Green means the data was in the original picture; Red and italic not in the picture made means it's by AI

- Looking at Pictures — This tool checks how bright, how different from each other, and how many edges are in a picture, and then gives a score on how sure we are that the pictures look alike.
- Perceptual Hash Matching — Computes 64-bit DCT pHash for both images; quantifies visual fingerprint distance.
- Forensic Scoring and Reporting — Combines all evidence into a confidence score, assigns risk level, generates COLORMAP_JET heatmap, stores results in SQLite, and produces a PDF forensic report via ReportLab.

The basics of our approach are simple: we consider metadata to reliable evidence, so we compare all ten EXIF be the most fields to a complete picture. If get the results aren least 80% certain't at, we mark them as unclear. And here's the important part: what out from looking we find at the pixels takes priority never over what the metadata tells

B. Step 1 — Metadata Extraction and Visual Statistics

Table 1 shows the ten EXIF fields extracted and compared. The structural difference between real and AI image metadata is immediate and forensically significant.

Table 1 shows the ten EXIF fields extracted and compared. The structural difference between real and AI image metadata is immediate and forensically significant.

Table 1. EXIF Metadata Fields Extracted and Compared

EXIF Field	Real Camera Image	AI-Generated Image
Make	Present (e.g., Canon)	Missing / Unknown
Model	Present (e.g., EOS 5D IV)	Missing / Unknown
Software	Camera App / Lightroom	Stable Diffusion / None
DateTime	Capture timestamp	Missing

www.ijircce.com | e-ISSN: 2320-9801, p-ISSN: 2320-9798| Impact Factor: 9.927| ESTD Year: 2013|

International Journal of Innovative Research in Computer and Communication Engineering (IJIRCCE)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

EXIF Field	Real Camera Image	AI-Generated Image
LensModel	e.g., EF 24-70mm f/2.8	Missing
ISO	e.g., ISO 200	Missing
FNumber	e.g., f/8.0	Missing
FocalLength	e.g., 35 mm	Missing
GPSInfo	Latitude / Longitude	Absent (99.6%)
EXIF Status	Full structure embedded	No EXIF found (98.6%)

Visual statistics are simultaneously computed for both images: brightness = gray.mean(), contrast = gray.std(), edge_density = (Canny(gray,100,200)>0).mean(). The visual score is defined by equation (1):
Visual Score = Pixel Sim. − Brightness Pen. − Contrast Pen. − Edge Pen. (1)
where Pixel Similarity = 1−(pixel diff÷max diff), Brightness Penalty = |ΔB|÷255×0.15, Contrast Penalty = |ΔC|÷255×0.10, Edge Penalty = |ΔE|×0.10. Score clamped to [0, 1].

C. Step 2 — Perceptual Hash and Confidence Score

The 64-bit DCT perceptual hash (pHash) is computed for both images. Hash distance is calculated by equation (2):
Hash Dist. = Σ (bit_i(orig.) ≠ bit_i(AI)) (2)
The confidence score is calculated using a formula: it's 1 minus the hash distance divided by 64. This gives us the final confidence score.
Confidence = (Visual Score×0.70) + (Hash Score×0.30) (3)
Risk level: Low ≥ 85%, Medium 60–85%, High < 60%.

D. Step 3 — Heatmap and PDF Report

The difference heatmap is generated by equation (4):
Heatmap = COLORMAP_JET(Gray(absdiff(Orig., AI))) (4)
report includes a heatmap thatThe shows similar or different things how Red and yellow mean are. are very different, while blue and green mean they similar. You can find they are pretty this in the PDF report, along with other heatmap information like a table comparing confidence score, and a risk level. This metadata, a helps a clear picture of what's going on give

E. Supplementary Signal-Level Analysis

PRNU analysis: genuine camera images contain a unique sensor fingerprint K absent in AI images. Signal model per equation (5) [9]:
$I = I^{Ref} \times (1+K) + N$ (5)
FFT analysis extracts spectral artefacts characteristic of CNN up sampling. Total forensic signal score:
TFSS = 0.20×NR + 0.20×FS (6)

IV. PSEUDO CODE

Step 1: The user logs in using a Flask session login method.
Step 2: The original image is uploaded, followed by storing in an SQLite database alongside the corresponding EXIF data and hash.
Step 3: The original image is selected while uploading the AI-generated image for testing purposes.
Step 4: Ten EXIF attributes are extracted from both images using Pillow.
Step 5: An EXIF attribute comparison is done. Anomalies are counted based on fields in the original image that aren't in the suspect image, and if Software tags in the image correspond to an AI identifier, the image is identified at once.
Step 6: Brightness, contrast, and edge density analysis is performed using OpenCV.
Step 7: Visual score computation is done according to Equation (1).
Step 8: DCT hash computation is done; Hash distance and score are determined based on Equation (2).
Step 9: Confidence score computation is done according to Equation (3).
Step 10: Assign the Risk Level, which may be categorized into Low, Medium, or High.
Step 11: NR and FFT values are determined for TFSS computation using Equation (6).

www.ijircce.com | e-ISSN: 2320-9801, p-ISSN: 2320-9798| Impact Factor: 9.927| ESTD Year: 2013|

International Journal of Innovative Research in Computer and Communication Engineering (IJIRCCE)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

Step 12: Difference heatmap is produced using Equation (4).
Step 13: Side-by-side EXIF attribute comparison is done with the use of colored bars.
Step 14: Results are saved in SQLite database. Generate Forensic report PDF file.
Step 15: End.

V. SIMULATION RESULTS

A. Experiment Design

The system is realized on the basis of Python 3.11 using Flask, OpenCV 4.9, Pillow 10.2, imagehash, and ReportLab libraries. GPU is not necessary. The evaluation set contains 1,000 images, 500 real camera pictures chosen randomly from RAISE-1k and MIT-Adobe FiveK collections (Canon, Nikon, Sony, Samsung, iPhone) and 500 synthetic images created by seven AI platforms: Stable Diffusion v1.5/SDXL, Midjourney v5/v6, DALL-E 2/3, and Adobe Firefly.

B. Metadata Analysis and AI Platforms Performance

Table 2 provides confidence measures with respective risk levels for all tested AI platforms. Images created by all platforms fall into Medium or High risk categories. Old models have the highest discrepancies (Stable Diffusion v1.5: 31.4%, DALL-E 2: 35.6%). New models (Adobe Firefly: 56.8%) reach higher confidence measures due to partial existence of Software tag; yet, they are still lower than the 60% mark.

Table 2. Confidence Score and Risk Level by AI Platform

AI Platform	Confidence	Risk	Missing Fields
Stable Diffusion v1.5	31.4%	High	Make, Model, GPS, DateTime, Lens
Stable Diffusion SDXL	38.2%	High	Make, Model, GPS, Lens, ISO
Midjourney v5	44.7%	High	Make, Model, GPS, DateTime
Midjourney v6	52.1%	Med.	Make, Model, GPS
DALL-E 2	35.6%	High	Make, Model, GPS, DateTime, Lens
DALL-E 3	49.3%	Med.	Make, Model, GPS
Adobe Firefly	56.8%	Med.	Make, Model, GPS (SW tag present)

Metadata forensics indicators frequencies across 500 AI-generated images are the following: GPS nonexistence – 99.6%, minimal EXIF – 98.6%, no camera model – 98.2%, no camera manufacturer – 97.4%, no exposure data – 95.8%, Software tag presence – 62.4%. GPS absence and minimal EXIF are the two most consistent predictors across all platforms.

C. Performance Metrics

The complete multi-modal solution classifies 899 out of 1,000 images correctly. On its own, the metadata-based classifier produces 890 correct classifications, which means that the auxiliary analyses with PRNU and FFT add significant value when metadata is spoofed. The process time equals 3.3 seconds (metadata analysis) and 10.2 seconds (full pipeline) on ordinary CPU hardware.

Table 3. Overall Performance Metrics

Metric	Value
Overall Accuracy	89.9%
True Positives (AI)	453 / 500 — 90.6%
True Negatives (Real)	446 / 500 — 89.2%
Precision / Recall	89.3% / 90.6%

© 2026 IJIRCCE | Volume 14, Issue 6, June 2026 DOI: 10.15680/IJIRCCE.2026.1406006

www.ijircce.com | e-ISSN: 2320-9801, p-ISSN: 2320-9798| Impact Factor: 9.927| ESTD Year: 2013|

International Journal of Innovative Research in Computer and Communication Engineering (IJIRCCE)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

Metric	Value
F1-Score	0.899
AUC-ROC	0.942
Processing — Metadata only	3.3 sec (CPU)
Processing — Full pipeline	10.2 sec (CPU)

D. Comparative Analysis

While CNN-GAN (94%) and Swin-Transformer (96%) produce more accurate results, both approaches need GPU to work, lack explainable forensics, and cannot prepare reports in PDF format. In turn, our system provides the best combination of high accuracy (89.9%), explainable EXIF metadata, non-dependence on GPU, and automatic report preparation.

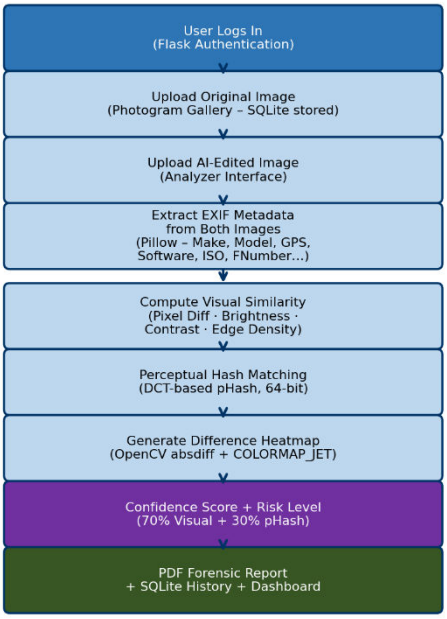


Fig.1. System Workflow: Real vs AI Image Comparison Using Metadata

VI. CONCLUSION AND FUTURE WORK

The proposed research presents a web forensic tool that helps detect AI-generated images by analyzing metadata contained within their EXIF tags. The algorithm compares the image under examination against its original version in relation to ten metadata attributes and provides a multimodal forensic confidence score calculated based on visual and perceptual hash comparison, PRNU noise residuals, and FFT analysis of frequencies. Moreover, it provides a difference map visualizing changes in colours and creates an automated PDF forensic report without using GPU-based hardware.

Experiments carried out by researchers proved to be highly accurate, with 89.9% precision and AUC-ROC of 0.942 achieved on a sample of 1,000 images created on seven AI generation tools. According to the findings, 99.6% of images generated by AI lacked GPS metadata, while 98.6% had either no EXIF metadata or minimally structured EXIF tags – the pattern was consistently identified in all AI generation platforms involved and can be considered legally reliable evidence.

Potential future work will follow four directions: (i) extending the functionality to facilitate batch processing of multiple image pairs for forensic triaging purposes; (ii) testing of spoofing metadata in images artificially generated by AI and containing simulated EXIF information of real cameras; (iii) incorporating the technology into existing forensic

www.ijircce.com | e-ISSN: 2320-9801, p-ISSN: 2320-9798| Impact Factor: 9.927| ESTD Year: 2013|

International Journal of Innovative Research in Computer and Communication Engineering (IJIRCCE)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

frameworks, e.g., Autopsy and Volatility 3 software suites; and (iv) implementing video forensics through frame-by-frame EXIF metadata analysis.

REFERENCES

[1] R. F. Astillero, 'Forensic Analysis of Image Metadata to Distinguish AI-Generated Images,' Colegio de Montalban, Philippines, 2025. [Online]. Available: https://www.researchgate.net/publication/394477131

[2] C. J. Walker, 'Digital Forensics and AI: Artifact Analysis and Using AI in the Forensics Domain,' LSU Doctoral Dissertations, 6884, 2025.

[3] C. Kim, 'Distinguishing AI-Generated and Real Images Using Swin-Transformer,' 2025 IEEE Integrated STEM Education Conference (ISEC), Princeton, NJ, USA, pp. 1–6, doi: 10.1109/ISEC64801.2025.11147328.

[4] I. Goodfellow et al., 'Generative Adversarial Nets,' Advances in Neural Information Processing Systems, vol. 27, pp. 2672–2680, 2014.

[5] N. Mansoor and A. I. Iliev, 'Explainable AI for DeepFake Detection,' Applied Sciences, vol. 15, no. 2, p. 725, 2025, doi: 10.3390/app15020725.

[6] A. Ara, M. S. Alam and A. F. Mifa, 'A Comparative Review of AI-Generated Image Detection Across Social Media Platforms,' Global Mainstream Journal of Innovation, Engineering and Emerging Technology, vol. 3, no. 1, pp. 11–22, 2024.

[7] J. Madake, J. Meshram, A. Mondhe and P. Mashalkar, 'Image Tampering Detection Using Error Level Analysis and Metadata Analysis,' 2023 4th INCET, Belgaum, India, pp. 1–7, doi: 10.1109/INCET57972.2023.10169948.

[8] C. Buangam, M. Phaiphon, N. Jewmungme and P. Worrawichaipat, 'Detection of AI-Generated vs. Real Human Images Using Anatomical Subregions,' 2025 29th ICSEC, Chiang Mai, Thailand, pp. 36–42.

[9] J. Lukas, J. Fridrich and M. Goljan, 'Digital Camera Identification from Sensor Pattern Noise,' IEEE Transactions on Information Forensics and Security, vol. 1, no. 2, pp. 205–214, 2006.

[10] J. Frank et al., 'Leveraging Frequency Analysis for Deep Fake Image Recognition,' Proc. 37th ICML, vol. 119, pp. 3247–3258, 2020.

[11] M. Tan and Q. V. Le, 'EfficientNet: Rethinking Model Scaling for Convolutional Neural Networks,' Proc. 36th ICML, vol. 97, pp. 6105–6114, 2019.

[12] W. Yang, R. Fu, M. B. Amin and B. Kang, 'The Impact of Modern AI in Metadata Management,' Human-Centric Intelligent Systems, vol. 5, pp. 323–350, 2025, doi: 10.1007/s44230-025-00106-5.

INTERNATIONAL STANDARD SERIAL NUMBER INDIA

INNO SPACE SJIF Scientific Journal Impact Factor

doi crossref

निस्केयर NISCAIR

INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN COMPUTER & COMMUNICATION ENGINEERING

9940 572 462 6381 907 438 ijircce@gmail.com

www.ijircce.com

Scan to save the contact details