

 Presentation + Paper

21 May 2025

LEAD-AI: lightweight entropy analysis for distinguishing AI-generated images from genuine photographs

Monica Sudarsan, Nihal A. Poredi, Evan Maurer, Enoch Solomon, Yu Chen (/profile/Yu.Chen-47294)

[Author Affiliations + \(\)](#)

[Proceedings Volume 13480, Disruptive Technologies in Information Sciences IX: \(/conference-proceedings-of-spie/13480.toc\)](#)
134800N (2025) <https://doi-org.ez9.periodicos.capes.gov.br/10.1117/12.3055540> (<https://doi-org.ez9.periodicos.capes.gov.br/10.1117/12.3055540>)

Event: [SPIE Defense + Commercial Sensing \(/conference-proceedings-of-spie/browse/SPIE-Defense-Security/2025\)](#), 2025, Orlando, Florida, United States

ARTICLE ▼	FIGURES & TABLES	REFERENCES	CITED BY ▼
-----------	------------------	------------	------------

Abstract

PROCEEDINGS
15 PAGES + PRESENTATION

DOWNLOAD STARTED

SAVE TO MY LIBRARY

WATCH PRESENTATION

▼
GET CITATION

Advertisement


The proliferation of AI-generated images poses significant challenges for digital media authentication and security. While existing detection methods rely primarily on computationally intensive deep learning algorithms, we introduce LEAD-AI (Lightweight Entropy Analysis for Distinguishing AI-Generated Images), a computationally efficient approach that analyzes entropy features in the spatial domain to identify synthetic images. Our method identifies subtle distinctions in pixel arrangement and texture complexity that result from the synthetic image generation process. By examining low-frequency regions where AI generators typically leave distinctive artifacts, LEAD-AI effectively differentiates between AI-generated images and authentic photographs while maintaining minimal computational overhead. Experimental results demonstrate that our approach successfully distinguishes images created by popular generative models (including Stable Diffusion XL and Adobe Firefly) from genuine photographs, with accuracy rates exceeding 98% in controlled testing environments. Additionally, LEAD-AI can differentiate between various AI generation tools, providing valuable forensic attribution capabilities. This technique offers a practical solution for rapid image authentication in resource-constrained environments, addressing the growing need for efficient tools to combat visual misinformation.

Conference Presentation



RIGHTS & PERMISSIONS

[Get copyright permission](#)
(<https://marketplace.copyright.com/rs-ui-web/mp/search/all/10.1117%2f12.3055540>).

 Get copyright permission on Copyright Marketplace

KEYWORDS

[Diffusion](#)

[Image classification](#)



- [Data modeling](#)
- [Education and training](#)
- [Image segmentation](#)
- [Artificial intelligence](#)
- [Photography](#)
- [Show All Keywords](#)

RELATED CONTENT

AGC automated machine learning for garbage classification on mobile... (/conference-proceedings-of-spie/12287/122872E/AGC--automated-machine-learning-for-garbage-classification-on-mobile/10.1117/12.2640931.full)
Proceedings of SPIE (October 13 2022)

Diffusion model uniform manifold filtering for classification of small datasets... (/conference-proceedings-of-spie/13517/135170K/Diffusion-model-uniform-manifold-filtering-for-

1 Introduction

The advent of generative models, ranging from Generative Adversarial Networks (GANs) to diffusion-based techniques, has revolutionized digital image synthesis. These advances have expanded creative boundaries and introduced significant challenges in media authentication and digital forensics. With AI-generated images achieving near-photorealism, distinguishing between synthetic content and genuine photographs has become increasingly critical, particularly in contexts where misinformation and deepfakes pose serious risks.

Most existing detection methods heavily rely on deep learning architectures to classify images based on learned features. While they perform well, these approaches use complex and computationally intensive models that are unsuitable for real-time applications in resource-constrained environments. Moreover, the less-interpretable black-box nature of these techniques restricts understanding of the actual differences between real and generated images.

This raises questions about how we will spot AI-generated images in the future. An interesting area to explore is the frequency domain, where subtle differences in texture, pixel distribution, and structural regularities become significant features that sometimes go unnoticed in the spatial domain. By examining the energy distribution and entropy using well-known methodologies such as the Fast Fourier transform (FFT) and the Discrete Cosine Transform (DCT), we can uncover intrinsic differences introduced during the generation of synthetic images.

Furthermore, due to computational efficiency, frequency domain methods are particularly suitable for real-time applications and implementation in resource-constrained environments. They offer a promising balance between performance and feasibility in the field of digital capture verification.

While frequency domain analysis presents compelling advantages for detecting AI-generated images, our research identified an opportunity to develop a more straightforward yet effective approach that operates directly in the spatial domain. By strategically focusing on low-frequency areas within images, such as skies, walls, or other relatively uniform regions, we can identify distinctive entropy patterns without requiring full-frequency transformation. This spatial domain approach maintains the computational efficiency benefits while eliminating the processing overhead associated with domain transformation operations. Our method capitalizes on the observation that AI-generated images tend to maintain natural randomness in these low-complexity regions.

classification-of-small-datasets/10.1117/12.3055038.full)
Proceedings of SPIE (February 24 2025)

Identification of sea turtles with graphical convolutional networks and SLIC... (/conference-proceedings-of-spie/13136/131360/Identification-of-sea-turtles-with-graphical-convolutional-networks-and-SLIC/10.1117/12.3028445.full)
Proceedings of SPIE (September 30 2024)

Deep learning model for accurate vegetation classification using RGB image... (/conference-proceedings-of-spie/11398/113980/Deep-learning-model-for-accurate-vegetation-classification-using-RGB-image/10.1117/12.2557833.full)
Proceedings of SPIE (April 21 2020)

Canopy recognition of cherry fruit tree based on SegNet network... (/conference-proceedings-of-spie/11915/119150/Canopy-recognition-of-cherry-fruit-tree-based-on-SegNet-network/10.1117/12.2605032.full)
Proceedings of SPIE (September 15 2021)

A new way of deep learning combined with street view... (/conference-proceedings-of-spie/11913/119130/A-new-way-of-deep-learning-combined-with-street-view/10.1117/12.2605032.full)
Proceedings of SPIE (August 06 2021)

This paper introduces a **Lightweight Entropy Analysis** approach for **Distinguishing AI-generated Images** (LEAD-AI). Entropy, as a measure of randomness or disorder within the data, has long been used in signal processing and image analysis. By applying entropy analysis directly to strategic spatial regions of the image, our method captures subtle discrepancies in pixel distribution and texture complexity that arise from the synthetic generation process. These discrepancies not only facilitate the differentiation between AI-generated images and genuine photographs but also distinguish images produced by various AI models and tools, offering a balance of computational efficiency and detection accuracy.

Our contributions can be summarized as follows.

- We introduce an innovative, low-overhead approach for digital image authentication that uses entropy analysis applied to strategically selected spatial regions within images.
- We demonstrate that our method accurately differentiates between AI-generated and genuine photographs by analyzing the statistical properties of low-frequency areas where synthetic artifacts are most detectable.
- We show that the method can further differentiate between images produced by various AI models (specifically, Stable Diffusion XL and Adobe Firefly), offering a means of source attribution in forensic applications.
- We validate LEAD-AI's performance through extensive experiments, showcasing its potential for real-time deployment in situations with limited computational resources while maintaining high detection accuracy.
- We provide insights into the entropy distribution characteristics of different image sources, establishing a statistical foundation for lightweight detection methods that can evolve alongside advancing generative technologies.

The remainder of the paper is organized as follows. [Section 2 \(/conference-proceedings-of-spie/13480/3055540/LEAD-AI--lightweight-entropy-analysis-for-distinguishing-AI-generated/10.1117/12.3055540.full#s2\)](#) reviews related work on digital image authentication and the detection of AI-generated content. [Section 3 \(/conference-proceedings-of-spie/13480/3055540/LEAD-AI--lightweight-entropy-analysis-for-distinguishing-AI-generated/10.1117/12.3055540.full#s3\)](#) details the LEAD-AI methodology, including the entropy analysis framework and the Gaussian classifier. [Section 4 \(/conference-proceedings-of-spie/13480/3055540/LEAD-AI--lightweight-entropy-analysis-for-distinguishing-AI-generated/10.1117/12.3055540.full#s4\)](#) presents our experimental setup, datasets, and evaluation results, followed by [Section 5 \(/conference-proceedings-of-spie/13480/3055540/LEAD-AI--lightweight-entropy-analysis-for-distinguishing-AI-generated/10.1117/12.3055540.full#s5\)](#), which discusses the limitations of the proposed classifier along with the potential future works. Finally,

Section 6 (/conference-proceedings-of-spie/13480/3055540/LEAD-AI--lightweight-entropy-analysis-for-distinguishing-AI-generated/10.1117/12.3055540.full#s6) concludes the paper and discusses future research directions aimed at further improving the robustness and applicability of our approach.

2 Background and Related Work

Rapid growth of highly realistic synthetic images has been mainly caused by the development of Generative Adversarial Networks (GANs).¹² (/conference-proceedings-of-spie/13480/3055540/LEAD-AI--lightweight-entropy-analysis-for-distinguishing-AI-generated/10.1117/12.3055540.full#c12) The emergence of GANs in 2014,¹³ (/conference-proceedings-of-spie/13480/3055540/LEAD-AI--lightweight-entropy-analysis-for-distinguishing-AI-generated/10.1117/12.3055540.full#c13) Dall-E,¹⁴ (/conference-proceedings-of-spie/13480/3055540/LEAD-AI--lightweight-entropy-analysis-for-distinguishing-AI-generated/10.1117/12.3055540.full#c14) and text-to-image diffusion models in 2022¹⁵ (/conference-proceedings-of-spie/13480/3055540/LEAD-AI--lightweight-entropy-analysis-for-distinguishing-AI-generated/10.1117/12.3055540.full#c15) were among the significant accomplishments in the area of image generation. In contrast, current advanced models such as Dall-E 3¹⁶ (/conference-proceedings-of-spie/13480/3055540/LEAD-AI--lightweight-entropy-analysis-for-distinguishing-AI-generated/10.1117/12.3055540.full#c16) and stable diffusion¹⁷ (/conference-proceedings-of-spie/13480/3055540/LEAD-AI--lightweight-entropy-analysis-for-distinguishing-AI-generated/10.1117/12.3055540.full#c17) demonstrate hyperrealistic image generation in different styles.

These models use substantial datasets and advanced architecture to produce images based on hypothetical variables along with semantic outputs and text prompts. As a result, AI-generated content (AIGC) is now available on social media¹⁸ (/conference-proceedings-of-spie/13480/3055540/LEAD-AI--lightweight-entropy-analysis-for-distinguishing-AI-generated/10.1117/12.3055540.full#c18) and the entertainment industry,² (/conference-proceedings-of-spie/13480/3055540/LEAD-AI--lightweight-entropy-analysis-for-distinguishing-AI-generated/10.1117/12.3055540.full#c2) has infiltrated scientific fields,¹⁹ (/conference-proceedings-of-spie/13480/3055540/LEAD-AI--lightweight-entropy-analysis-for-distinguishing-AI-generated/10.1117/12.3055540.full#c19) and has been used for data augmentation.²⁰ (/conference-proceedings-of-spie/13480/3055540/LEAD-AI--lightweight-entropy-analysis-for-distinguishing-AI-generated/10.1117/12.3055540.full#c20) However, misuse of technology through deepfakes and misinformation campaigns has raised concerns about fake media. The emergence of malicious use of such technology requires methods and tools that distinguish natural images from AI images. Despite their realism, AI-generated images often harbor subtle, model-specific artifacts. For example, GAN-generated images may present artifacts, inconsistent textures, or abnormal frequency distributions due to limitations in convolution and upsampling operations.²¹ (/conference-proceedings-of-spie/13480/3055540/LEAD-AI--lightweight-entropy-analysis-for-distinguishing-AI-generated/10.1117/12.3055540.full#c21) Our previous work has exploited this fact to bypass this detection method in efforts to generate undetectable images.⁹ (/conference-proceedings-of-spie/13480/3055540/LEAD-AI--lightweight-entropy-analysis-for-distinguishing-AI-generated/10.1117/12.3055540.full#c9) Additionally, Electrical Network Frequency (ENF) has been used to detect deepfaked audio²² (/conference-proceedings-of-spie/13480/3055540/LEAD-AI--lightweight-entropy-analysis-for-distinguishing-AI-generated/10.1117/12.3055540.full#c22) and visual²³ (/conference-proceedings-of-spie/13480/3055540/LEAD-AI--lightweight-entropy-analysis-for-distinguishing-AI-generated/10.1117/12.3055540.full#c23) content, which has been proven to be a unique temporal fingerprint²⁴ (/conference-proceedings-of-spie/13480/3055540/LEAD-AI--lightweight-entropy-analysis-for-distinguishing-AI-generated/10.1117/12.3055540.full#c24) (/conference-proceedings-of-spie/13480/3055540/LEAD-AI--lightweight-entropy-analysis-for-distinguishing-AI-generated/10.1117/12.3055540.full#c25), ²⁶ (/conference-proceedings-of-spie/13480/3055540/LEAD-AI--lightweight-entropy-analysis-for-distinguishing-AI-generated/10.1117/12.3055540.full#c26)

$$H(X) = - \sum_x P(x) \log_2 P(x) \quad (1)$$

As defined by Eq. 1 ([/conference-proceedings-of-spie/13480/3055540/LEAD-AI--lightweight-entropy-analysis-for-distinguishing-AI-generated/10.1117/12.3055540.full#d1](#)), entropy in information theory quantifies the measure of unpredictability within a dataset.²⁷ ([/conference-proceedings-of-spie/13480/3055540/LEAD-AI--lightweight-entropy-analysis-for-distinguishing-AI-generated/10.1117/12.3055540.full#c27](#)) Assessment of visual data complexity and naturalness using information density depends on entropy in image forensics. Natural images exhibit entropy distributions aligned with natural scene statistics, representing different textures, lighting effects, and edge details.²⁸ ([/conference-proceedings-of-spie/13480/3055540/LEAD-AI--lightweight-entropy-analysis-for-distinguishing-AI-generated/10.1117/12.3055540.full#c28](#)) AI-generated images, on the other hand, sometimes reveal entropy irregularities caused by the model's properties, either showing unnaturally high regularity in certain regions or excessive disorder in specific areas of the image.

In the field of digital forensics, we need fast and efficient solutions. Entropy is a computationally inexpensive statistical measure that can be calculated using simple probability distributions. Natural images follow an established statistic, and the entropy calculation reflects different textures, edges, and lighting effects. AI-generated images often violate this statistic due to limitations in the architecture of the models. Thus, entropy is a very simple but potentially powerful method to detect subtle deviations between natural and fake images. Several studies have been conducted to investigate the use of entropy in AI-generated images. Cozzolino utilized a lossless image encoder to calculate the probability distribution of each pixel²⁹ ([/conference-proceedings-of-spie/13480/3055540/LEAD-AI--lightweight-entropy-analysis-for-distinguishing-AI-generated/10.1117/12.3055540.full#c29](#)) but does not rely on actual training data.

In the frequency domain, such as in DCT and FFT, low-frequency components capture broad structures, differences in brightness, and smooth gradients, which are factors of natural image perception. AI models such as GANs and other diffusion models often struggle to replicate the low-frequency distributions, which leads to the smoothing of large areas in an image and inconsistent brightness contrast. Meanwhile, high-frequency regions have too many artifacts, edges, noise patterns, and textures. They are less discriminative for generalization. The peculiarities of low-frequency regions are consistent and more robust in artifact detection in the images. Attention toward lightweight entropy metrics results in faster and more scalable solutions. By focusing on low-frequency areas, the global artifact left by AI images can be detected in the lower medium of the image, as demonstrated.¹⁰ ([/conference-proceedings-of-spie/13480/3055540/LEAD-AI--lightweight-entropy-analysis-for-distinguishing-AI-generated/10.1117/12.3055540.full#c10](#))

3 Lightweight Entropy Analysis: Rationale and Algorithms

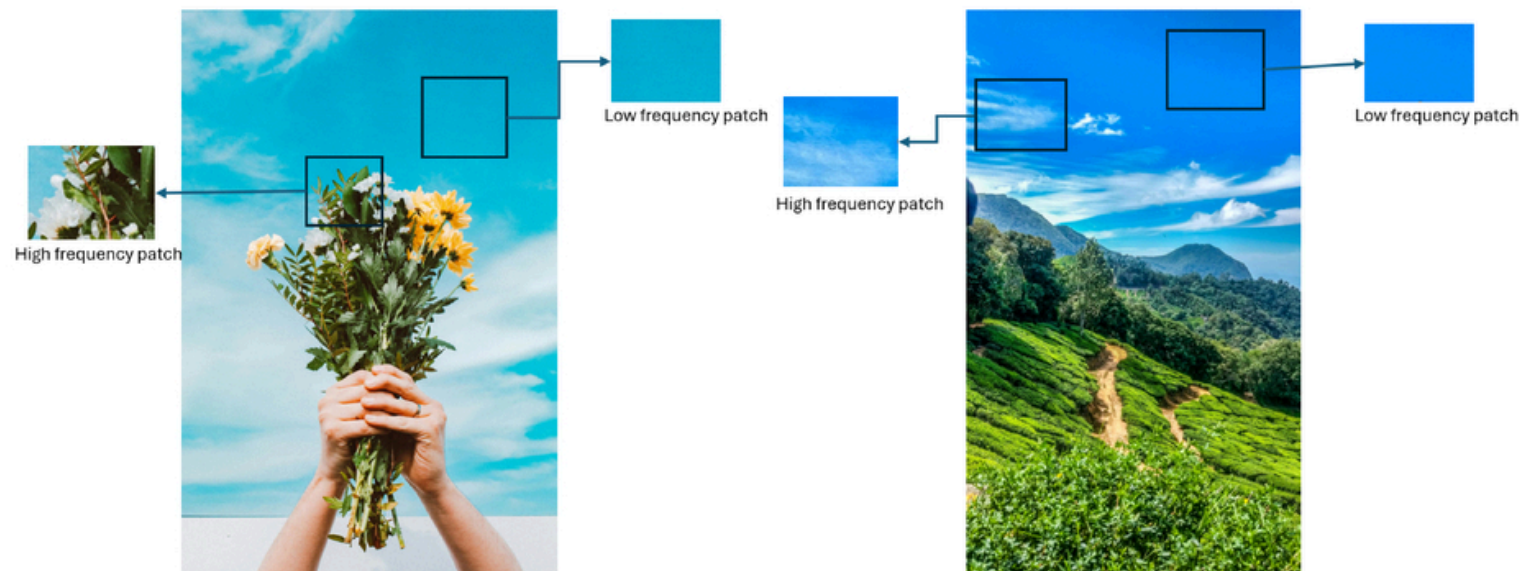
3.1 Rationale and Process

In natural images, the dispersion of pixel distribution across channels such as Red, Green, and Blue is very complex, often non-uniform, and also varying with lighting effects. The measure of entropy for these channels is usually higher and tends to be nonuniform, depending on the factors involved. On the other hand, AI-generated images are expected to tend to be more synthesized and exhibit less randomness and lower entropy compared to natural images. For comparative analysis, we have curated datasets for both AI-generated and natural images. For the AI-generated images, we have utilized the data set from a Stable Diffusion XL model hosted on OpenArt.ai and Adobe Firefly. A total of 600 images, 200 from each category (Stable Diffusion XL, Adobe Firefly, and Natural Images), were collected for training. Both models give us hyper-realistic images that look nearly identical to their natural counterparts. Natural images were sourced from genuine publicly available photographs in the category “natural” scenes.

To focus the analysis on low-frequency regions, we followed the patch-based approach to detect the entropy values of each and every image, as shown in [Figure 1 \(/conference-proceedings-of-spie/13480/3055540/LEAD-AI--lightweight-entropy-analysis-for-distinguishing-AI-generated/10.1117/12.3055540.full#f1\)](#). The extracted patch is 100 × 100 in size and plain patches that do not overlap. Opting for a 100 × 100 pixel size gives a balance between spatial statistics. We have implemented an approach that automates the segmentation of pixel patches according to the required pixel patch size. Isolating clean patches in an image is a common preprocess, but implementing them manually is tedious and labor intensive when scaling up the data set count. Although manual cropping seems very easy, it introduces errors and inconsistency in datasets, and what one person recognizes as the best region may differ from the other. Consistency in the derivation of the patch size and the yield of high-quality segments results in maximum performance and definitive values.

Fig 1 [Download \(/proceedings/DownloadFigures?url=/ContentImages/Proceedings/13480/134800N/FigureImages/00018_PSIISDG13480_134800N_page_4_1.jpg\)](#)

Patch selection and extraction



The algorithm ensures that we follow a grid approach and is applied to the whole image. Converting the image into an HSV color space from RGB is more robust under different lighting conditions. The hue (H) typically has information regarding colors, whereas the saturation (S) and value (V) capture low-saturation areas, which include some overlappings in the plain patch. The patch is then automatically segmented, where the saturation and the value are adjusted to get the plain patches without overlapping.

For classification purposes, we have calculated the entropy for each channel (Red, Green, Blue) separately for both AI-generated and natural images. When distinguishing between natural and AI-generated images, the calculation of only entropy values does not detail local variations and any irregularities in the patches. To capture both factors, we computed the mean and standard deviation for the segregated patches and color channels.

3.2 Lightweight Gaussian Classifier

For classification purposes, we developed a multivariate Gaussian-based classifier that uses the distinctive entropy features extracted from the color channels of image patches. This lightweight probabilistic approach provides an efficient alternative to computationally intensive deep learning methods while maintaining robust discrimination capability.

Given a set of entropy features extracted from an image patch, we model each class distribution (Adobe Firefly, Stable Diffusion XL, and Natural images) as a multivariate Gaussian distribution. Let $\mathbf{X} = [X_R, X_G, X_B]^T$ represent the three-dimensional feature vector comprising entropy values from the red, green, and blue channels respectively. For each class $c \in \{\text{Firefly}, \text{SDXL}, \text{Natural}\}$, we estimate the mean vector $\boldsymbol{\mu}_c = [\mu_{c,R}, \mu_{c,G}, \mu_{c,B}]^T$ and covariance matrix Σ_c from the training data.

The probability density function for class c is given by:

$$p(\mathbf{X}|c) = \frac{1}{(2\pi)^{d/2} |\Sigma_c|^{1/2}} \exp \left(-\frac{1}{2} (\mathbf{X} - \boldsymbol{\mu}_c)^T \Sigma_c^{-1} (\mathbf{X} - \boldsymbol{\mu}_c) \right) \quad (2)$$

where $d = 3$ is the dimensionality of our feature space and $|\Sigma_c|$ denotes the determinant of the covariance matrix. To simplify computation while preserving classification efficacy, we employ a diagonal covariance matrix which assumes independence between the RGB channels. This simplification reduces the multivariate Gaussian to the product of univariate Gaussians:

where

$$\sigma_{c,i}^2$$

(ContentImages/Proceedings/13480/134800N/FigureImages/00018_PSISDG13480_134800N_page_5_3.jpg) represents the variance of channel i for class c . For classification, we employ the maximum likelihood estimation approach. Given a test image with the entropy feature vector \mathbf{X}_{test} , we compute the likelihood for each class and assign the image to the class with the highest probability:

$$\hat{c} = \arg \max_{c \in \{\text{Firefly}, \text{SDXL}, \text{Natural}\}} p(\mathbf{X}_{test}|c) \quad (4)$$

This classifier offers several advantages for our application. First, it provides a probabilistic interpretation of class membership, allowing for confidence assessment in the classification results. Second, its computational simplicity enables real-time processing even in resource-constrained environments. Finally, despite its simplicity, the Gaussian model effectively captures the distinctive entropy distributions we observed in the different image categories, as demonstrated by our experimental results.

Algorithm 1 (</conference-proceedings-of-spie/13480/3055540/LEAD-AI--lightweight-entropy-analysis-for-distinguishing-AI-generated/10.1117/12.3055540.full#A1>) for the Gaussian classifier is presented below:

Algorithm 1 [Download \(/proceedings/DownloadFigures?url=/ContentImages/Proceedings/13480/134800N/FigureImages/00018_PSIDG13480_134800N_page_6_2.jpg\)](/proceedings/DownloadFigures?url=/ContentImages/Proceedings/13480/134800N/FigureImages/00018_PSIDG13480_134800N_page_6_2.jpg)

Classification using Gaussian PDF

Require: $test_val$ (row vector of color-channel values), $statistics$ (training datasets mean and std)

Ensure: Best matching dataset

$max_prob \leftarrow -\infty$

$best_match \leftarrow \text{None}$

for each dataset in $statistics$ **do**

$means \leftarrow statistics[dataset][\text{"means"}]$

$stds \leftarrow statistics[dataset][\text{"stds"}]$

Compute element-wise Gaussian PDF:

$$prob \leftarrow \frac{1}{\sqrt{2\pi\sigma_{c,i}^2}} \exp\left(-\frac{(X_i - \mu_{c,i})^2}{2\sigma_{c,i}^2}\right)$$

Compute total probability:

$$total_prob \leftarrow \prod prob$$

if $total_prob > max_prob$ **then**

$max_prob \leftarrow total_prob$

$best_match \leftarrow dataset$

end if

end for

return $best_match$

The classifier captures differences in the distributions of AI-generated and natural images and has proven to be very efficient when the data set is limited. The testing data set we collected is made up of 135 Adobe Firefly images, 140 Stable Diffusion XL images, and 100 Natural images. This data set was processed to collect their channel entropy values, which were then read into the classifier. Examples of the test values input into the classifier can be seen in the [Tables 1](https://www.spiedigitallibrary.org.ez9.periodicos.capes.gov.br/conference-proceedings-of-spie/13480/3055540/LEAD-AI--lightweig) ([/conference-proceedings-of-spie/13480/3055540/LEAD-AI--lightweig](https://www.spiedigitallibrary.org.ez9.periodicos.capes.gov.br/conference-proceedings-of-spie/13480/3055540/LEAD-AI--lightweig)

[entropy-analysis-for-distinguishing-AI-generated/10.1117/12.3055540.full#T1](#)), [2 \(/conference-proceedings-of-spie/13480/3055540/LEAD-AI--lightweight-entropy-analysis-for-distinguishing-AI-generated/10.1117/12.3055540.full#T2\)](#), and [3 \(/conference-proceedings-of-spie/13480/3055540/LEAD-AI--lightweight-entropy-analysis-for-distinguishing-AI-generated/10.1117/12.3055540.full#c3\)](#) below. A proper input into the classifier would be a vector of the image's three color-channel entropy values. For example, an input for the Firefly image “*Img1*” would look like [1.2695, 5.9388, 6.5078]. We ignore the entropy of the overall segment when classifying the images, as it is a composite value of the channel entropies a does not demonstrate the uniqueness of the images and the channel values.

Table 1

Adobe Firefly Training data.

Firefly	Entropy(100x100)	Red Entropy	Green Entropy	Blue Entropy
Img1	6.1570	1.2695	5.9388	6.5078
Img2	5.4922	1.1021	4.9770	5.6428
Img3	5.2829	0.9826	4.2912	5.8333
Img4	6.1862	1.2052	5.9725	6.6699
Img5	5.2468	1.2613	4.6347	5.0893

Table 2

SDXL Training data.

SDXL	Entropy(100x100)	Red Entropy	Green Entropy	Blue Entropy
Img1	4.6730	3.0204	3.3821	2.8616
Img2	5.1085	3.8017	3.6493	3.1197
Img3	4.9565	3.0788	3.4610	3.5748
Img4	5.4051	3.4697	4.0463	3.9445
Img5	5.0080	3.1866	3.5096	3.5730

Table 3

Natural Images Training data.

Natural	Entropy(100x100)	Red Entropy	Green Entropy	Blue Entropy
Img1	6.1298	5.2535	4.5463	3.8346

Img2	5.8049	4.6113	4.2476	3.8009
Img3	6.5894	4.7518	5.1164	5.1582
Img4	3.8684	4.3131	3.2948	3.5554
Img5	3.9811	0.1340	3.9616	3.22681

4 Experimental Study

4.1 Experimental Results

The entropy characteristics of the training data were carefully analyzed to determine distinctive patterns for image classification. [Figures 2 \(/conference-proceedings-of-spie/13480/3055540/LEAD-AI--lightweight-entropy-analysis-for-distinguishing-AI-generated/10.1117/12.3055540.full#f2\)](#), [3 \(/conference-proceedings-of-spie/13480/3055540/LEAD-AI-lightweight-entropy-analysis-for-distinguishing-AI-generated/10.1117/12.3055540.full#f3\)](#), [4 \(/conference-proceedings-of-spie/13480/3055540/LEAD-AI--lightweight-entropy-analysis-for-distinguishing-AI-generated/10.1117/12.3055540.full#f4\)](#), and [5 \(/conference-proceedings-of-spie/13480/3055540/LEAD-AI--lightweight-entropy-analysis-for-distinguishing-AI-generated/10.1117/12.3055540.full#f5\)](#) illustrate the normal distributions of the entropy values for the overall image, as well as individual red, green, and blue channels in our three image categories (Adobe Firefly, Stable Diffusion XL, and natural photographs). [Table 4 \(/conference-proceedings-of-spie/13480/3055540/LEAD-AI--lightweight-entropy-analysis-for-distinguishing-AI-generated/10.1117/12.3055540.full#T4\)](#) provides the corresponding numerical values for the mean and standard deviation.

Fig 2 [Download \(/proceedings/DownloadFigures?url=/ContentImages/Proceedings/13480/134800N/FigureImages/00018_PSIDG13480_134800N_page_8_1.jpg\)](#)

Normal Distributions of Training Data Overall Entropy

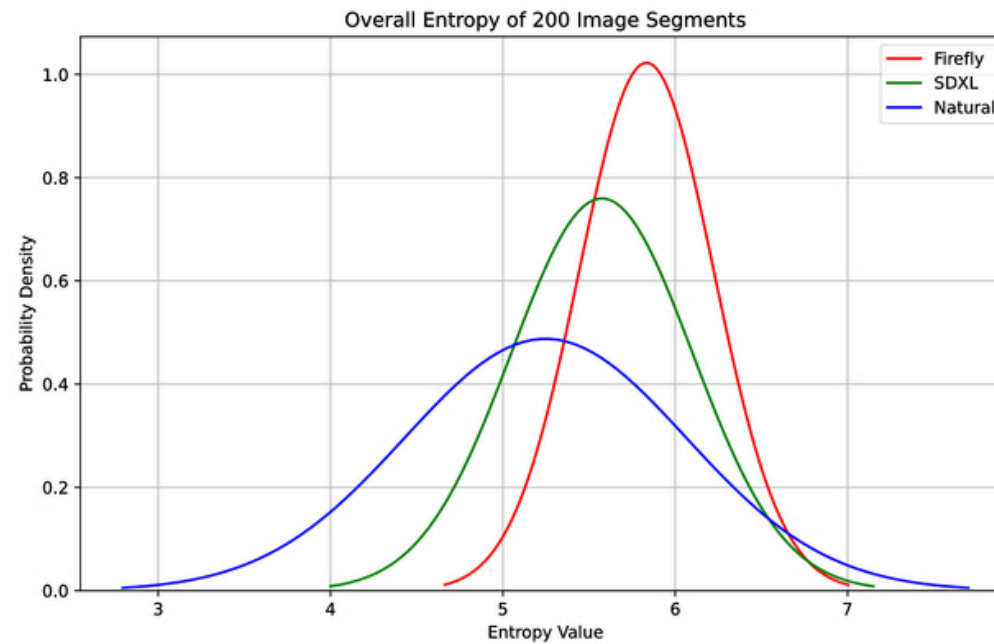


Fig 3 [Download \(/proceedings/DownloadFigures?url=/ContentImages/Proceedings/13480/134800N/FigureImages/00018_PSISDG13480_134800N_page_8_2.jpg\).](#)

Normal Distributions of Training Data Red Entropy

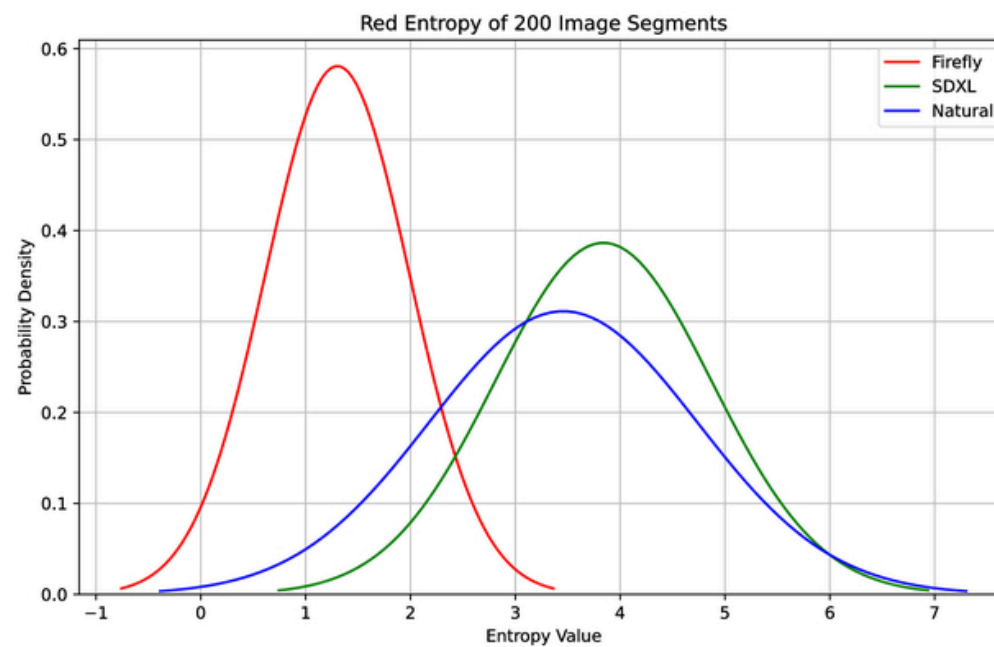


Fig 4 [Download \(/proceedings/DownloadFigures?url=/ContentImages/Proceedings/13480/134800N/FigureImages/00018_PSiSDG13480_134800N_page_9_1.jpg\).](#)

Normal Distributions of Training Data Green Entropy

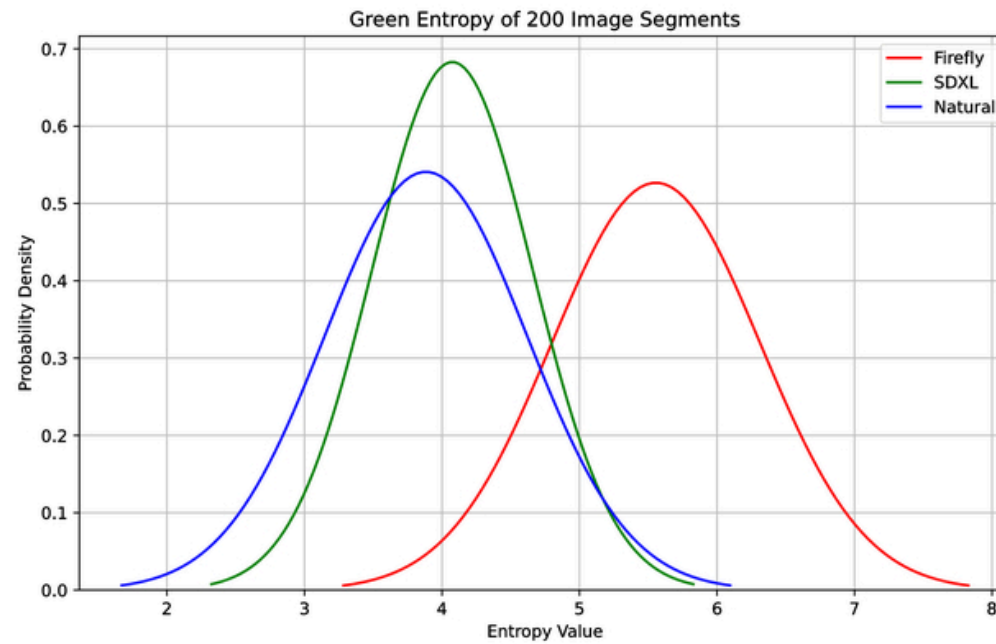


Fig 5 [Download \(\(proceedings/DownloadFigures?url=/ContentImages/Proceedings/13480/134800N/FigureImages/00018_PSiSDG13480_134800N_page_9_2.jpg\).](#)

Normal Distributions of Training Data Blue Entropy

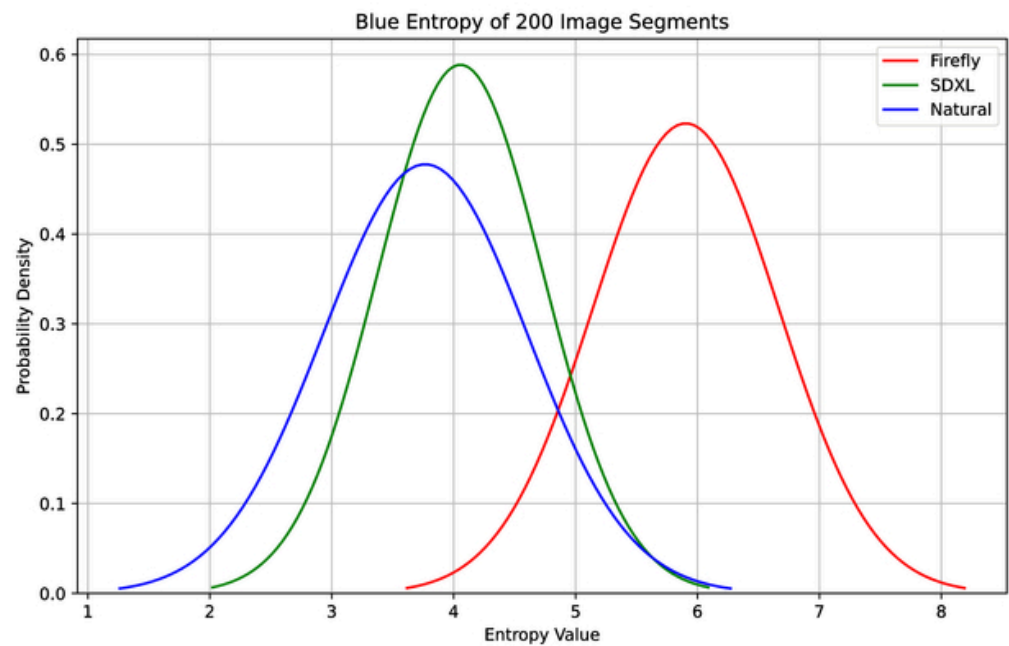


Table 4

Mean and Standard Deviation Values for Each Dataset and Their Respective Channels.

Dataset Metrics	Segment Entropy	Red Entropy	Green Entropy	Blue Entropy
Firefly Mean	5.8335	1.3028	5.5559	5.9039
Firefly STD	0.3911	0.6886	0.7592	0.7643
SDXL Mean	5.5748	3.8404	4.0754	4.0552
SDXL STD	0.5263	1.0347	0.5856	0.6795
Natural Mean	5.2473	3.4570	3.8831	3.7664
Natural STD	0.8199	1.2848	0.7394	0.8371

For the following discussion, please note that Adobe Firefly is represented as red, Stable Diffusion XL as green, and Natural images as blue in the previous figures. We will first focus c attention on the Adobe Firefly curves shown above. By observation, we can see that the entropy distribution for the Firefly class is narrowly centered around an average of 1.3028 (see [Table 4 \(/conference-proceedings-of-spie/13480/3055540/LEAD-AI--lightweight-entropy-analysis-for-distinguishing-AI-generated/10.1117/12.3055540.full#T4\)](#)), showing that there is consistently low complexity in the red channel. However, Firefly exhibits a wider distribution similar to their natural image counterparts, albeit shifted by an entropy value of roughly two both the green and blue channels.

The Stable Diffusion XL red channel curve exhibits a broader distribution with an average of 3.8404. Stable Diffusion XL acts conversely to Firefly, as its distribution becomes more narrow in the green and blue channels, but still remains fairly close to the natural image averages for each channel. The natural image red channel curve shows a mean of 3.4570 and has a broader distribution of variability in the entropy values of every channel. Generally, the natural image distributions follow what one would expect from the images used for training

During classification, Firefly images tended to be more easily distinguished from the rest, while the classifier had trouble recognizing a Stable Diffusion XL image versus a Natural image. Based on our previous discussion, this is most likely due to the similarity in their normal distribution graphs, where the Stable Diffusion XL distribution, while possessing a slightly narrower distribution and higher average in every instance, mimics the shape of the natural image distribution.

The classification results provide further proof that the firefly images are distinct and the stable diffusion XL images are more similar to natural images. The confusion matrix created using the classification results further exemplifies this phenomenon, which can be seen in [Figure 6 \(/conference-proceedings-of-spie/13480/3055540/LEAD-AI--lightweight-entropy-analysis-for-distinguishing-AI-generated/10.1117/12.3055540.full#f6\)](#). The classifier has no problem identifying Firefly images but tends to identify a large portion of Stable Diffusion XL images as belonging to the natural images dataset. [Table 5 \(/conference-proceedings-of-spie/13480/3055540/LEAD-AI--lightweight-entropy-analysis-for-distinguishing-AI-generated/10.1117/12.3055540.full#T5\)](#) presents detailed performance metrics, with the Firefly detector achieving 98.93% accuracy compared to 80.21% for SDXL and natural image detectors.

Fig 6 [Download \(/proceedings/DownloadFigures?url=/ContentImages/Proceedings/13480/134800N/FigureImages/00018_PSiSDG13480_134800N_page_11_1.jpg\)](#)

Confusion Matrix of 3D Gaussian Classification

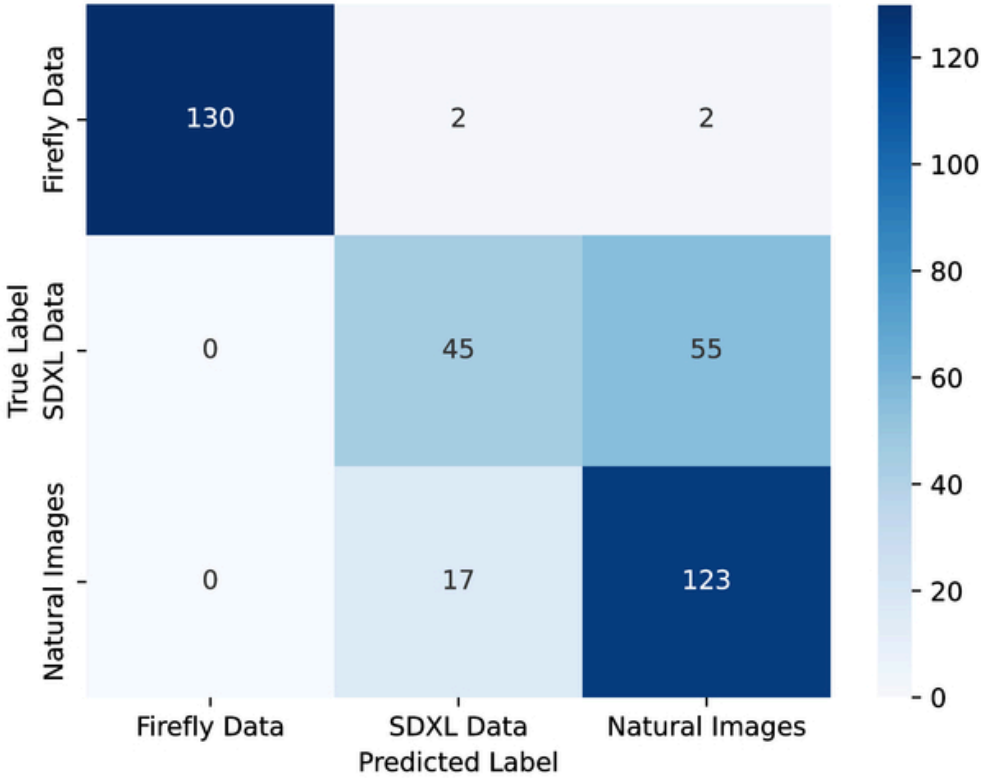


Table 5

Performance Metrics of SDXL, Firefly and Natural images in the Gaussian classifier.

Metric	SDXL	Firefly	Natural
False Positive Rate	0.3167	0.0000	0.2969
False Negative Rate	0.0876	0.0164	0.1774
True Negative Rate	0.8786	0.9701	0.4500

Negative Predictive Value	0.8786	0.9701	0.4500
False Discovery Rate	0.2436	0.0000	0.0693
Recall Score	0.8786	0.9701	0.4500
Precision Score	0.6833	1.0000	0.7031
Accuracy	80.2139%	98.9305%	80.2139%

Further testing with a binary classifier focused solely on distinguishing between the two AI models (Firefly and SDXL) yielded excellent results, achieving 98.54% accuracy as shown in [Figure 7 \(/conference-proceedings-of-spie/13480/3055540/LEAD-AI--lightweight-entropy-analysis-for-distinguishing-AI-generated/10.1117/12.3055540.full#f7\)](#) and [Table 6 \(/conference-proceedings-of-spie/13480/3055540/LEAD-AI--lightweight-entropy-analysis-for-distinguishing-AI-generated/10.1117/12.3055540.full#T6\)](#). These findings demonstrate that entropy analysis can effectively differentiate not only between AI-generated and natural images but also between different AI generation models.

Fig 7 [Download \(/proceedings/DownloadFigures?url=/ContentImages/Proceedings/13480/134800N/FigureImages/00018_PSiSDG13480_134800N_page_12_1.jpg\)](#)

Confusion Matrix of AI model Detector.

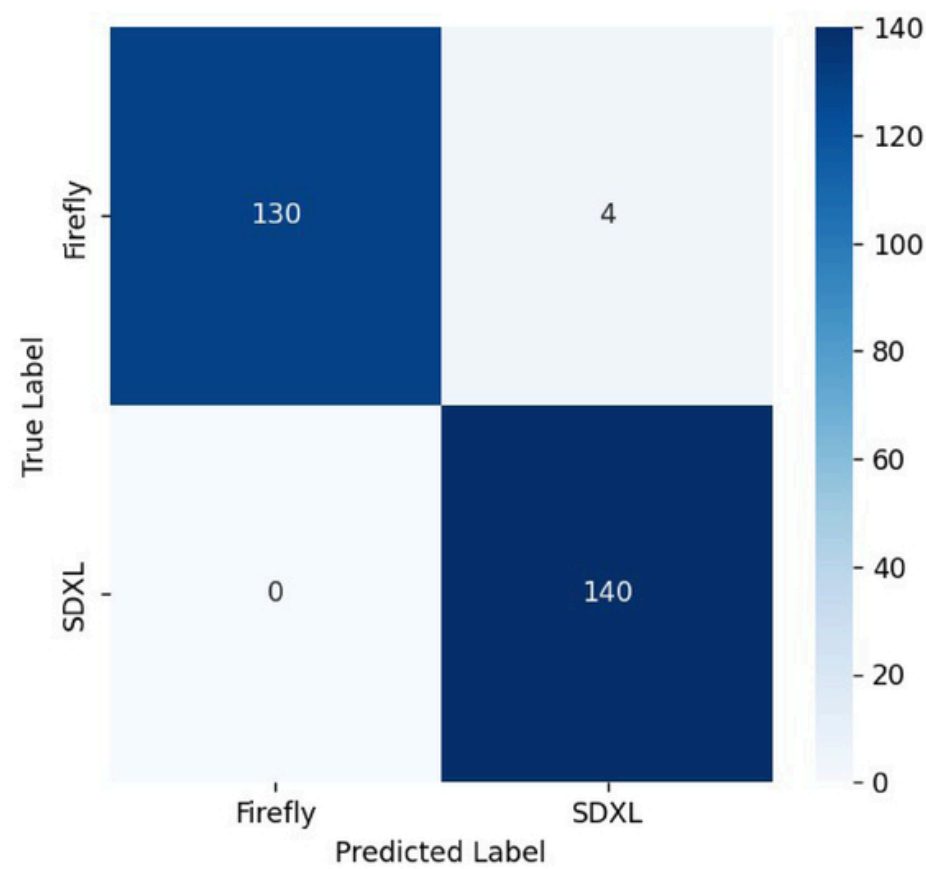


Table 6

Performance Metrics of AI Model Detectors.

Metric	SDXL Detector	Firefly Detector
False Positive Rate	0.0278	0.0000
False Negative Rate	0.0000	0.0278
True Negative Rate	0.9722	1.0000

Negative Predictive Value	1.0000	0.9701
False Discovery Rate	0.0299	0.0000
Recall Score	1.0000	0.9701
Precision Score	0.9722	1.0000
Accuracy	98.5401%	98.5401%

These results validate our approach of using channel-wise entropy as a discriminative feature for image authentication. The distinctive entropy signatures of different image sources provide a reliable foundation for lightweight classification methods, with particular success in identifying images from the Adobe Firefly model. While Stable Diffusion XL images more closely resemble natural images in their entropy characteristics, our method still provides valuable discriminative power with minimal computational requirements.

4.2 A Comparison Discussion

To contextualize LEAD-AI performance, we compare it with two established benchmarks: a lightweight convolutional neural network (CNN) approach introduced by Bird and Lotfi, CIFAKE classifier,⁶ ([/conference-proceedings-of-spie/13480/3055540/LEAD-AI--lightweight-entropy-analysis-for-distinguishing-AI-generated/10.1117/12.3055540.full#c6](#)) and a zero-shot entropy-based method, the Cozzolino et al. ZED detector.²⁹ ([/conference-proceedings-of-spie/13480/3055540/LEAD-AI--lightweight-entropy-analysis-for-distinguishing-AI-generated/10.1117/12.3055540.full#c29](#)) CIFAKE, trained on the CIFAKE dataset (60,000 real CIFAR-10 images and 60,000 synthetic equivalents), achieves a peak accuracy of 93.55% using two convolutional layers of 32 filters followed by one dense layer 64 neurons, with earlier feature extractor results reaching 92.93% (32 filters) and 92.98% (128 filters).⁶ ([/conference-proceedings-of-spie/13480/3055540/LEAD-AI--lightweight-entropy-analysis-for-distinguishing-AI-generated/10.1117/12.3055540.full#c6](#)) While its CNN architecture, executed on an Nvidia RTX 3080Ti GPU, suggests greater computational demand than simpler methods, exact inference times are unreported. In contrast, ZED employs a multiresolution lossless encoder (SReC) trained on real images, achieving a balanced accuracy of 92.2% on the RAISE dataset across various generative models (e.g., Stable Diffusion XL, Adobe Firefly), with AUC scores exceeding 95% for several cases.²⁹ ([/conference-proceedings-of-spie/13480/3055540/LEAD-AI--lightweight-entropy-analysis-for-distinguishing-AI-generated/10.1117/12.3055540.full#c29](#)) Its multi-scale processing implies a moderate computational load, though timings are not specified. LEAD-AI, operating in the spatial domain with a lightweight Gaussian classifier, achieves 98.93% accuracy for Adobe Firefly and 80.21% for Stable Diffusion XL on our dataset, leveraging low-frequency entropy features to minimize computational overhead. ZED excels in generalizability in unseen models without synthetic training data, while LEAD-AI performs both in accuracy for specific models and efficiency, making it ideal for rapid forensic deployment in resource-constrained settings, where the CNN-based approach of CIFAKE may be less practical despite its competitive accuracy.

5 Limitations and Future Work

While the LEAD-AI approach has demonstrated promising results in distinguishing between AI-generated and authentic images, several limitations warrant consideration and provide directions for future research.

The current implementation of LEAD-AI relies on a relatively simple Gaussian classifier, which, while computationally efficient, may not capture more complex decision boundaries that exist between AI-generated and natural images. This limitation becomes particularly evident when analyzing images from models like Stable Diffusion XL, which produce outputs that closely mimic the statistical properties of natural images.

Our evaluation was limited to images from two AI models (Adobe Firefly and Stable Diffusion XL), representing only a subset of the diverse generative technologies currently available. As new models emerge with increasingly sophisticated techniques for photorealistic image synthesis, the entropy patterns identified in this study may evolve, potentially reducing the effectiveness of our current approach.

Additionally, our testing was conducted on clean, unmodified images. In real-world scenarios, images often undergo various post-processing operations such as compression, resizing or filtering, which might alter the entropy characteristics we rely on for detection. The robustness of our method against such transformations requires further investigation.

For future work, we plan to extend LEAD-AI in several directions. First, we intend to apply our entropy analysis approach in the frequency domain, which may reveal additional discriminative features that are not apparent in the spatial domain. Combining spatial and frequency domain analysis could potentially improve detection accuracy while maintaining computational efficiency.

Second, we aim to integrate our lightweight entropy features with more sophisticated machine learning techniques, potentially developing hybrid models that balance the computational efficiency of statistical methods with the discriminative power of deep learning approaches. For example, entropy features could be fed into a shallow neural network optimized for edge devices, balancing efficiency and accuracy. This could include ensemble methods that combine entropy-based classifiers with compact neural networks specifically designed for edge deployment.

Third, we will expand our evaluation to include a broader range of AI image generation models, including newer versions and alternative architectures. This comprehensive evaluation will help assess the generalizability of our approach and identify model-specific patterns that might inform more targeted detection strategies.

Finally, we plan to investigate the resilience of our method against adversarial manipulations specifically designed to evade detection. As detection systems improve, it is likely that generation models will adapt to produce images with more natural entropy distributions, necessitating an ongoing evolution of detection techniques.

By addressing these limitations and pursuing these research directions, we aim to enhance the capabilities of LEAD-AI and contribute to the development of more robust, efficient, and accessible tools for authenticating visual media in an increasingly AI-generated visual landscape.

6 Conclusions

The channel-wise entropy analysis presented in this paper demonstrates that AI-generated and natural images can be effectively differentiated using entropy values as discriminative features. Our LEAD-AI approach provides a statistically sound method for detecting synthetic images while maintaining minimal computational overhead. The key advantages of LEAD-AI include its computational efficiency for real-time detection in resource-constrained environments, high accuracy in distinguishing between different image sources, and interpretability that is often lacking in deep learning methods. The experimental results confirm not only the method's ability to identify AI-generated content but also to differentiate between images produced by different generative models. To the best of our knowledge, this is the first work that leverages the distribution of entropy values across color channels to detect AI-generated images on an individual basis. While state-of-the-art methods may achieve similar accuracy, they typically employ computationally intensive approaches like convolutional neural networks. The lightweight nature of LEAD-AI, coupled with its high accuracy, makes it an ideal candidate for quick and reliable deployment in scenarios where computational resources are limited but authentication needs are critical. As generative models continue to evolve, approaches like LEAD-AI provide a solid foundation for media authentication systems that can serve as a first line of defense against AI-generated misinformation. Future work will focus on enhancing the method's robustness across a broader range of generative models and exploring hybrid approaches that maintain efficiency while adapting to the evolving landscape of synthetic media.

References

- 1 G. Bansal, A. Nawal, V. Chamola, et al., "Revolutionizing visuals: the role of generative ai in modern image generation," ACM Transactions on Multimedia Computing, Communications and Applications, 20 (11), 1 –22 (2024). <https://doi-org.ez9.periodicos.capes.gov.br/10.1145/3689641> (<https://doi-org.ez9.periodicos.capes.gov.br/10.1145/3689641>) [Google Scholar \(http://scholar.google.com.ez9.periodicos.capes.gov.br/scholar_lookup?title=Revolutionizing+visuals:+the+role+of+generative+ai+in+modern+image+generation&author=G.+Bansal&author=A.+Nawal&author=V.+Chamola&journal=ACM+Transactions+on+Multimedia+Computing,+Communications+and+Applications&volume=20&issue=11&publication_year=2024&pages=1-22\)](http://scholar.google.com.ez9.periodicos.capes.gov.br/scholar_lookup?title=Revolutionizing+visuals:+the+role+of+generative+ai+in+modern+image+generation&author=G.+Bansal&author=A.+Nawal&author=V.+Chamola&journal=ACM+Transactions+on+Multimedia+Computing,+Communications+and+Applications&volume=20&issue=11&publication_year=2024&pages=1-22).
- 2 H. Cao, C. Tan, Z. Gao, et al., "A survey on generative diffusion models," IEEE Transactions on Knowledge and Data Engineering, (2024). <https://doi-org.ez9.periodicos.capes.gov.br/10.1109/TKDE.2024.3361474> (<https://doi-org.ez9.periodicos.capes.gov.br/10.1109/TKDE.2024.3361474>) [Google Scholar \(http://scholar.google.com.ez9.periodicos.capes.gov.br/scholar_lookup?title=A+survey+on+generative+diffusion+models&author=H.+Cao&author=C.+Tan&author=Z.+Gao&journal=IEEE+Transactions+on+Knowledge+and+Data+Engineering&publication_year=2024\)](http://scholar.google.com.ez9.periodicos.capes.gov.br/scholar_lookup?title=A+survey+on+generative+diffusion+models&author=H.+Cao&author=C.+Tan&author=Z.+Gao&journal=IEEE+Transactions+on+Knowledge+and+Data+Engineering&publication_year=2024).
- 3 I. Amerini, M. Barni, S. Battiato, et al., "Deepfake media forensics: Status and future challenges," Journal of Imaging, 11 (3), 73 (2025). <https://doi-org.ez9.periodicos.capes.gov.br/10.3390/jimaging11030073> (<https://doi-org.ez9.periodicos.capes.gov.br/10.3390/jimaging11030073>) [Google Scholar \(http://scholar.google.com.ez9.periodicos.capes.gov.br/scholar_lookup?title=Deepfake+media+forensics:+Status+and+future+challenges&author=I.+Amerini&author=M.+Barni&author=S.+Battiato&journal=Journal+of+Imaging&publication_year=2025\)](http://scholar.google.com.ez9.periodicos.capes.gov.br/scholar_lookup?title=Deepfake+media+forensics:+Status+and+future+challenges&author=I.+Amerini&author=M.+Barni&author=S.+Battiato&journal=Journal+of+Imaging&publication_year=2025).

http://scholar.google.com.ez9.periodicos.capes.gov.br/scholar_lookup?

[title=Deepfake+media+forensics:+Status+and+future+challenges&author=L.+Amerini&author=M.+Barni&author=S.+Battiato&journal=Journal+of+Imaging&volume=11&issue=3&publication_year=2025&pages=73](http://scholar.google.com.ez9.periodicos.capes.gov.br/scholar_lookup?title=Deepfake+media+forensics:+Status+and+future+challenges&author=L.+Amerini&author=M.+Barni&author=S.+Battiato&journal=Journal+of+Imaging&volume=11&issue=3&publication_year=2025&pages=73)

4 X. Yu, Y. Wang, Y. Chen, et al., “Fake artificial intelligence generated contents (faigc): A survey of theories, detection methods, and opportunities,” arXiv preprint arXiv:2405.00711, (2024). [Google Scholar \(http://scholar.google.com.ez9.periodicos.capes.gov.br/scholar_lookup?title=Fake+artificial+intelligence+generated+contents+\(faigc\):+A+survey+of+theories,+detection+methods,+and+opportunities&author=X.+Yu&author=Y.+Wang&author=Y.+Chen&journal=arXiv+preprint+arXiv:2405.00711&publication_year=2024\)](http://scholar.google.com.ez9.periodicos.capes.gov.br/scholar_lookup?title=Fake+artificial+intelligence+generated+contents+(faigc):+A+survey+of+theories,+detection+methods,+and+opportunities&author=X.+Yu&author=Y.+Wang&author=Y.+Chen&journal=arXiv+preprint+arXiv:2405.00711&publication_year=2024)

5 A. Boutadjine, F. Harrag, and K. Shaalan, “Human vs. machine: A comparative study on the detection of ai-generated content,” ACM Transactions on Asian and Low-Resource Language Information Processing, 24 (2), 1 –26 (2025). <https://doi-org.ez9.periodicos.capes.gov.br/10.1145/3708889> (<https://doi-org.ez9.periodicos.capes.gov.br/10.1145/3708889>) [Google Scholar \(http://scholar.google.com.ez9.periodicos.capes.gov.br/scholar_lookup?title=Human+vs.+machine:+A+comparative+study+on+the+detection+of+ai-generated+content&author=A.+Boutadjine&author=F.+Harrag&author=K.+Shaalan&journal=ACM+Transactions+on+Asian+and+Low-Resource+Language+Information+Processing&volume=24&issue=2&publication_year=2025&pages=1-26\)](http://scholar.google.com.ez9.periodicos.capes.gov.br/scholar_lookup?title=Human+vs.+machine:+A+comparative+study+on+the+detection+of+ai-generated+content&author=A.+Boutadjine&author=F.+Harrag&author=K.+Shaalan&journal=ACM+Transactions+on+Asian+and+Low-Resource+Language+Information+Processing&volume=24&issue=2&publication_year=2025&pages=1-26)

6 J. J. Bird and A. Lotfi, “Cifake: Image classification and explainable identification of aigenerated synthetic images,” IEEE Access, 12 15642 –15650 (2024). <https://doi-org.ez9.periodicos.capes.gov.br/10.1109/ACCESS.2024.3356122> (<https://doi-org.ez9.periodicos.capes.gov.br/10.1109/ACCESS.2024.3356122>) [Google Scholar \(http://scholar.google.com.ez9.periodicos.capes.gov.br/scholar_lookup?title=Cifake:+Image+classification+and+explainable+identification+of+aigenerated+synthetic+images&author=J.+J.+Bird&author=A.+Lotfi&journal=IEEE+Access&volume=12&publication_year=2024&pages=15642-15650\)](http://scholar.google.com.ez9.periodicos.capes.gov.br/scholar_lookup?title=Cifake:+Image+classification+and+explainable+identification+of+aigenerated+synthetic+images&author=J.+J.+Bird&author=A.+Lotfi&journal=IEEE+Access&volume=12&publication_year=2024&pages=15642-15650)

7 J. Frank, T. Eisenhofer, L. Schönherr, et al., “Leveraging frequency analysis for deep fake image recognition,” in International conference on machine learning, 3247 –3258 (2020). [Google Scholar \(http://scholar.google.com.ez9.periodicos.capes.gov.br/scholar_lookup?title=Leveraging+frequency+analysis+for+deep+fake+image+recognition&author=J.+Frank&author=T.+Eisenhofer&author=L.+Schönherr&conference=International+conference+on+machine+learning&publication_year=2020&pages=3247-3258\)](http://scholar.google.com.ez9.periodicos.capes.gov.br/scholar_lookup?title=Leveraging+frequency+analysis+for+deep+fake+image+recognition&author=J.+Frank&author=T.+Eisenhofer&author=L.+Schönherr&conference=International+conference+on+machine+learning&publication_year=2020&pages=3247-3258)

8 L. Lin, N. Gupta, Y. Zhang, et al., “Detecting multimedia generated by large ai models: A survey,” arXiv preprint arXiv:2402.00045, (2024). [Google Scholar \(http://scholar.google.com.ez9.periodicos.capes.gov.br/scholar_lookup?](http://scholar.google.com.ez9.periodicos.capes.gov.br/scholar_lookup?)

[title=Detecting+multimedia+generated+by+large+ai+models:+A+survey&author=L.+Lin&author=N.+Gupta&author=Y.+Zhang&journal=arXiv+preprint+arXiv:2402.00045&publication_year=2024](#)).

9 N. Poredi, M. Sudarsan, E. Solomon, et al., “Generative adversarial networks-based aigenerated imagery authentication using frequency domain analysis,” Disruptive Technologies Information Sciences VIII, 13058 376 –390 SPIE(2024). [Google Scholar \(http://scholar.google.com.ez9.periodicos.capes.gov.br/scholar_lookup?title=Generative+adversarial+network+based+aigenerated+imagery+authentication+using+frequency+domain+analysis&author=N.+Poredi&author=M.+Sudarsan&author=E.+Solomon&volume=13058&publication_year=2024&pages=376-390\)](#).

10 N. Poredi, D. Nagothu, and Y. Chen, “Ausome: authenticating social media images using frequency analysis,” Disruptive Technologies in Information Sciences VII, 12542 44 –56 SPIE(2023). [Google Scholar \(http://scholar.google.com.ez9.periodicos.capes.gov.br/scholar_lookup?title=Ausome:+authenticating+social+media+images+using+frequency+analysis&author=N.+Poredi&author=D.+Nagothu&author=Y.+Chen&volume=12542&publication_year=2023&pages=44-56\)](#).

11 I. Alam, M. S. Muneer, and S. S. Woo, “Ugad: Universal generative ai detector utilizing frequency fingerprints,” in Proceedings of the 33rd ACM International Conference on Information and Knowledge Management, 4332 –4340 (2024). [Google Scholar \(http://scholar.google.com.ez9.periodicos.capes.gov.br/scholar_lookup?title=Ugad:+Universal+generative+ai+detector+utilizing+frequency+fingerprints&author=I.+Alam&author=M.+S.+Muneer&author=S.+S.+Woo&conference=Proceedings+of+the+33rd+International+Conference+on+Information+and+Knowledge+Management&publication_year=2024&pages=4332-4340\)](#).

12 I. Goodfellow, J. Pouget-Abadie, M. Mirza, et al., “Generative adversarial networks,” Communications of the ACM, 63 (11), 139 –144 (2020). <https://doi-org.ez9.periodicos.capes.gov.br/10.1145/3422622> (<https://doi-org.ez9.periodicos.capes.gov.br/10.1145/3422622>) [Google Scholar \(http://scholar.google.com.ez9.periodicos.capes.gov.br/scholar_lookup?title=Generative+adversarial+networks&author=I.+Goodfellow&author=J.+Pouget-Abadie&author=M.+Mirza&journal=Communications+of+the+ACM&volume=63&issue=11&publication_year=2020&pages=139-144\)](#).

13 I. J. Goodfellow, J. Pouget-Abadie, M. Mirza, et al., “Generative adversarial nets,” Advances in neural information processing systems, 27 (2014). [Google Scholar \(http://scholar.google.com.ez9.periodicos.capes.gov.br/scholar_lookup?title=Generative+adversarial+nets&author=I.+J.+Goodfellow&author=J.+Pouget-Abadie&author=M.+Mirza&journal=Advances+in+neural+information+processing+systems&volume=27&publication_year=2014\)](#).

14 A. Ramesh, M. Pavlov, G. Goh, et al., “Zero-shot text-to-image generation,” in International conference on machine learning, 8821 –8831 (2021). [Google Scholar \(http://scholar.google.com.ez9.periodicos.capes.gov.br/scholar_lookup?title=Zero-shot+text-to-](#)

image+generation&author=A.+Ramesh&author=M.+Pavlov&author=G.+Goh&conference=International+conference+on+machine+learning&publication_year=2021&pages=8821-8831

15 J. Ho, A. Jain, and P. Abbeel, "Denoising diffusion probabilistic models," Advances in neural information processing systems, 33 6840 –6851 (2020). [Google Scholar \(http://scholar.google.com.ez9.periodicos.capes.gov.br/scholar_lookup?title=Denoising+diffusion+probabilistic+models&author=J.+Ho&author=A.+Jain&author=P.+Abbeel&journal=Advances+in+neural+information+processing+systems&volume=33&publication_year=2020&pages=6840-6851\)](https://scholar.google.com.ez9.periodicos.capes.gov.br/scholar_lookup?title=Denoising+diffusion+probabilistic+models&author=J.+Ho&author=A.+Jain&author=P.+Abbeel&journal=Advances+in+neural+information+processing+systems&volume=33&publication_year=2020&pages=6840-6851)

16 S. Srdarov and T. Leaver, "Generative ai glitches: The artificial everything," M/C Journal, 27 (6), (2024). <https://doi-org.ez9.periodicos.capes.gov.br/10.5204/mcj.3123> (<https://doi-org.ez9.periodicos.capes.gov.br/10.5204/mcj.3123>) [Google Scholar \(http://scholar.google.com.ez9.periodicos.capes.gov.br/scholar_lookup?title=Generative+ai+glitches:+The+artificial+everything&author=S.+Srdarov&author=T.+Leaver&journal=M/C+Journal&volume=27&issue=6&publication_year=2024\)](http://scholar.google.com.ez9.periodicos.capes.gov.br/scholar_lookup?title=Generative+ai+glitches:+The+artificial+everything&author=S.+Srdarov&author=T.+Leaver&journal=M/C+Journal&volume=27&issue=6&publication_year=2024)

17 P. Esser, S. Kulal, A. Blattmann, et al., "Scaling rectified flow transformers for highresolution image synthesis," in Forty-first international conference on machine learning, (2024). [Google Scholar \(http://scholar.google.com.ez9.periodicos.capes.gov.br/scholar_lookup?title=Scaling+rectified+flow+transformers+for+highresolution+image+synthesis&author=P.+Esser&author=S.+Kulal&author=A.+Blattmann&conference=Forty-first+international+conference+on+machine+learning&publication_year=2024\)](http://scholar.google.com.ez9.periodicos.capes.gov.br/scholar_lookup?title=Scaling+rectified+flow+transformers+for+highresolution+image+synthesis&author=P.+Esser&author=S.+Kulal&author=A.+Blattmann&conference=Forty-first+international+conference+on+machine+learning&publication_year=2024)

18 Y. Wei and G. Tyson, "Understanding the impact of ai-generated content on social media: The pixiv case," in Proceedings of the 32nd ACM International Conference on Multimedia 6813 –6822 (2024). [Google Scholar \(http://scholar.google.com.ez9.periodicos.capes.gov.br/scholar_lookup?title=Understanding+the+impact+of+ai-generated+content+on+social+media:+The+pixiv+case&author=Y.+Wei&author=G.+Tyson&conference=Proceedings+of+the+32nd+ACM+International+Conference+on+Multimedia&publication_year=2024&pages=6813-6822\)](http://scholar.google.com.ez9.periodicos.capes.gov.br/scholar_lookup?title=Understanding+the+impact+of+ai-generated+content+on+social+media:+The+pixiv+case&author=Y.+Wei&author=G.+Tyson&conference=Proceedings+of+the+32nd+ACM+International+Conference+on+Multimedia&publication_year=2024&pages=6813-6822)

19 G. Tang and S. E. Eaton, "A rapid investigation of artificial intelligence generated content footprints in scholarly publications," Journal of Scholarly Publishing, 55 (3), 337 –355 (2024). <https://doi-org.ez9.periodicos.capes.gov.br/10.3138/jsp-2023-0079> (<https://doi-org.ez9.periodicos.capes.gov.br/10.3138/jsp-2023-0079>) [Google Scholar \(http://scholar.google.com.ez9.periodicos.capes.gov.br/scholar_lookup?title=A+rapid+investigation+of+artificial+intelligence+generated+content+footprints+in+scholarly+publications&author=G.+Tang&author=S.+E.+Eaton&journal=Journal+of+Scholarly+Publishing&volume=55&issue=3&publication_year=2024&pages=337-355\)](http://scholar.google.com.ez9.periodicos.capes.gov.br/scholar_lookup?title=A+rapid+investigation+of+artificial+intelligence+generated+content+footprints+in+scholarly+publications&author=G.+Tang&author=S.+E.+Eaton&journal=Journal+of+Scholarly+Publishing&volume=55&issue=3&publication_year=2024&pages=337-355)

20 P. Zhao, H. Zhang, Q. Yu, et al., "Retrieval-augmented generation for ai-generated content: A survey," arXiv preprint arXiv:2402.19473, (2024). [Google Scholar \(http://scholar.google.com.ez9.periodicos.capes.gov.br/scholar_lookup?title=Retrieval-augmented+generation+for+ai-](http://scholar.google.com.ez9.periodicos.capes.gov.br/scholar_lookup?title=Retrieval-augmented+generation+for+ai-)

generated+content:+A+survey&author=P.+Zhao&author=H.+Zhang&author=Q.+Yu&journal=arXiv+preprint+arXiv:2402.19473&publication_year=2024).

21 X. Zhang, S. Karaman, and S.-F. Chang, "Detecting and simulating artifacts in gan fake images," in 2019 IEEE international workshop on information forensics and security (WIFS), 1–6 (2019). [Google Scholar \(http://scholar.google.com.ez9.periodicos.capes.gov.br/scholar_lookup?title=Detecting+and+simulating+artifacts+in+gan+fake+images&author=X.+Zhang&author=S.+Karaman&author=S.-F.+Chang&conference=2019+IEEE+international+workshop+on+information+forensics+and+security+\(WIFS\)&publication_year=2019&pages=1-6\)](http://scholar.google.com.ez9.periodicos.capes.gov.br/scholar_lookup?title=Detecting+and+simulating+artifacts+in+gan+fake+images&author=X.+Zhang&author=S.+Karaman&author=S.-F.+Chang&conference=2019+IEEE+international+workshop+on+information+forensics+and+security+(WIFS)&publication_year=2019&pages=1-6).

22 A. Pazytkarim, D. Nagothu, and Y. Chen, "A lightweight deep learning model for rapid detection of fabricated enf signals from audio sources," Disruptive Technologies in Information Sciences VIII, 13058 363–375 SPIE(2024). [Google Scholar \(http://scholar.google.com.ez9.periodicos.capes.gov.br/scholar_lookup?title=A+lightweight+deep+learning+model+for+rapid+detection+of+fabricated+enf+signals+from+audio+sources&author=A.+Pazytkarim&author=D.+Nagothu&author=Y.+Chen&volume=13058&publication_year=2024&pages=363-375\)](http://scholar.google.com.ez9.periodicos.capes.gov.br/scholar_lookup?title=A+lightweight+deep+learning+model+for+rapid+detection+of+fabricated+enf+signals+from+audio+sources&author=A.+Pazytkarim&author=D.+Nagothu&author=Y.+Chen&volume=13058&publication_year=2024&pages=363-375).

23 D. Nagothu, Y. Chen, E. Blasch, et al., "Detecting malicious false frame injection attacks on surveillance systems at the edge using electrical network frequency signals," Sensors, (11), 2424 (2019). <https://doi-org.ez9.periodicos.capes.gov.br/10.3390/s19112424> (<https://doi-org.ez9.periodicos.capes.gov.br/10.3390/s19112424>) [Google Scholar \(http://scholar.google.com.ez9.periodicos.capes.gov.br/scholar_lookup?title=Detecting+malicious+false+frame+injection+attacks+on+surveillance+systems+at+the+edge+using+electrical+network+frequency+signals&author=D.+Nagothu&author=Y.+Chen&author=E.+Blasch&journal=Sensors&volume=19&issue=11&publication_year=2019&pages=2424\)](http://scholar.google.com.ez9.periodicos.capes.gov.br/scholar_lookup?title=Detecting+malicious+false+frame+injection+attacks+on+surveillance+systems+at+the+edge+using+electrical+network+frequency+signals&author=D.+Nagothu&author=Y.+Chen&author=E.+Blasch&journal=Sensors&volume=19&issue=11&publication_year=2019&pages=2424).

24 D. Nagothu, R. Xu, Y. Chen, et al., "Defakepro: Decentralized deepfake attacks detection using enf authentication," IT Professional, 24 (5), 46–52 (2022). <https://doi-org.ez9.periodicos.capes.gov.br/10.1109/MITP.2022.3172653> (<https://doi-org.ez9.periodicos.capes.gov.br/10.1109/MITP.2022.3172653>) [Google Scholar \(http://scholar.google.com.ez9.periodicos.capes.gov.br/scholar_lookup?title=Defakepro:+Decentralized+deepfake+attacks+detection+using+enf+authentication&author=D.+Nagothu&author=R.+Xu&author=Y.+Chen&journal=IT+Professional&volume=24&issue=5&publication_year=2022&pages=46-52\)](http://scholar.google.com.ez9.periodicos.capes.gov.br/scholar_lookup?title=Defakepro:+Decentralized+deepfake+attacks+detection+using+enf+authentication&author=D.+Nagothu&author=R.+Xu&author=Y.+Chen&journal=IT+Professional&volume=24&issue=5&publication_year=2022&pages=46-52).

25 N. Poredi, D. Nagothu, Y. Chen, et al., "Robustness of electrical network frequency signals as a fingerprint for digital media authentication," in 2022 IEEE 24th International Workshop on Multimedia Signal Processing (MMSP), 1–6 (2022). [Google Scholar \(http://scholar.google.com.ez9.periodicos.capes.gov.br/scholar_lookup?title=Robustness+of+electrical+network+frequency+signals+as+a+fingerprint+for+digital+media+authentication&author=N.+Poredi&author=D.+Nagothu&author=Y.+Chen&conference=2022+IEEE+24th+International+Workshop+on+Multimedia+Signal+Processing+\(MMSP\)&publication_year=2022&pages=1-6\)](http://scholar.google.com.ez9.periodicos.capes.gov.br/scholar_lookup?title=Robustness+of+electrical+network+frequency+signals+as+a+fingerprint+for+digital+media+authentication&author=N.+Poredi&author=D.+Nagothu&author=Y.+Chen&conference=2022+IEEE+24th+International+Workshop+on+Multimedia+Signal+Processing+(MMSP)&publication_year=2022&pages=1-6).

26

D. Nagothu, R. Xu, Y. Chen, et al., "Deterring deepfake attacks with an electrical network frequency fingerprints approach," Future Internet, 14 (5), 125 (2022). <https://doi-org.ez9.periodicos.capes.gov.br/10.3390/fi14050125> (<https://doi-org.ez9.periodicos.capes.gov.br/10.3390/fi14050125>) [Google Scholar](#) (http://scholar.google.com.ez9.periodicos.capes.gov.br/scholar_lookup?title=Deterring+deepfake+attacks+with+an+electrical+network+frequency+fingerprints+approach&author=D.+Nagothu&author=R.+Xu&author=Y.+Chen&journal=Future+Internet&volume=14&issue=5&publication_year=2022&pages=125).

27

C. Shannon, "Claude shannon," Information Theory, 3 224 (1948). [Google Scholar](#) (http://scholar.google.com.ez9.periodicos.capes.gov.br/scholar_lookup?title=Claude+shannon&author=C.+Shannon&journal=Information+Theory&volume=3&publication_year=1948&pages=224).

28

D. Ruderman, "The statistics of natural images," NETWORK-BRISTOL- 6, 105 –105 (1995). [Google Scholar](#) (http://scholar.google.com.ez9.periodicos.capes.gov.br/scholar_lookup?title=The+statistics+of+natural+images&author=D.+Ruderman&journal=NETWORK-BRISTOL-+6&publication_year=1995&pages=105-105).

29

D. Cozzolino, G. Poggi, M. Nießner, et al., "Zero-shot detection of ai-generated images," in European Conference on Computer Vision, 54 –72 (2024). [Google Scholar](#) (http://scholar.google.com.ez9.periodicos.capes.gov.br/scholar_lookup?title=Zero-shot+detection+of+ai-generated+images&author=D.+Cozzolino&author=G.+Poggi&author=M.+Nießner&conference=European+Conference+on+Computer+Vision&publication_year=2024&pages=54-72).

(2025) Published by SPIE. Downloading of the abstract is permitted for personal use only.

Citation [Download Citation](#) ▼

Monica Sudarsan, Nihal A. Poredi, Evan Maurer, Enoch Solomon, and Yu Chen (/profile/Yu.Chen-47294)
"LEAD-AI: lightweight entropy analysis for distinguishing AI-generated images from genuine photographs",
Proc. SPIE 13480, Disruptive Technologies in Information Sciences IX, 134800N (21 May 2025);
<https://doi-org.ez9.periodicos.capes.gov.br/10.1117/12.3055540> (<https://doi-org.ez9.periodicos.capes.gov.br/10.1117/12.3055540>).

