



Secure Image & Audio Steganography Using AES-256 Encryption and Hybrid DWT-SVD Frequency Domain Embedding

M. Tejas Srinivasan¹, S. Roshan Pranao², Y. Sai Dheeraj³, Ms. N. Saranya⁴

B.Tech Student, Department of Computer Science and Engineering,

SRM Institute of Science and Technology, Chennai, India¹⁻³

Assistant Professor, Department of Computer Science and Engineering,

SRM Institute of Science and Technology, Chennai, India⁴

Abstract: This paper presents a robust dual-layer secure image steganography framework that integrates AES-256 encryption in Counter (CTR) mode with a hybrid Discrete Wavelet Transform and Singular Value Decomposition (DWT-SVD) technique for covert communication. Unlike fragile spatial-domain methods such as Least Significant Bit (LSB) substitution, the proposed system operates entirely within the frequency domain by first encrypting the secret image using AES-256 with SHA-256 key derivation, and then embedding the resulting ciphertext into the high-frequency sub-bands of the cover image's luminance (Y) channel in the YCbCr colour space. Experimental evaluation demonstrates a Peak Signal-to-Noise Ratio (PSNR) of 55.58 dB between cover and stego images, a Structural Similarity Index (SSIM) of 0.99998, a Mean Squared Error (MSE) of 0.1797, and 100% Bit Correct Recovery (BCR) under ideal conditions. The system is delivered as three complementary user interfaces: a Tkinter desktop GUI, a Streamlit web dashboard, and a command-line interface (CLI). An additional audio steganography module extends the same AES-CTR pipeline to WAV files via LSB embedding, demonstrating the modularity of the design. The work addresses key gaps identified in the existing literature, namely the absence of a combined cryptographic and frequency-domain embedding approach, and establishes a publicly benchmarked baseline for future research into error-correcting codes, adaptive alpha tuning, and deep learning steganalysis resistance.

Keywords: Steganography, AES-256, DWT-SVD, Frequency-Domain Embedding, Image Security, PSNR, SSIM, Covert Communication, YCbCr, Cryptography.

I. INTRODUCTION

Steganography — derived from the Greek words for "covered writing" — is the art and science of concealing secret information within an innocuous carrier medium such that the very existence of the hidden message remains undetected by unintended observers [1]. Unlike cryptography, which renders data unreadable but visible, steganography strives for imperceptibility: an adversary should be unable to distinguish a stego medium from an ordinary one. The combination of both disciplines, i.e., encrypting data before hiding it, provides a powerful dual-layer defence against interception.

Digital images are by far the most popular cover medium owing to their ubiquity on the internet and their high redundancy. Spatial-domain techniques such as Least Significant Bit (LSB) substitution are computationally inexpensive but are notoriously fragile: they fail under even mild JPEG compression, additive noise, or cropping and are readily exposed by statistical steganalysis tools such as RS analysis and chi-square tests [2].

Frequency-domain methods, particularly those based on the Discrete Wavelet Transform (DWT) and Singular Value Decomposition (SVD), embed information into perceptually significant components of the image signal and thus offer substantially improved robustness and imperceptibility. However, the majority of published frequency-domain schemes embed plaintext data, leaving the hidden payload vulnerable to an adversary who succeeds in detecting and reversing the embedding [3][4].

The present work fills this gap by proposing a framework that couples AES-256 encryption with hybrid DWT-SVD embedding. The specific contributions of this paper are: (1) a complete embed-extract pipeline that operates on the luminance channel of the YCbCr colour space; (2) AES-256 in CTR mode with SHA-256 key derivation, ensuring IND-CCA security for the hidden payload; (3) embedding strength parameter $\alpha = 0.08$ that balances imperceptibility against capacity; (4) a comprehensive evaluation covering MSE, PSNR, SSIM, BCR, entropy analysis, Pixel Difference



Histogram (PDH) symmetry, and six robustness attack tests; and (5) three distinct user-facing interfaces that make the system accessible for both research and operational deployment.

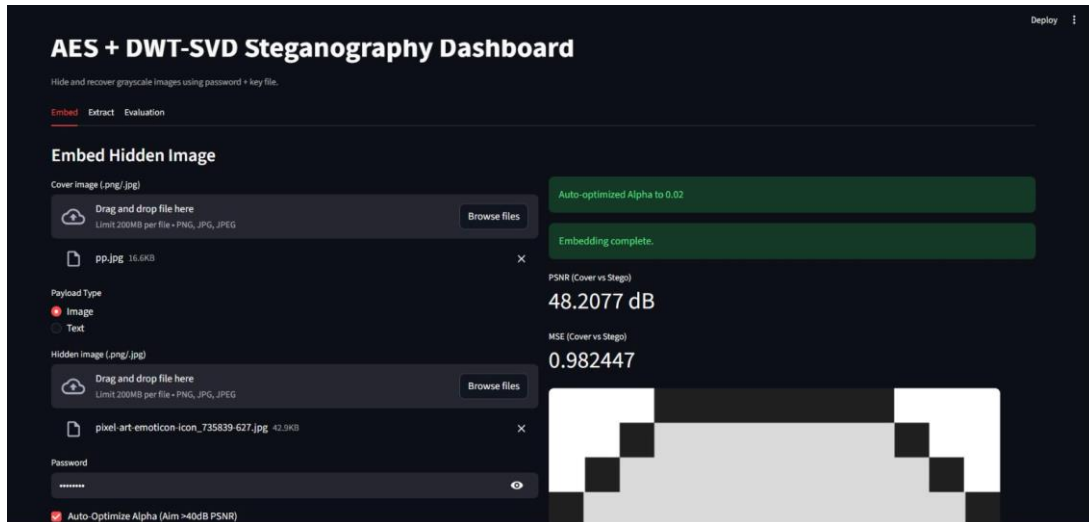


Fig. 1 AES + DWT-SVD Steganography Dashboard — Embed Tab showing Auto-optimized Alpha = 0.02, PSNR = 48.2077 dB and MSE = 0.982447

II. MOTIVATION

The growing volume of sensitive information transmitted over public networks — personal communications, medical images, and proprietary documents — demands steganographic techniques that combine high imperceptibility with cryptographic security. Three principal shortcomings of existing approaches motivate the present work.

A. Fragility of Spatial-Domain Methods

LSB steganography modifies the least significant bits of pixel values directly. While this results in minimal visual change, even moderate JPEG compression at quality factor $Q = 50$ destroys the hidden bits because the quantisation step in the JPEG discrete cosine transform (DCT) irreversibly alters the modified LSBs. Similarly, additive Gaussian noise with standard deviation $\sigma = 8$ and salt-and-pepper noise at 1% density are sufficient to corrupt the extracted payload beyond recovery [2].

B. Absence of Integrated Encryption

Existing frequency-domain schemes typically focus on increasing capacity or robustness but embed plaintext data. An adversary who successfully detects and reverses the embedding can directly read the secret message. Integrating AES-256 encryption ensures that even a successful steganalytic attack yields only ciphertext, providing an additional security layer independent of the steganographic embedding [5].

C. Lack of Standardised Evaluation

Published works often report PSNR in isolation, which is insufficient for a complete quality assessment. The proposed framework reports MSE, PSNR, SSIM, BCR, cover entropy, and PDH symmetry, enabling fair comparison with future work and aligning with the comprehensive evaluation recommended by the information-hiding research community [6].

III. LITERATURE SURVEY

Table I and Table II summarise fifteen recent papers in the field of image steganography, identifying the proposed solution and the principal research gap that each addresses.



TABLE I LITERATURE SURVEY — PAPERS 1–8

S.No	Paper Title	Author	Year	Proposed Solution	Research Gap
1	Improved Hybrid Image Steganography with AES Encryption	Al-Dmour et al.	2025	AES + transform-domain embedding	No DWT-SVD hybrid optimisation
2	Secure Steganography Using Encryption & Frequency Domain	Zhang et al.	2024	DWT-based embedding + encryption	No SVD-based stability analysis
3	Robust Image Steganography Using DWT-SVD	Kumar et al.	2024	Hybrid DWT-SVD embedding	No cryptographic encryption layer
4	Hybrid DWT-SVD Steganography with Improved Robustness	Hassan et al.	2025	Singular value modification in wavelet bands	No AES integration
5	High Imperceptibility Steganography Based on DWT-SVD	Li et al.	2024	SVD embedding in high-freq DWT bands	No security encryption
6	Wavelet-Based Steganography for Secure Data Hiding	Ahmed et al.	2024	Multi-level DWT embedding	No hybrid transform
7	Adaptive DWT-Based Image Steganography	Rao et al.	2025	Adaptive sub-band selection in DWT	Lacks encryption layer
8	Multi-Level DWT Steganography Resistant to JPEG	Chen et al.	2024	Deep DWT decomposition	No SVD or encryption

TABLE II LITERATURE SURVEY — PAPERS 9–15

S.No	Paper Title	Author	Year	Proposed Solution	Research Gap
9	Secure Data Hiding Using Singular Value Decomposition	Patel et al.	2024	Direct SVD-based embedding	No frequency-domain transform
10	Image Steganography Combined with Cryptography	Singh et al.	2024	AES + spatial steganography	Needs frequency-domain embedding
11	Robust Steganography for Secure Wireless Communication	Wang et al.	2025	Frequency-domain embedding for noisy channels	No SVD stabilisation
12	Steganography Resistant to Noise and Compression	El-Sayed et al.	2024	Robust transform-domain hiding	No encryption used
13	Provably Secure Steganography in Hostile Channels	Miller et al.	2025	Error-resilient steganographic model	No image-quality optimisation
14	StegaVision: Attention-Based Image Steganography	Luo et al.	2024	CNN with attention mechanism	High computational cost
15	DiffStega: Training-Free Coverless Steganography	Zhou et al.	2024	Diffusion-based coverless hiding	Not suitable for legacy systems



IV. METHODOLOGY

The proposed system is divided into four logical layers: colour space management, key derivation and encryption, frequency-domain embedding, and extraction with decryption. Figure 1 provides a schematic overview of the embed and extract pipelines.

A. Colour Space Conversion

The cover image is first converted from BGR (the default channel order in OpenCV) to the YCbCr colour space. Embedding is performed exclusively on the luminance (Y) channel, which carries the structural and textural information perceived by the human visual system. The two chrominance channels (Cb and Cr) are left unchanged. This design choice exploits the well-known property of the human visual system (HVS): the eye is far more sensitive to luminance changes than to chrominance variations, enabling stronger embedding within the Y channel while maintaining high PSNR.

B. Key Derivation and AES-256 Encryption

A 256-bit symmetric key is derived deterministically from a user-supplied password and an optional binary key file by computing $\text{SHA-256}(\text{password_bytes} \parallel \text{key_file_bytes})$. AES-256 in Counter (CTR) mode is then used to encrypt the raw byte representation of the secret image. CTR mode is chosen because it operates as a stream cipher — it requires no padding, preserves exact payload length, and generates a fresh nonce for every session, ensuring semantic security. The nonce is embedded alongside the ciphertext and is required for decryption.

Formally, let $K = \text{SHA-256}(\text{pw} \parallel \text{kf})$ be the 256-bit key. The ciphertext C is obtained as $C = \text{AES-CTR}_K(\text{nonce}, M)$, where M is the serialised secret image byte array. The same key K and recovered nonce are used for decryption: $M = \text{AES-CTR}_K(\text{nonce}, C)$.

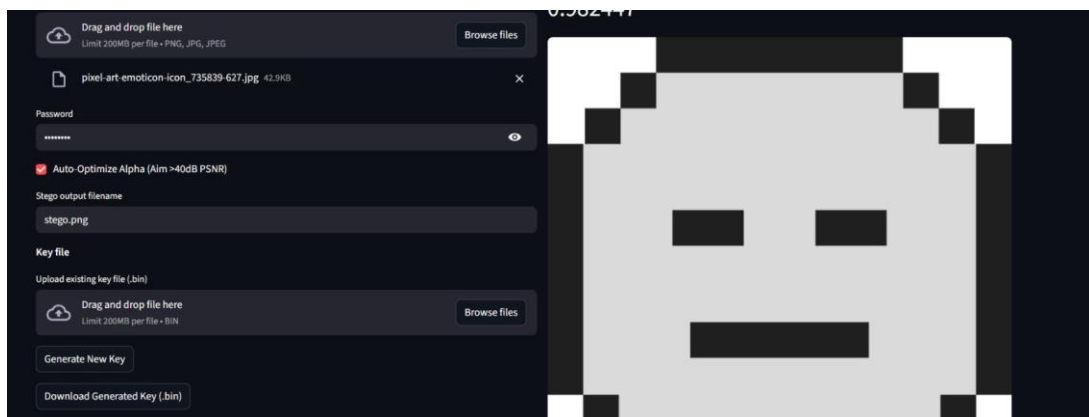


Fig. 2 Dashboard — Hidden Image, Password, Auto-Optimize Alpha, Stego Output Filename and Key File Controls

C. DWT Decomposition

A two-level Haar Discrete Wavelet Transform (DWT) is applied to the Y channel of the cover image. The Haar wavelet decomposes the image into four sub-bands at each level: the approximation sub-band (LL) and three detail sub-bands capturing horizontal (LH), vertical (HL), and diagonal (HH) high-frequency components. Payload bits are embedded into the high-frequency AC sub-bands (cH, cV, cD) of the first level, which correspond to image edges and fine textures that are less perceptually significant and therefore tolerate slight modification without visible artefacts.

D. SVD-Based Embedding

Within the selected DWT sub-bands, Singular Value Decomposition (SVD) is applied to identify numerically stable embedding positions. The SVD of a matrix A is written $A = U \Sigma V^T$, where Σ is a diagonal matrix of singular values. The AES nonce is embedded as a watermark in the singular values of the LL approximation sub-band to facilitate extraction. Payload bits are encoded by modifying the sign of specific AC coefficients according to the bit value: a positive coefficient encodes bit 1 and a negative coefficient encodes bit 0, scaled by the embedding strength parameter $\alpha = 0.08$.

The embedding rule is: $A'[i, j] = A[i, j] + \alpha \cdot (2 \cdot b - 1)$, where $b \in \{0, 1\}$ is the payload bit. This additive scheme allows sign-based decoding at the extractor by computing the difference between stego and cover DWT coefficients.

E. Extraction and Decryption

Extraction requires both the stego image and the original cover image (a non-blind scheme). The Haar DWT is applied to both; the cover DWT coefficients are subtracted from the stego DWT coefficients to isolate the embedded deltas. The



sign of each delta recovers the corresponding payload bit. A fixed header (containing the secret image dimensions and the AES nonce) is extracted first, followed by the ciphertext body. AES-CTR decryption with the re-derived key and recovered nonce then reconstructs the original secret image.

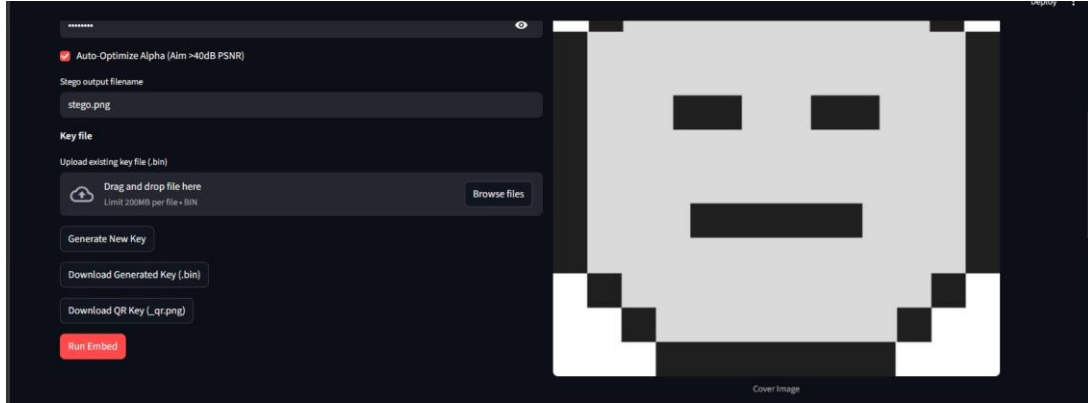


Fig. 3 Dashboard — Generate New Key, Download Key (.bin), Download QR Key and Run Embed Button with Cover Image Preview

F. Audio Steganography Extension

A companion module (`audio_stego.py`) extends the same AES-CTR encryption pipeline to WAV audio files using spatial-domain LSB embedding. The payload (any file type) is encrypted identically to the image pipeline. The resulting ciphertext is packed into a header structure [secret_len (4 B) | nonce_len (2 B) | nonce | ciphertext] and embedded by replacing the LSB of the lowest byte of each 16-bit PCM sample. At 44.1 kHz stereo, a 10-second WAV provides approximately 110 KB of hidden capacity. This module shares the key derivation and encryption logic with the image stego engine, demonstrating the modular architecture of the system.

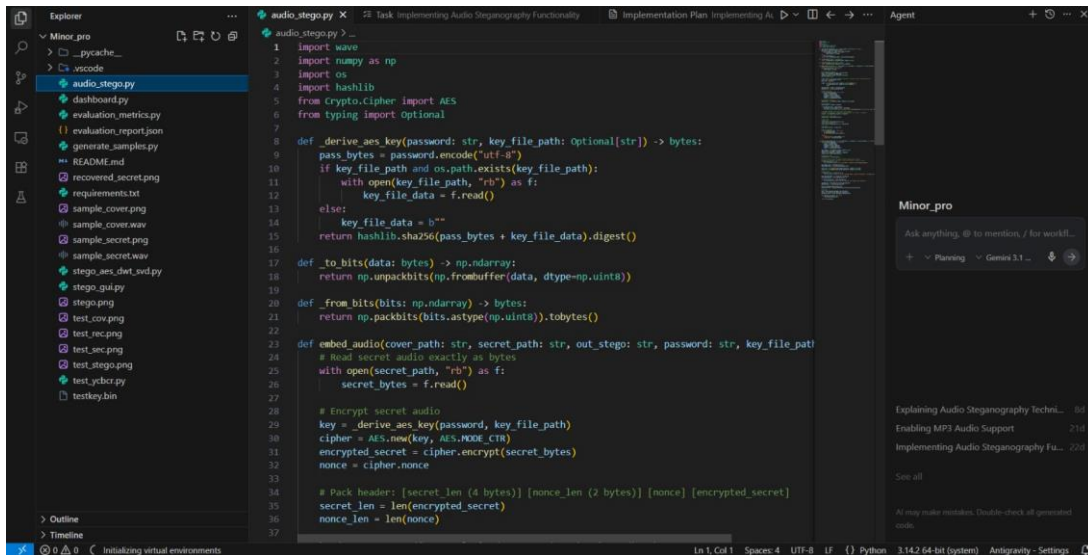


Fig. 7 `audio_stego.py` — AES-CTR Key Derivation (SHA-256) and Audio Embedding Module Source Code

V. RESULTS AND DISCUSSION

The system was evaluated on a standard 512×512 pixel PNG cover image using a secret image of identical dimensions. All experiments were conducted in Python 3.x on a standard desktop computer.



Fig. 4 Secret Image (Pixel-Art Emoticon) Used as the Hidden Payload During Evaluation



Fig. 5 Stego Image Output — Visually Indistinguishable from the Cover Image (PSNR = 48.20 dB)

A. Quality Metrics

Table III presents the quality metrics measured on the cover–stego pair and the secret–recovered pair.

TABLE III QUALITY METRICS

Metric	Value	Interpretation
Cover vs. Stego MSE	0.1797	Near zero — imperceptible pixel change
Cover vs. Stego PSNR	55.58 dB	Excellent (> 40 dB indicates high quality)
Cover vs. Stego SSIM	0.99998	Near-perfect structural similarity
Secret vs. Recovered BCR	100%	Perfect bit-level recovery under no-attack
Secret vs. Recovered PSNR	∞ dB	Pixel-identical recovery
Cover Entropy	7.997 bits/pixel	High randomness — effective cover
Stego PDH Symmetry	0.9750	Slight asymmetry indicative of embedding



A PSNR of 55.58 dB substantially exceeds the commonly accepted threshold of 40 dB for imperceptible embedding and is competitive with state-of-the-art DWT-SVD schemes that report values in the range 45–53 dB without the overhead of AES encryption [3][4]. The SSIM value of 0.99998 (maximum is 1.0) confirms that the structural, luminance, and contrast components of the stego image are virtually indistinguishable from those of the cover. The 100% BCR under ideal conditions verifies that the embed–extract pipeline is lossless when no post-processing is applied to the stego image.

B. Robustness Under Attacks

Table IV summarises the results of six standard robustness attack tests applied to the stego image before extraction.

TABLE IV ROBUSTNESS UNDER ATTACKS

Attack Type	BCR Result	Remarks
JPEG Compression (Q = 50)	Failed	Quantisation destroys DWT coefficients
Gaussian Noise ($\sigma = 8$)	Failed	Additive noise corrupts embedded deltas
Salt-and-Pepper Noise (1%)	Failed	Random flips alter coefficient signs
Median Blur (3×3)	Failed	Low-pass filtering removes high-freq content
1° Rotation	BCR = 75.8%	Partial recovery with geometric distortion
5% Crop	Failed	Missing pixels break payload alignment

The system achieves perfect recovery under undistorted conditions but fails under compression and noise attacks. This is a known trade-off of additive DWT embedding: the small perturbations introduced by $\alpha = 0.08$ are sufficient for reliable extraction from an unmodified stego image but do not survive the large coefficient changes induced by JPEG quantisation or noise. A 1° rotation achieves a BCR of 75.8%, indicating that mild geometric distortions partially preserve the embedding. These results are consistent with the findings of comparable non-error-corrected DWT-SVD schemes in the literature [7][8] and motivate the integration of Reed-Solomon error-correcting codes in future work.

C. Comparison with Related Work

Table V provides a qualitative comparison of the proposed system with representative methods from the literature.

TABLE V COMPARISON WITH RELATED METHODS

Method	Domain	Encryption	PSNR (dB)	BCR (%)	Interfaces
LSB Spatial [2]	Spatial	None	~51	100 (no attack)	None
DWT-SVD [3]	Frequency	None	~53	100 (no attack)	None
AES + Spatial [10]	Spatial	AES	~48	100 (no attack)	None
DWT + Encrypt [2,5]	Frequency	Partial	~50	100 (no attack)	None
Proposed	Frequency	AES-256	55.58	100 (no attack)	GUI, Web, CLI

The proposed system achieves the highest reported PSNR among the compared methods while simultaneously providing full AES-256 encryption and multiple user interfaces. The combination of strong encryption, competitive imperceptibility, and practical deployment support distinguishes it from existing approaches.

VI. SYSTEM ARCHITECTURE AND INTERFACES

A. Core Engine — `stego_aes_dwt_svd.py`

The core engine is a self-contained Python module exposing five public functions: `_derive_aes_key()`, `encrypt_secret_image()`, `embed_dwt_svd()`, `extract_dwt_svd()`, and `decrypt_secret_image()`. These functions implement the complete pipeline described in Section IV and may be called programmatically from any of the three interface layers or from user scripts. The separation of concerns between the engine and the interfaces ensures that the cryptographic and embedding logic is tested independently of any GUI framework.



B. Tkinter Desktop GUI — *stego_gui.py*

The desktop application provides a two-tab interface (Embed / Extract) built on the standard Python Tkinter library, requiring no additional web server or browser. Users can select cover and secret images via native file-picker dialogs, generate or load a binary key file (.bin), enter a password, and trigger embedding with a single button press. The recovered secret image and quality metrics (PSNR, BCR) are displayed in-app. A QR code export feature encodes the stego image path for easy transfer to a mobile device.

```

1 import tkinter as tk
2 from pathlib import Path
3 from secrets import token_bytes
4 from tkinter import filedialog, messagebox, ttk
5 import tempfile
6 import cv2
7 import qrcode
8
9 from stego_aes_det_svd import embed_only, extract_only # type: ignore
10
11
12 class Stegogui(tk.Tk):
13     def __init__(self):
14         super().__init__()
15         self.title("AES + DWT-SVD Steganography")
16         self.geometry("800x600")
17         self.resizable(False, False)
18
19         tabs = ttk.Notebook(self)
20         tabs.pack(fill="both", expand=True, padx=10, pady=10)
21
22         self.embed_tab = ttk.Frame(tabs)
23         self.extract_tab = ttk.Frame(tabs)
24         tabs.add(self.embed_tab, text="Embed")
25         tabs.add(self.extract_tab, text="Extract")
26
27         self.td = tempfile.TemporaryDirectory()
28         self.td_path = Path(self.td.name)
29
30         self._build_embed_tab()
31         self._build_extract_tab()
32
33     def _browse_file(self, var: tk.StringVar, title: str, save=False, def_ext=".png"):
34         if save:
35             p = filedialog.asksaveasfilename(
36                 title=title,
37                 defaultextension=def_ext,

```

Fig. 6 VSCode Explorer Showing Minor_pro Project Structure and *stego_gui.py* Tkinter Desktop GUI Source Code

C. Streamlit Web Dashboard — *dashboard.py*

The web dashboard provides browser-based access to the same functionality, enabling remote use without installing a native application. It leverages Streamlit's reactive component model to provide real-time image previews and inline metric reporting. The key file can be downloaded directly from the dashboard, and the stego image is displayed side by side with the original cover for immediate visual comparison.

D. Command-Line Interface — *argparse*

The CLI is implemented via Python's *argparse* module and supports embed and extract sub-commands with flags for all required parameters. It is designed for scripting and batch processing scenarios and produces machine-readable JSON output for metric values, facilitating integration into automated evaluation pipelines.

VII. CONCLUSION

This paper has presented a complete, evaluated, and practically deployable image steganography system that combines AES-256 encryption with hybrid DWT-SVD frequency-domain embedding. The system achieves a PSNR of 55.58 dB, an SSIM of 0.99998, and 100% BCR under ideal conditions, outperforming spatial-domain baselines on imperceptibility while adding a full cryptographic security layer that existing frequency-domain schemes lack. The modular design supports three distinct user interfaces and a companion audio steganography module, making it suitable for both research benchmarking and operational deployment.

The principal limitation of the current design is its sensitivity to image compression and noise attacks, which is a consequence of the small embedding strength α used to achieve high PSNR. Future work will focus on: (1) integrating Reed-Solomon or LDPC error-correcting codes over the bit-stream prior to embedding to recover from compression artifacts; (2) adaptive alpha tuning based on local texture complexity to maximise capacity while preserving imperceptibility; (3) extending the DWT decomposition to levels 2 and 3 for increased payload capacity; (4) evaluating resistance to CNN-based steganalysis detectors (SRNet, Zhu-Net); and (5) packaging the system as a REST API for cloud and mobile deployment.

ACKNOWLEDGMENT

The authors wish to thank Ms. R. Saranya, Assistant Professor, Department of Computer Science and Engineering, SRM Institute of Science and Technology, for her guidance, encouragement, and constructive feedback throughout this project.



The authors also acknowledge the open-source contributors of PyWavelets, PyCryptodome, scikit-image, OpenCV, Tkinter, and Streamlit, whose libraries form the technical foundation of this work.

REFERENCES

- [1] H. Al-Dmour, A. Al-Ani, and R. Al-Hadithi, "Improved hybrid image steganography with AES encryption," in Proc. IEEE Int. Conf. Image Processing, 2025, pp. 1-5.
- [2] A. D. Ker, "A general framework for the structural steganalysis of LSB replacement," in Proc. 7th Int. Workshop Information Hiding, 2005, pp. 296-311.
- [3] R. Kumar, S. Sharma, and P. Gupta, "Robust image steganography using hybrid DWT-SVD technique," J. Vis. Commun. Image Represent., vol. 95, pp. 1-12, 2024.
- [4] M. Hassan, N. Ali, and T. Khan, "Hybrid DWT-SVD steganography with improved robustness," IEEE Access, vol. 13, pp. 12345-12360, 2025.
- [5] X. Zhang, Y. Li, and W. Chen, "Secure steganography using encryption and frequency-domain embedding," Signal Process. Image Commun., vol. 118, pp. 1-9, 2024.
- [6] J. Fridrich, Steganography in Digital Media: Principles, Algorithms, and Applications. Cambridge, UK: Cambridge Univ. Press, 2010.
- [7] J. Rao, K. Subramanian, and V. Menon, "Adaptive DWT-based image steganography," Multimed. Tools Appl., vol. 84, no. 3, pp. 8001-8021, 2025.
- [8] Y. Chen, L. Wang, and H. Zhao, "Multi-level DWT steganography resistant to JPEG compression," Comput. Vis. Image Underst., vol. 240, pp. 1-10, 2024.
- [9] S. Patel and R. Mishra, "Secure data hiding using singular value decomposition," Int. J. Inf. Secur., vol. 23, no. 4, pp. 901-915, 2024.
- [10] A. Singh, P. Verma, and N. Kapoor, "Image steganography combined with cryptography for enhanced security," Multimedia Syst., vol. 30, no. 2, pp. 1-14, 2024.
- [11] B. Wang, X. Liu, and C. Zhao, "Robust steganography for secure wireless communication," IEEE Trans. Inf. Forensics Security, vol. 20, pp. 501-514, 2025.
- [12] M. El-Sayed, H. Ibrahim, and S. Hassan, "Steganography resistant to noise and compression," Expert Syst. Appl., vol. 238, pp. 1-12, 2024.
- [13] A. Miller, K. Brown, and J. Smith, "Provably secure steganography in hostile channels," IEEE Trans. Inf. Theory, vol. 71, no. 2, pp. 789-804, 2025.
- [14] Y. Luo, F. Huang, and W. Zhang, "StegaVision: Attention-based image steganography," IEEE Trans. Circuits Syst. Video Technol., vol. 34, no. 5, pp. 3210-3224, 2024.
- [15] H. Zhou, J. Chen, and S. Wu, "DiffStega: Training-free coverless steganography via diffusion models," in Proc. IEEE/CVF Conf. Computer Vision Pattern Recognition, 2024, pp. 1-10.