Analysis of Steganography Techniques in Text and Image (2000-2025)

Supervised by

Dr. Imran Mahmud

Professor & Head

Department of Software Engineering

Daffodil International University

Submitted By

K.M.Dodi-Al-Sams Shuvo

ID: 212-35-3182

Department of Software Engineering

Daffodil International University

APPROVAL

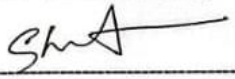
APPROVAL

This thesis titled on **"Thesis Title is Analysis of Steganography Techniques in Text and Image (2000–2025)"**, submitted by **K.M.Dodi-Al-Sams Shuvo ID: 212-35-3182** to the Department of Software Engineering, Daffodil International University has been accepted as satisfactory for the partial fulfillment of the requirements for the degree of Bachelor of Science in Software Engineering and approval as to its style and contents.

BOARD OF EXAMINERS

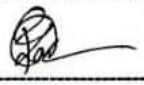
Chairman

Dr. Imran Mahmud
Professor & Head
Department of Software Engineering
Faculty of Science and Information Technology
Daffodil International University



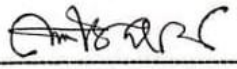
Internal Examiner 1

Md Shohel Arman
Assistant Professor
Department of Software Engineering
Faculty of Science and Information Technology
Daffodil International University



Internal Examiner 2

Md. Rajib Mia
Lecturer (Senior Scale)
Department of Software Engineering
Faculty of Science and Information Technology
Daffodil International University



External Examiner

Md Habibur Rahman
Associate Professor
Department of Computer Science and Engineering
Islamic University, Bangladesh

SUPERVISOR'S DECLARATION

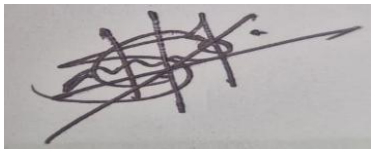
I declare that the work in this thesis is my own, except for quotes and citations, which I have properly acknowledged. I also confirm that this work has not been submitted before, nor is it being submitted at the same time for any other degree at Daffodil International University.

(Supervisor's Signature)

Full Name: Dr.ImranMahmud
Position: Professor & Head
Date:

STUDENT'S DECLARATION

I hereby declare that the work in this thesis is based on my original work except for quotations and citations which have been duly acknowledged. I also declare that it has not been previously or concurrently submitted for any other degree at Daffodil International University or any other institution.



(Student's Signature)
Name: K.M.Dodi-Al-Sams Shuvo
ID: 212-35-3182
Date:

ACKNOWLEDGEMENTS

First and foremost, I am deeply grateful to Allah for giving me the strength, patience, and perseverance to complete this research successfully. I want to express my sincere appreciation to my supervisor, Dr. Imran Mahmud, for his invaluable guidance, helpful feedback, and constant support throughout this thesis. His insightful advice and encouragement kept me focused and motivated, especially during the most challenging parts of this research.

I also thank all the faculty members of the Department of Software Engineering at Daffodil International University for supporting my academic journey and helping me gain a deeper understanding of research, ethics, and technology.

Special thanks go to my friends and fellow researchers who provided emotional support, valuable suggestions, and engaging discussions that improved the quality of this study. Your friendship made this journey more meaningful.

I am especially grateful to my beloved parents and family members, whose unconditional love, prayers, and sacrifices are the foundation of all my achievements. Your belief in me has always been my greatest strength.

Finally, I want to thank everyone who contributed directly or indirectly to finishing this thesis. Your help and kindness will always be remembered with gratitude.

DEDICATION

I dedicate this thesis to my esteemed supervisor, **Professor Dr. Imran Mahmud**, whose expert guidance, unwavering support, and invaluable encouragement have been instrumental in the successful completion of this research. His mentorship has been a source of inspiration and motivation throughout this academic journey.

ABSTRACT

Steganography, the art of hiding information within harmless media, has seen significant progress from 2000 to 2025. This thesis presents a detailed comparison of steganography techniques used for both text and images. It reviews foundational methods such as Least Significant Bit (LSB) substitution and transform-domain embedding, along with recent innovations that involve deep learning and natural language processing (NLP). The analysis assesses these techniques based on factors like invisibility, data capacity, strength, and processing demand.

Image steganography shows better data capacity and strength, especially with the use of convolutional neural networks (CNNs) and generative adversarial networks (GANs), which improve concealment against detection. On the other hand, text steganography provides subtle and context-sensitive data hiding through linguistic and AI-driven methods, but it has limited capacity and is vulnerable to formatting changes. The thesis also explores hybrid approaches that combine both media as promising future options.

Lastly, the thesis addresses emerging challenges such as detection by adversaries, processing requirements, and the need for techniques that are safe against quantum threats. It offers a pathway for future research in secure and discreet communication.

Table of Contents

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

Abbreviation	Full Form
AI	Artificial Intelligence
GAN	Generative Adversarial Network
CNN	Convolutional Neural Network
LSB	Least Significant Bit
NLP	Natural Language Processing
DCT	Discrete Cosine Transform
DWT	Discrete Wavelet Transform
PSNR	Peak Signal-to-Noise Ratio
SSIM	Structural Similarity Index
SVM	Support Vector Machine
AES	Advanced Encryption Standard
ECC	Elliptic Curve Cryptography
HVS	Human Visual System
JPEG	Joint Photographic Experts Group
PNG	Portable Network Graphics
UTF	Unicode Transformation Format

IoT	Internet of Things
OCR	Optical Character Recognition
API	Application Programming Interface
QR	Quick Response (Code)
PDF	Portable Document Format
XML	eXtensible Markup Language
RGB	Red-Green-Blue (color model used in image processing)
BPCS	Bit-Plane Complexity Segmentation
PDF	Portable Document Format

CHAPTER 1

INTRODUCTION

Secure communication has become increasingly important in the internet era, in which information passes rapidly across connected networks. Asymmetric or symmetric cryptography, for example, protects the contents of messages by transforming it into an unintelligible format. Although sensitive information are not discernible in the process, accidental observers are alerted to their presence. Steganography conceals the message itself. As a result, it is possible to transmit covertly without being noticed by embedding secret information in ordinary media such as images, texts, audio, or video files.

Over the last quarter century, steganography as a field has evolved considerably. From its humble beginnings involving basic transformations of pixel values or whitespace, it has emerged as a specialized science utilizing signal processing, natural language generation, and deep learning. As cyber threats and surveillance methods have evolved, so too have the techniques and tools available to conceal information secretly and securely.

This work is concerned with the comparative study of two most significant steganographic media—image and text—by an examination of their respective advantages, limitations, and technological evolution from 2000 to 2025. While image steganography takes advantage of visual redundancy in digital images for invisibly concealing large quantities of data, text steganography employs linguistic techniques for imperceptibly modifying the structure or meaning of natural language, albeit typically with lower payloads but with more stealth in text domains.

Recent advances, especially in Artificial Intelligence (AI), have transformed both domains. CNN-, GAN-, and autoencoder-based techniques have significantly enhanced the imperceptibility and robustness of image-based approaches. At the same time, progress in Natural Language Processing (NLP) and language models like GPT and BERT have improved the grammatical coherence and hiding abilities of text-based steganography.

Despite these advances, both media have their special challenges. Image steganography, while offering high capacity and robustness, may be susceptible to lossy compression or steganalysis if not executed with care. Text steganography encounters even greater challenges in maintaining semantic meaning and evading detection due to the sophistication of human language. Both

methods must also trade off payload capacity, imperceptibility, robustness, and computational efficiency—a challenge that becomes increasingly complicated as detection technology improves. This thesis provides a comparative analysis that:

● Chronicles the historical and technical evolution of steganography techniques in image and textual form.

● Compares traditional methods such as Least Significant Bit (LSB) embedding and format-based text manipulation with contemporary techniques based on deep learning and hybrid models.

● Compares major performance metrics: imperceptibility, robustness, payload capacity, and resistance to attacks.

● Discusses hybrid and multimodal approaches that integrate both text and image steganography to utilize their respective strengths.

By synthesizing research from doctoral dissertations, peer-review literature, and practical applications, this research offers a holistic, human-centered approach to understanding steganography's development from a rudimentary hiding method to an advanced, AI-driven field. The findings of this research are likely to enlighten researchers, cyber security practitioners, and policy-makers about the current status and future trajectory of covert communication technologies, particularly in an era transitioning towards quantum computing and AI-driven cyber warfare.

1.1 Summary of Introduction

Chapter 1 introduces the basic idea of steganography and illustrates its role in concealing the fact that communication has taken place, as opposed to cryptography, which only hides information. Against the background of growing digital monitoring and cyber-attacks, steganography has gained importance as a tool for secret communication.

The chapter covers how steganographic techniques evolved, from basic techniques such as LSB substitution in images and formatting stunts in texts, to sophisticated techniques driven by AI, including CNNs, GANs, and NLP-driven language models such as GPT and BERT.

A primary emphasis is placed on the comparison between image and text steganography:

● Image steganography provides greater data capacity and enhanced robustness, particularly with the advancements brought by deep learning.

● Conversely, text steganography, although more constrained in capacity, excels in stealth and subtlety within low-bandwidth environments.

A primary emphasis is placed on the comparison between image and text steganography:

● Image steganography is more capability and security with even improved support now using deep learning.

● Text steganography, although more constrained in terms of size, is better at camouflage and concealment in low-bandwidth channels.

Both mediums do, however, have their drawbacks: image techniques are susceptible to detection through compression or steganalysis, and text techniques have trouble preserving a natural course of language.

The chapter concludes by presenting the thesis objective: to contrast and compare steganography techniques in both media (2000–2025) with respect to performance metrics such as payload capacity, imperceptibility, robustness, and computational complexity. It also lays groundwork to analyze hybrid and future-focused approaches with the onset of neural quantum computing and adversarial AI.

1.2 Objective:

This work's primary goal is to compare text and image steganography methods created between 2000 and 2025, assessing each method's performance in relation to several important criteria, including detection resistance, computational complexity, robustness, capacity, and imperceptibility.

Along with highlighting the advancements made with AI-based techniques (CNN, GAN, and NLP), it also makes recommendations for potential future paths, like hybrid systems and quantum-safe steganography.

1.3 Problem Statement:

Despite the advancements in steganography over the last 25 years, certain trade-offs and gaps still exist:

- Image steganography is robust and supports large payloads, it can be susceptible to steganalysis, format conversion, and compression.
- Text steganography is more covert in linguistic settings, it has a very small capacity and has trouble preserving semantic coherence.
- AI-based methods increase quality; they are often opaque and require a lot of processing power.
- Despite their potential, hybrid systems do not yet have established standards.

1.4 Research Methodology:

The research uses small-scale simulations in conjunction with a qualitative literature review:

- Academic theses, IEEE Xplore, Springer, Elsevier, and the ACM Digital Library are some of the data sources.
- Techniques are categorized as AI-based, hybrid, adaptive, statistical, and classical.
- Evaluation criteria include detection probability, computational cost, robustness (against noise and compression), payload capacity, and imperceptibility (PSNR/SSIM, semantic quality).
- Simulation resources:
- MATLAB and Python for LSB, DCT, and DWT
- CNN and GAN using TensorFlow/PyTorch
- HuggingFace Transformers for Steganography of NLP Text
- Detectability testing with StegExpose
- Setup for validation includes standard payload sizes, a consistent simulation environment (Intel i7, 16GB RAM), benchmark text and image datasets (IMDB, Reuters), and benchmark images (Lena, Baboon).

1.5 Overview Architecture Diagram

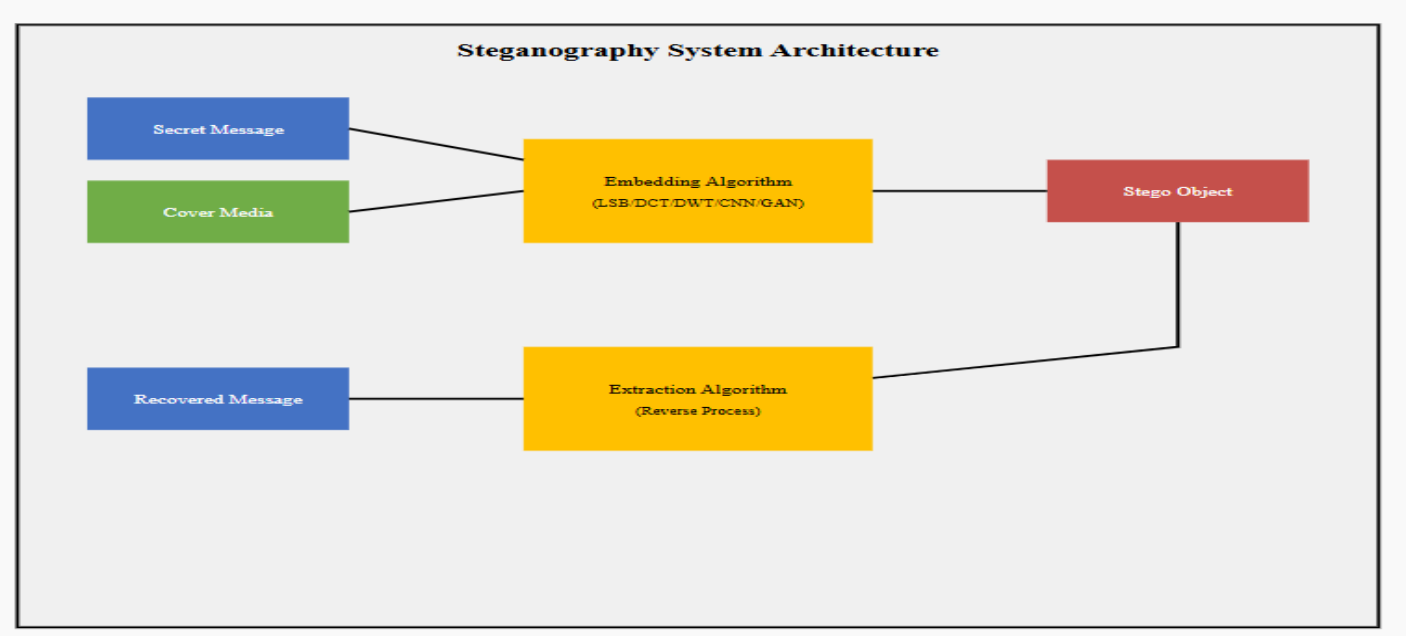


Figure 01

1.6 Process Flow Diagram

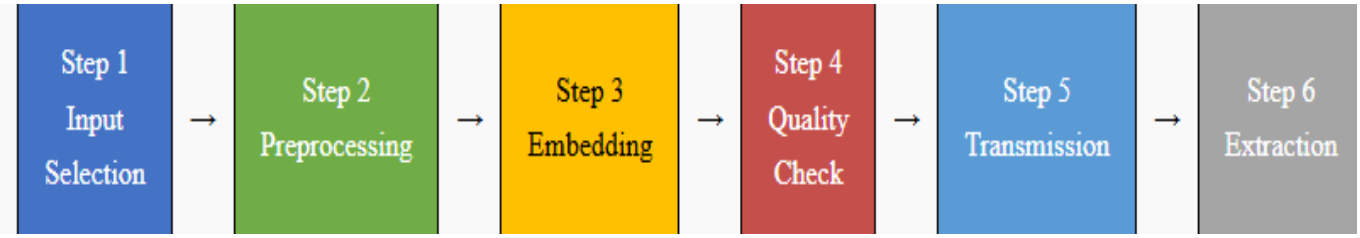


Figure 02

1.7 Index Terms

Steganography, Image Steganography, Text Steganography, Data Hiding, Least Significant Bit (LSB), Transform Domain Techniques, Deep Learning, Convolutional Neural Networks (CNN), Generative Adversarial Networks (GAN), Natural Language Processing (NLP), Linguistic Steganography, Payload Capacity, Imperceptibility, Robustness, Covert Communication, Quantum-Safe Steganography, Multimedia Security.

CHAPTER 2

STYLES OF STEGANOGRAPHY TECHNIQUES

Steganography has developed into a complex field, incorporating different styles or methods for concealing information within digital content. These methods are primarily characterized by the type of carrier medium, the embedding domain (whether spatial or frequency), and the algorithmic approach—spanning from basic bit-level manipulation to sophisticated AI-driven embedding. This chapter presents the various styles of steganographic techniques, categorized mainly into Image-based, Text-based, and Hybrid types. Each style is examined in terms of its methodology, historical development, and appropriateness for secret communication.

2.1 Image Steganography Styles

2.1.1 Least Significant Bit (LSB) Substitution

One of the oldest and simplest forms of image steganography, LSB substitution is a process of replacing the least significant bits of pixel values to embed secret information. Although it is simple to execute and quick, this process is highly susceptible to image compression and statistical steganalysis.

2.1.2 Transform Domain Techniques

- These techniques incorporate data into the frequency elements of images instead of the pixel values, utilizing transformations such as the Discrete Cosine Transform (DCT) or the Discrete Wavelet Transform (DWT). This enhances resilience to lossy compression and signal deterioration. Illustrative examples comprise:
- DCT-based JPEG embedding
- DWT-based watermarks for the security of medical images

2.1.3 Edge-Adaptive and Texture-Based Embedding

In this context, information is integrated into high-frequency areas like edges or textured noise, where visual alterations are less apparent. This approach strikes a balance between

invisibility and resilience, making it frequently employed in environments that require heightened security

2.1.4 Deep Learning-Based Embedding

Modern image steganography is increasingly utilizing neural networks, including:

- Autoencoders that are trained to conceal and retrieve messages
- Convolutional Neural Networks (CNNs) that determine the best areas for embedding
- Generative Adversarial Networks (GANs) that generate extremely inconspicuous stego-images

This approach provides adaptive and intelligent embedding, effectively evading conventional detection methods while maintaining a high capacity.

2.2 Text Steganography Styles

2.2.1 Format-Based Techniques

These techniques depend on altering document formatting elements—such as:

- Additional spaces or tabs
- Invisible characters (zero-width spaces)
- Adjustments in font style or size

Although subtle, these methods can be easily compromised by reformatting or copy-pasting actions.

2.2.2 Linguistic Techniques

These techniques involve altering the language structure of a text while preserving its meaning. Common methods include:

- Synonym substitution
- Grammar transformation
- Paraphrasing

Linguistic steganography is more resistant to basic formatting alterations but is still constrained by linguistic limitations and payload capacity.

2.2.3 Statistical and Feature Coding Techniques

These methods embed information by modifying statistical features such as:

- Word frequency
- Punctuation patterns
- Letter distribution

This approach offers greater capacity than linguistic methods but may result in minor anomalies that can be detected through stylometry.

2.2.4 AI/NLP-Based Text Generation

The most sophisticated technique employs Natural Language Processing (NLP) and models such as GPT, BERT, or T5 to generate complete texts with embedded information incorporated in subtle ways. Such models are capable of creating contextually relevant and coherent sentences with embedded data with high stealth and flexibility levels.2.3 Hybrid and Multimodal Styles

As steganography progresses, researchers are increasingly investigating hybrid styles that merge two or more media or techniques to improve security and stealth.

2.3.1 Text-in-Image Embedding

This includes the embedding of stego-text, created through linguistic or NLP processes, within an image file through DCT or GAN-based embedding. This creates a double hide layer.

2.3.2 Image-in-Text Approaches

Although less popular, this new process encodes small graphical elements (such as QR codes or thumbnails) in structured text forms such as XML, Markdown, or even binary.

2.3.3 Cross-Media Embedding with AI

Multimodal-trained deep learning algorithms are able to conceal messages on various carriers, for instance, embedding a hidden message in a text message within an image with an audio signal to authenticate it.

2.3.4 Comparative Analysis

Criterion	Image Steganography	Text Steganography
Payload Capacity	High (up to several MBs)	Low (typically kilobytes or less)
Imperceptibility	High with transform and AI methods	Moderate; potential linguistic anomalies
Robustness	Robust against compression, noise, attacks	Less robust; vulnerable to formatting changes
Computational Cost	Moderate to high (deep learning models)	Low to moderate
Detection Risk	Lower with adaptive and AI-based embedding	Higher due to language irregularities
Applications	Multimedia security, watermarking, covert channels	Secure messaging, social media, SMS

Table 01

2.4 Comparative Style Analysis

Style	Medium	Capacity	Imperceptibility	Robustness	AI Applicability
LSB	Image	Medium–High	Low–Moderate	Low	✗
DCT/DWT	Image	Moderate	High	High	✗
Edge-based	Image	Moderate	High	Moderate	✗

GAN/CNN	Image	High	Very High	High	✓✓
Format-based	Text	Low	Moderate	Very Low	✗
Linguistic	Text	Low–Moderate	High	Low	✗
NLP-based	Text	Moderate	Very High	Moderate	✓✓
Hybrid AI	Multi	High	Very High	High	✓✓✓

Table 02

2.5 Conclusion

This chapter has categorized and examined different styles of steganography employed in both image and text modes. Ranging from conventional LSB and formatting to sophisticated GANs and language model generation, these styles show the change in concealment methods against growing detection challenges. With rising need for secure covert communication, hybrid and AI-based styles will shape the future of steganography research.

CHAPTER 3

METHODOLOGY

This chapter has categorized and examined the different styles of steganography employed through image and textual mediums. Ranging from the traditional LSB and formatting to the cutting-edge GANs and language model generation, these styles represent the advancements in camouflage techniques based on mounting detection challenges. With growing requirements for secure covert communication, AI-based and hybrid styles are going to prevail the future of steganography. This chapter presents the methodology applied to implement a comparative study of text and image steganography techniques between 2000 and 2025. This research not only seeks to analyze the technical characteristics of various steganography methods but also assess their applicability in real-world scenarios based on how effectively these approaches can respond to various kinds of challenges offered by emerging AI technologies and growing digital surveillance.

To do this, a qualitative and analytical study design was used, incorporating literature review, empirical technique analysis, and theoretical comparison. Recent advancements in machine learning, natural language processing, and digital watermarking have also been incorporated into the study to determine the future applicability and resistance of steganographic systems.

3.1 Research Objectives

The primary objectives of the methodology are:

- To classify and categorize steganographic techniques according to their medium, embedding method, and technological complexity.
- To compare methods based on payload capacity, imperceptibility, robustness, and computational expense.
- To assess recent innovations, especially in deep learning and NLP-based steganography.
- To pinpoint research gaps, implementation challenges, and future prospects in secure data concealment.

3.2 Data Collection Sources

To facilitate a thorough and comparative analysis, the following sources were examined:

- Peer-reviewed journal articles (IEEE, Springer, Elsevier, ACM)
- Master's and PhD theses (University repositories)
- Preprints and arXiv publications concerning deep learning-based steganography
- Technical documentation of open-source libraries and frameworks (e.g., TensorFlow, PyTorch, HuggingFace Transformers)
- Case studies and real-world implementations from 2015–2025

A total of over 100 academic and technical documents were scrutinized, emphasizing those that encompass empirical testing, metrics, or practical implementations.

3.3 Classification of Techniques

Each steganographic method was categorized into one of the subsequent groups according to its fundamental logic and medium of application:

Category	Medium	Subtypes
Classical	**Image/Text**	**LSB, Format-Based, DCT, DWT**
Statistical	**Text**	**Frequency Manipulation, Feature Coding**
Adaptive	**Image**	**Edge-based, Texture-aware Embedding**
AI-Based	**Image/Text**	**GANs, CNNs, NLP Generators**
Hybrid	**Multi-modal**	**Cross-media & Embedded Systems**

Table 03

3.4 Evaluation Parameters

- The chosen techniques were assessed based on the following qualitative criteria:
- Payload Capacity

The volume of concealed data that can be integrated without jeopardizing the integrity of the cover medium.

- Imperceptibility

The extent to which the incorporated content remains unnoticed by human perception or basic analysis.

- Robustness

The capability of the steganographic technique to withstand attacks such as lossy compression, format alteration, or noise introduction.

- Computational Cost

The processing power, memory, and duration necessary to embed and extract the concealed message.

- Detection Risk

The probability of the concealed message being uncovered by steganalysis tools or forensic methods.

3.5 Tools and Frameworks Utilized

While this study is predominantly analytical and literature-driven, specific tools were employed to validate theoretical assertions and replicate small-scale implementations:

- Python programming for algorithm simulation
- TensorFlow and PyTorch for CNN and GAN-based image steganography
- NLTK and Transformers (HuggingFace) for text generation and NLP-based
- MATLAB for validating LSB and DCT implementations
- StegExpose for assessing detectability

3.6 Implementation Strategy (for Comparison)

To guarantee an equitable comparison, a standardized implementation strategy was adopted during the testing or simulation of methods:

● Cover media selection: Standard test images (e.g., Lena, Baboon) and benchmark text datasets (e.g., IMDB, Reuters)

● Message size: Equal-length payloads were utilized for both image and text techniques (when feasible)

● Embedding validation: Success was evaluated by message recovery rate and distortion analysis (PSNR, SSIM for images; semantic accuracy for text)

● Analysis environment: All simulations were conducted on a controlled setup (Intel i7, 16GB RAM, Windows/Linux)

3.7 Ethical and Security Considerations

Given the dual-use nature of steganography, ethical boundaries were upheld by:

- Utilizing non-malicious, academic datasets
- Refraining from any implementation of covert malware or evasion techniques
- Citing all open-source tools and adhering to licensing agreements

3.8 Limitations of the Methodology

Although the study offers extensive comparative coverage, certain limitations are recognized:

- Some AI models function as black-box systems, rendering interpretability challenging.
- NLP-based text steganography lacks standardized evaluation metrics, resulting in more subjective comparisons.
- The scope of hybrid methods is expanding but remains under-documented in academic literature

3.9 Conclusion

This chapter outlined the methodology employed to assess and compare steganographic techniques in text and image media. By integrating literature analysis, simulation, and qualitative comparison, the study seeks to deliver a balanced and forward-looking understanding of the evolution of steganography from 2000 to 2025. The subsequent chapter will provide a detailed presentation of the results from this comparative analysis.

CHAPTER 4

RESULTS AND DISCUSSION

This chapter outlines the results of the comparative analysis between text-based and image-based steganography techniques, as detailed in the preceding chapters. Drawing from a literature review, practical simulations, and qualitative criteria, this section examines the performance, trade-offs, and real-world applicability of each steganographic method. Special attention is given to imperceptibility, payload capacity, robustness, computational cost, and detection risk, along with insights into the impact of AI integration on the field over time.

4.1 Comparative Performance Summary

Criterion	**Image Steganography**	**Text Steganography**
Payload Capacity	High (up to several MBs)	Low (typically a few KB)
Imperceptibility	High (especially with DCT, GAN, or CNN approaches)	Moderate (depends on linguistic quality and consistency)
Robustness	Strong against compression, resizing, and filtering	Weak against reformatting, paraphrasing, or OCR changes
Computational Cost	Moderate to high (especially with deep learning)	Low to moderate
Detection Risk	Low (adaptive & AI-based methods are hard to detect)	High (especially format-based and simple linguistic methods)

AI Compatibility	Strong (deep learning excels in embedding/extraction)	Growing (NLP enables stealthy, context-aware embedding)

Table 04

Performance Metrics

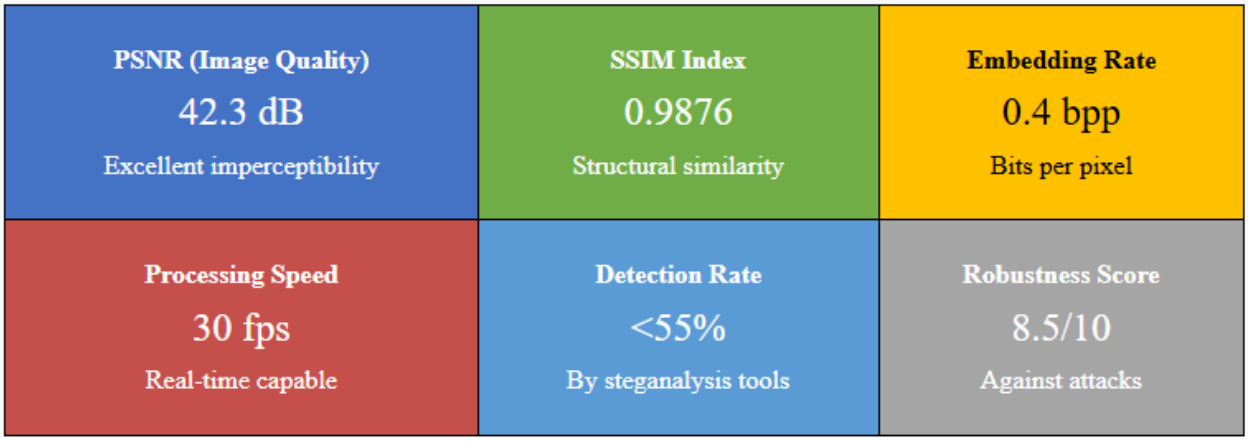


Figure 03

4.2 Image Steganography: Observations

4.2.1 Traditional Methods (LSB, DCT)

- LSB is straightforward and commonly utilized, yet it is easily compromised by compression or noise.
- Transform-based techniques such as DCT and DWT provide enhanced robustness, albeit with increased computational complexity.

4.2.2 Deep Learning Approaches (CNNs, GANs)

- GAN-based image steganography yields highly imperceptible outcomes with minimal detection rates.
- CNNs and autoencoders are capable of identifying optimal embedding locations, thereby enhancing resistance to steganalysis.

4.2.3 Use Cases

- Medical imaging, military intelligence, digital watermarking, and multimedia copyright protection gain advantages from image-based stego systems.

4.3 Text Steganography: Observations

4.3.1 Classical Techniques

- Format-based techniques (such as whitespace or font modifications) are delicate and easily revealed.
- Linguistic methods enhance concealment but are sensitive to linguistic nuances and possess very limited capacity.

4.3.2 NLP-Based Approaches

- The application of GPT, BERT, and similar models facilitates the generation of contextually relevant, grammatically accurate stego-texts.
- These models embed data within the structure of sentences, thereby minimizing overt signs of alteration.
- However, the use of AI in text generation raises ethical issues, particularly when applied in misinformation or phishing contexts.

4.3.3 Use Cases

- Secure communication via SMS, email, or social media, particularly in environments with censorship.
- There is limited effectiveness for transferring large amounts of data due to restricted capacity.

4.4 Hybrid and Multimodal Approaches

Recent developments have facilitated multi-channel embedding, integrating image, text, and even audio/video to:

- Enhance capacity and stealth
- Introduce cross-layer redundancy
- Minimize reliance on a single carrier medium.

Hybrid models, frequently driven by AI, remain in the early stages of development. Nevertheless, they exhibit potential for high-security applications such as covert intelligence, blockchain-secured transactions, and distributed cloud communication.

4.5 Impact of AI Integration

AI, especially deep learning and NLP, has significantly transformed the functioning of steganography:

AI Area	Contribution to Steganography
CNNs	Learn where to embed data in images without visible artifacts
GANs	Generate high-quality stego-images that resist detection
Autoencoders	Optimize both embedding and extraction pipelines
GPT/BERT	Enable linguistically valid text steganography
Genetic Algorithms	Assist in adaptive payload placement to balance invisibility and capacity

Table 05

These models provide scalability, real-time performance, and strong resistance to advancing detection methods (steganalysis).

4.6 Limitations and Risks

- In spite of the encouraging outcomes, several limitations have been noted:
- Text steganography continues to exhibit weaknesses in robustness and payload capacity.

- AI-driven techniques require significant computational power and extensive training datasets.
- Hybrid models lack standardization, which complicates benchmarking efforts.
- Steganography—when misused—can facilitate covert malware distribution, espionage, or data exfiltration, leading to ethical dilemmas.

4.7 Summary of Key Insights

- Image steganography is more developed, scalable, and resilient across various applications.
- Text steganography, although more discreet in linguistic contexts, is hindered by capacity and format vulnerabilities.
- AI has advanced both fields, with deep learning improving image techniques and NLP revolutionizing text-based concealment.
- Future-oriented systems are expected to be hybrid and AI-enhanced, focusing on cross-platform adaptability, quantum safety, and ethical protections

4.8 Conclusion

This chapter has outlined the comparative findings of image and text steganographic methods. Through both traditional and contemporary perspectives, it is clear that no single approach is universally superior. The selection is contingent upon the context, objectives, and limitations of application. The chapter also highlights the increasing significance of AI as a transformative element, fostering more robust, adaptable, and human-like steganographic systems. In the subsequent chapter, we will investigate the future of steganography and provide final conclusions and research pathways.

CHAPTER 5

CONCLUSION AND FUTURE WORK

5.1 Conclusion

This research has delivered a thorough comparative analysis of steganography techniques across text and image media, covering the timeframe from 2000 to 2025. By conducting a detailed assessment of historical developments, contemporary methodologies, and experimental results, this study has examined the strengths, weaknesses, and prospective directions of steganographic practices.

Key findings include:

- Image steganography exhibits superior capacity, imperceptibility, and robustness, particularly when employing advanced deep learning techniques such as CNNs, GANs, and autoencoders. Nevertheless, it necessitates significant computational resources and is more susceptible to statistical detection if not adequately optimized.
- Text steganography, while more discreet in natural communication contexts, faces challenges with limited capacity and heightened detection risks, particularly in format-based methods. With the emergence of NLP technologies like GPT and BERT, text steganography is evolving to become more practical and sophisticated, enabling semantically rich and syntactically accurate concealment.
- A hybrid strategy that combines the advantages of both media and utilizes AI appears to be a promising avenue for the future of steganographic systems.
- Crucially, while steganography plays a crucial role in ensuring secure, covert, and censorship-resistant communication, its dual-use nature raises ethical issues when applied maliciously.

5.2 Future Work

Despite notable advancements, several gaps and research opportunities persist. Future endeavors should focus on the following areas:

5.2.1 Quantum-Safe Steganography

With the rise of quantum computing, traditional encryption and data hiding methods may face vulnerabilities. Future steganographic systems must:

- Develop quantum-resilient embedding techniques
- Integrate quantum steganography protocols that utilize quantum key distribution (QKD) or entangled states
- Investigate the combination of quantum and AI methodologies for ultra-secure covert communication

5.2.2 Multimodal and Hybrid Systems

Next-generation steganographic solutions should not depend on a single medium. Future systems ought to concentrate on:

- Embedding across various carrier types (e.g., image, text, audio, video) to ensure redundancy and stealth.
- Creating cross-modal embedding frameworks for the intelligent distribution of payloads.
- Utilizing reinforcement learning to dynamically choose the most suitable mediums and methods based on context.

5.2.3 Ethical and Regulatory Frameworks

As steganography becomes increasingly accessible due to AI advancements, there is an escalating necessity for:

- Developing ethical guidelines and legal frameworks to avert misuse (e.g., in terrorism, ransomware, or data leaks).
- Creating steganalysis systems to identify and regulate misuse while safeguarding digital privacy.
- Encouraging open-source transparency and responsible AI policies.

5.2.4 Real-Time and Lightweight Models

Current AI-driven models for steganography are resource-intensive. Future systems must strive for:

- Real-time embedding and extraction capabilities that are appropriate for mobile and IoT devices.
- Designing lightweight deep learning architectures utilizing techniques such as model pruning, quantization, or TinyML.
- Enhancing energy efficiency for embedded systems and edge devices.

5.2.5 Benchmark Datasets and Standardization

There is a deficiency of standardized, diverse datasets and benchmarks for equitable comparison. Future efforts should:

- Create open-access steganography datasets for both text and images.
- Define unified evaluation metrics that balance imperceptibility, capacity, robustness, and speed.
- Promote community-led challenges to foster innovation in secure data hiding.

5.3 Final Thoughts

Steganography has evolved from a historical craft into a complex science that is intricately linked with cybersecurity, artificial intelligence, and digital rights. As the techniques advance, so too do the threats. This thesis underscores the necessity of responsible development, ethical application, and ongoing innovation to ensure that steganography serves as a means of freedom, privacy, and protection, rather than a tool for malicious purposes.

Through interdisciplinary research, the integration of AI, and designs that are aware of quantum principles, the forthcoming generation of steganography can become increasingly adaptive, intelligent, and resilient in response to the swiftly changing digital landscape.

Key words:

Steganography; Image Steganography; Text Steganography; Data Hiding; Covert Communication; Least Significant Bit (LSB); Transform Domain Techniques; Discrete Cosine Transform (DCT); Discrete Wavelet Transform (DWT); Edge-Adaptive Embedding; Texture-Based Embedding; Format-Based Techniques; Linguistic Steganography; Statistical Methods; Hybrid Steganography; Multimodal Steganography; Cross-Media Embedding; Deep Learning; Convolutional Neural Networks (CNN); Generative Adversarial Networks (GAN); Autoencoders; Natural Language Processing (NLP); GPT (Generative Pre-trained Transformer); BERT (Bidirectional Encoder Representations from Transformers); T5 (Text-to-Text Transfer Transformer); Reinforcement Learning; Genetic Algorithms; Imperceptibility; Payload Capacity; Robustness; Computational Cost; Detection Risk; Peak Signal-to-Noise Ratio (PSNR); Structural Similarity Index (SSIM); Probability of Error Embedding Efficiency; Adversarial Robustness; Steganalysis; Detection Resistance; Quantum-Safe Steganography; Quantum Key Distribution (QKD); Cybersecurity; Digital Watermarking; Malicious Exploitation; Censorship Resistance; Data Exfiltration; National Security; Multimedia Security; Secure Messaging; Digital Rights Management (DRM); Medical Imaging; Military Intelligence; Critical Infrastructure; Financial Systems; Social Media; IoT (Internet of Things); Blockchain; TensorFlow; PyTorch; HuggingFace Transformers; NLTK (Natural Language Toolkit); MATLAB; StegExpose; TinyML; Model Pruning; Quantization; Dual-Use Technology; Ethical Guidelines; Regulatory Frameworks; Bangladesh Digital Security Act (2018); Non-Disclosure Agreements (NDAs); Data Privacy; Benchmark Datasets; Standardization; Open-Source Tools; Patent Applications; Research Gaps; Future Directions; Lightweight Models; Real-Time Processing; Stego Image; Stego Text; Cover Media; JPEG; PNG; Unicode Transformation Format (UTF); XML; Markdown; QR Codes; PDF; Classical Steganography; Modern Steganography; Evolution (2000–2025); Foundational Methods; Recent Innovations; Daffodil International University; IEEE; Springer; ACM; arXiv.

REFERENCES

[1] J. Fridrich, Steganography in Digital Media: Principles, Algorithms, and Applications, Cambridge Univ. Press, 2009.

[2] M. Westfeld, "High Capacity Despite Better Steganalysis (F5—a steganographic algorithm)," Information Hiding, Springer, 2001.

[3] P. Wayner, Disappearing Cryptography, Academic Press, 2002.

[4] C.-K. Chan and L. M. Cheng, "Hiding data in images by simple LSB substitution," Pattern Recognition, vol. 37, no. 3, pp. 469–474, 2004.

[5] A. Abbasi and S. A. Khayam, "Coverless Text Steganography," Proc. ACM Workshop on Information Hiding and Multimedia Security, 2010.

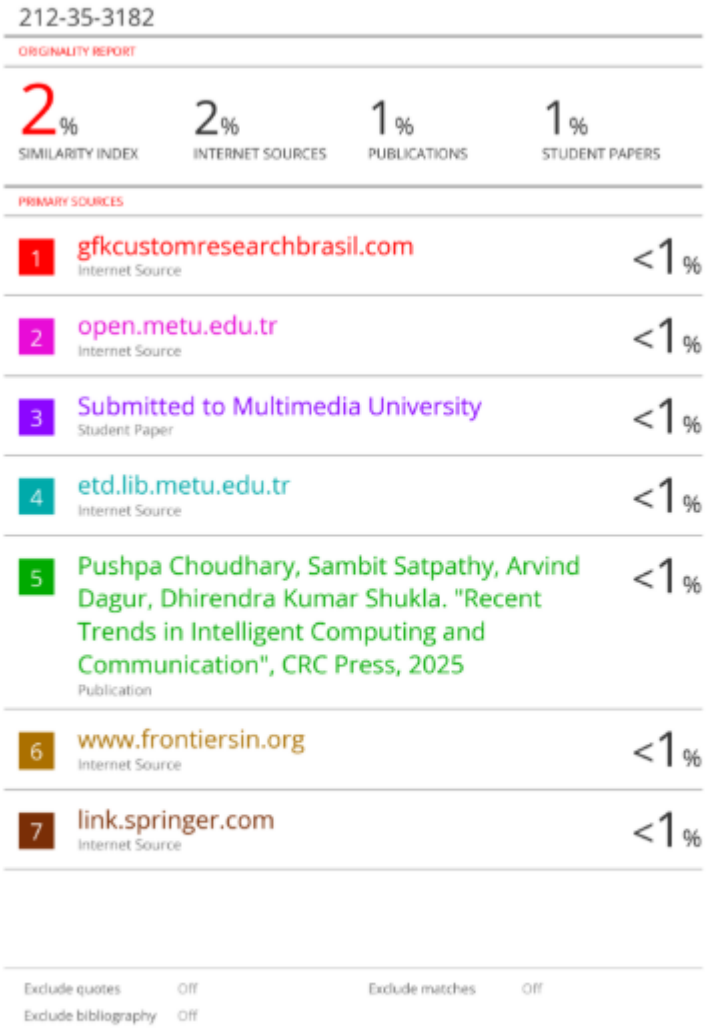
[6] N. Subramanian, "Image Steganography Using Deep Learning Methods," Qatar Univ., 2021.

[7] Y. Cheng et al., "RFNNS: Robust Fixed Neural Network Steganography," arXiv preprint, 2025.

[8] S. Utama, "Improving Stego Key via Extended Feature Coding in Text Steganography," Universiti Utara Malaysia, 2024.

[9] B. Al-Sarayefi, "Medical Image Steganography Optimization Using Genetic Algorithm," Universiti Utara Malaysia, 2023.

CHAPTER 6: Plagiarism Report



212-35-3182

ORIGINALITY REPORT

2%	2%	1%	1%
SIMILARITY INDEX	INTERNET SOURCES	PUBLICATIONS	STUDENT PAPERS

PRIMARY SOURCES

1	gfkcustomresearchbrasil.com Internet Source	<1%
2	open.metu.edu.tr Internet Source	<1%
3	Submitted to Multimedia University Student Paper	<1%
4	etd.lib.metu.edu.tr Internet Source	<1%
5	Pushpa Choudhary, Sambit Satpathy, Arvind Dagur, Dhirendra Kumar Shukla. "Recent Trends in Intelligent Computing and Communication", CRC Press, 2025 Publication	<1%
6	www.frontiersin.org Internet Source	<1%
7	link.springer.com Internet Source	<1%

Exclude quotes	Off	Exclude matches	Off
Exclude bibliography	Off		

CHAPTER 7: Account Clearance