International Research Journal of Innovations in Engineering and Technology (IRJIET)

ISSN (online): 2581-3048

Volume 10, Issue 5, pp 61-70, May-2026

https://doi.org/10.47001/IRJIET/2026.105009

A Taxonomy-Driven Survey of Deep Learning-Based Semantic Mapping Methods for Coverless Steganography

[1]Yaseen Hikmat Ismaiel, [2]Ali Farag Sultan, [3]Sadiq Sardar Sadiq, [4]Qutaiba Salim Murad, [5]Ahmed Waad Mohammed

[1,2,3,4,5]Department of Computer Science, College of Computer Sciences and Mathematics, University of Mosul, Mosul, Iraq

E-mail: [1]yaseen-hikmat@uomosul.edu.iq, [2]ali.25csp78@student.uomosul.edu.iq, [3]Sadiq.25csp67@student.uomosul.edu.iq, [4]qutaiba.25csp63@student.uomosul.edu.iq, [5]ahmed.25csp66@student.uomosul.edu.iq

Abstract **- Coverless steganography has emerged as a promising alternative to traditional information hiding techniques by eliminating direct modification of carrier media. Recently, deep learning (DL) has enabled a new paradigm known as deep semantic mapping, where secret information is encoded through high-level semantic representations extracted from images rather than pixel-level embedding. This approach significantly improves resistance against modern steganalysis methods based on statistical and neural network analysis. In this study, we aim to provide a systematic taxonomy and comparative analysis of deep learning-based coverless steganography methods with a focus on semantic mapping techniques. A comprehensive systematic survey of recent literature (2019 onward) is conducted, covering retrieval-based methods, generative adversarial networks (GANs), diffusion models, and Vision Transformer (ViT)-based approaches. A structured taxonomy is developed to classify existing methods into retrieval-based, generative-based, diffusion-based, transformer-based, and hybrid semantic mapping frameworks. A comparative analysis is performed based on accuracy, payload capacity, robustness, and visual quality. The analysis reveals a clear evolution from database-driven retrieval systems to advanced generative and transformer-based architectures. Recent diffusion and ViT-based models demonstrate superior robustness against compression, noise, and steganalysis attacks while maintaining high visual fidelity and improved semantic consistency. Deep semantic mapping significantly enhances the security and efficiency of coverless steganography. However, challenges remain in balancing payload capacity, computational complexity, and adversarial robustness. Future research should focus on hybrid architectures and real-time lightweight semantic encoding models.**

Keywords: Coverless steganography, Deep semantic mapping, Information hiding, Diffusion models, Vision Transformers.

I. INTRODUCTION

The rapid growth of digital communication systems and multimedia data exchange has significantly increased the demand for secure and covert information transmission [1]-[3]. While cryptographic techniques are widely used to protect the confidentiality of messages, they do not conceal the existence of communication itself. This visibility often attracts attention from adversaries, making cryptography insufficient in highly sensitive environments. To address this limitation, steganography has been introduced as a complementary security mechanism that focuses on hiding the existence of information within innocuous carrier media such as images, audio, and video [3].

Traditional image steganography techniques, such as Least Significant Bit (LSB) substitution and transform-domain embedding, directly modify pixel or frequency components of the cover image to encode secret data [4]. Although these modifications are typically imperceptible to human vision, they introduce statistical artifacts that can be detected by modern steganalysis techniques. In particular, the emergence of deep learning-based steganalysis models, especially those using Convolutional Neural Networks (CNNs), has significantly increased the ability to detect hidden information with high accuracy [5]. This has exposed a fundamental weakness in conventional steganographic approaches, making them increasingly unsuitable for high-security applications.

To overcome these limitations, a new paradigm known as coverless steganography has been introduced. Unlike traditional methods, coverless approaches do not modify the carrier image. Instead, they establish a mapping relationship between secret information and existing image content or semantic features [6]. Early coverless methods relied on large-scale image databases and handcrafted feature extraction techniques such as Scale-Invariant Feature Transform (SIFT). However, these approaches suffered from limited payload capacity, high retrieval complexity, and poor robustness against geometric and environmental transformations [7].

ISSN (online): 2581-3048
Volume 10, Issue 5, pp 61-70, May-2026
https://doi.org/10.47001/IRJIET/2026.105009

With the advancement of deep learning, particularly representation learning and generative modeling, a more powerful form of coverless steganography has emerged, known as Deep Semantic Mapping (DSM) [8]. In this paradigm, deep neural networks are used to extract high-level semantic features from images, enabling the encoding of secret messages into latent semantic spaces rather than pixel-level modifications. This shift from low-level feature matching to high-level semantic representation has significantly improved both robustness and scalability [9].

Recent developments have further transformed this field through the integration of advanced deep learning architectures such as Generative Adversarial Networks (GANs) [8]-[13], diffusion models [14]-[17], and Vision Transformers (ViTs) [18]. GAN-based approaches enable the synthesis of realistic images conditioned on secret data, while diffusion models provide improved image quality and resistance to compression and noise-based attacks [19]. Meanwhile, Vision Transformers enhance semantic feature extraction by capturing long-range dependencies within image representations. These advancements have collectively strengthened the performance of coverless steganography systems in terms of robustness, visual fidelity, and resistance to steganalysis [20].

Despite these improvements, several challenges remain unresolved, including the trade-off between payload capacity and semantic consistency, computational complexity of generative models, and vulnerability under adaptive adversarial attacks. Furthermore, there is still a lack of unified taxonomy and standardized evaluation frameworks for comparing different deep semantic mapping approaches.

Therefore, this study presents a systematic taxonomy-driven survey and comparative analysis of deep learning-based coverless steganography methods, with a particular focus on deep semantic mapping techniques. The main contributions of this work are as follows: (i) a structured classification of existing methods into retrieval-based, generative-based, diffusion-based, transformer-based, and hybrid approaches [21]–[23]; (ii) a comprehensive comparative analysis based on key performance metrics, including accuracy, payload capacity, robustness, and visual quality; and (iii) the identification of critical research gaps along with future research directions for the development of secure and efficient semantic information hiding systems.

The remainder of this paper is organized as follows: Section 2 introduces the taxonomy of existing approaches, Section 3 provides comparative analysis, Section 4 discusses research challenges, and Section 5 concludes the paper with future research directions, followed by the references.

II. DATA AND METHODOLOGY

In this section of our paper, we present the methodological framework adopted in this study to systematically review, classify, and analyze recent advances in Coverless Deep Semantic Mapping (CDSM) for steganography. The proposed methodology follows a structured survey-driven approach that combines literature collection, inclusion–exclusion filtering, taxonomy design, and comparative evaluation criteria definition.

2.1 Research Design

This study is conducted as a taxonomy-driven systematic survey of existing literature on deep learning-based coverless steganography. Unlike experimental research, this work does not propose a new algorithm; instead, it focuses on organizing, synthesizing, and critically analyzing existing approaches in the domain. The survey process is structured into four main stages:

1) Collection of relevant literature from scientific databases
2) Selection of high-quality studies using predefined criteria
3) Development of a taxonomy for method categorization
4) Comparative analysis using standardized evaluation metrics

2.2 Literature Collection Strategy

Relevant research articles were collected from major scientific databases, including IEEE Xplore, SpringerLink, Elsevier ScienceDirect, and Google Scholar. The search was conducted using combinations of the following keywords:

"coverless steganography"

"deep semantic mapping"

"information hiding deep learning"

"GAN steganography"

"diffusion model steganography"

"vision transformer security"

The time range of selected studies spans from 2019 onward to ensure coverage of both early and recent developments in the field.

2.3 Inclusion and Exclusion Criteria

In order to ensure the quality and relevance of selected studies, the following criteria were applied:

- **Inclusion Criteria:**
  1) Peer-reviewed journal and conference papers.

ISSN (online): 2581-3048
Volume 10, Issue 5, pp 61-70, May-2026
https://doi.org/10.47001/IRJIET/2026.105009

2) Studies focusing on coverless or semantic-based steganography.
3) Approaches using deep learning techniques (CNN, GAN, Diffusion, ViT).
4) Papers providing evaluation metrics (accuracy, robustness, payload, or visual quality).

➢ **Exclusion Criteria:**

1) Traditional steganography methods based only on pixel embedding.
2) Non-peer-reviewed articles, blogs, or technical reports.
3) Studies without experimental or comparative evaluation.
4) Duplicate or irrelevant publications.

2.4 Taxonomy Development Approach

A hierarchical taxonomy-driven framework is proposed to systematically classify existing CDSM methods based on their underlying deep learning paradigms and information encoding strategies. The objective of this taxonomy is to provide a structured understanding of the evolution of semantic information hiding techniques and to support a consistent comparative analysis across different methodological families. The proposed taxonomy categorizes existing approaches into five primary classes:

2.4.1 Retrieval-Based Methods: Retrieval-based approaches represent the earliest form of coverless steganography [7], where secret information is mapped to pre-existing images without modifying their content. These methods rely on handcrafted feature descriptors such as Scale-Invariant Feature Transform (SIFT) or learned representations extracted using Convolutional Neural Networks (CNNs). The mapping process typically involves matching secret data patterns with indexed image features stored in large-scale databases. Although these methods provide inherent resistance to steganalysis due to the absence of embedding, they suffer from limited payload capacity and high computational complexity during retrieval.

2.4.2 Generative-Based Methods: Generative-based approaches employ Generative Adversarial Networks (GANs) to synthesize images conditioned on secret information [6], [16], [19]. In this paradigm, the secret message is embedded into the latent space of the generator, enabling the creation of realistic images that inherently encode semantic content. This eliminates the need for external image databases and improves system scalability. However, such methods may introduce challenges related to training instability and the presence of visual artifacts in generated outputs.

2.4.3 Diffusion-Based Methods: Diffusion-based methods utilize probabilistic generative models to iteratively refine noise into high-quality images guided by latent semantic representations [10], [17], [21]. These approaches offer improved stability and superior visual fidelity compared to GAN-based frameworks. Additionally, diffusion models demonstrate strong resistance to compression and noise-based attacks, making them highly suitable for secure semantic information hiding applications.

2.4.4 Transformer-Based Methods: Transformer-based approaches leverage Vision Transformers (ViT) or related attention mechanisms to model long-range dependencies within image representations [11], [22]. Unlike convolutional architectures, Transformers enable global semantic feature extraction, enhancing the accuracy of message-to-feature mapping. These methods provide improved robustness and scalability, particularly in complex visual environments with high semantic variability.

2.4.5 Hybrid Methods (Emerging Paradigm): Hybrid approaches [23] integrate multiple deep learning paradigms, such as diffusion models combined with multi-image fusion or transformer-enhanced generative systems. These frameworks aim to improve security, robustness, and payload capacity simultaneously by leveraging complementary strengths of different architectures. Although still in early development stages, hybrid models represent a promising direction for next-generation coverless steganography systems. Figure (1) illustrates the categorizes of existing CDSM approaches.

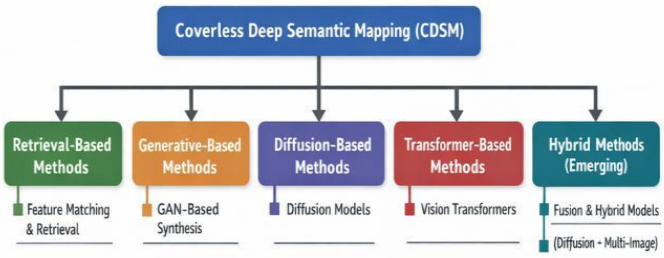

Figure 1: Classification of existing CDSM approaches based on their methodological characteristics

Accordingly, this taxonomy serves as the foundational framework for the comparative analysis presented in subsequent sections, enabling a structured evaluation of existing methods in terms of robustness, payload capacity, visual quality, and computational efficiency.

In order to complement the proposed visualization of CDSM, Table (1) presents a structured tabular representation accompanied by an analytical description. This hybrid format enhances interpretability, reproducibility, and academic rigor by combining visual taxonomy concepts with formal comparative structuring.

ISSN (online): 2581-3048
Volume 10, Issue 5, pp 61-70, May-2026
https://doi.org/10.47001/IRJIET/2026.105009

Table 1: Taxonomy of Coverless Deep Semantic Mapping Methods

Category	Core Principle	Technique	Key Advantage	Main Limitation
Retrieval-Based Methods	Map secret data to pre-existing images using feature matching	SIFT-based retrieval, CNN feature indexing	No modification of carrier images; inherent steganalysis resistance	Low payload capacity; high retrieval cost
Generative-Based Methods	Generate images conditioned on latent semantic representations	GAN, Conditional GAN	High scalability; eliminates database dependency	Training instability; possible artifacts
Diffusion-Based Methods	Iterative denoising guided by semantic embeddings	Latent Diffusion Models	High image fidelity; strong robustness	High computational cost
Transformer-Based Methods	Global semantic modeling using attention mechanisms	Vision Transformer (ViT)	Strong semantic representation; high accuracy	Resource-intensive training
Hybrid Methods	Integration of multiple deep learning paradigms	Diffusion + Transformer, Multi-image fusion	Best overall performance and robustness	High complexity and deployment cost

The proposed taxonomy highlights the fundamental shift in CDSM from traditional retrieval-based mechanisms toward advanced deep learning-driven semantic modeling. Early retrieval-based approaches rely on static feature descriptors and database indexing, which provide basic security benefits but suffer from limited scalability and constrained information capacity. Generative-based methods introduce a significant paradigm shift by leveraging adversarial learning frameworks to directly synthesize images conditioned on semantic representations. This eliminates the dependency on pre-stored image databases and improves system flexibility; however, it introduces challenges related to training stability and visual consistency. Diffusion-based models represent a more recent advancement, offering a probabilistic generative process that gradually refines noise into semantically consistent and visually realistic images. These methods significantly improve robustness and perceptual quality but require substantial computational resources. Transformer-based architectures further enhance semantic encoding by capturing global contextual dependencies through attention mechanisms. This leads to improved accuracy in mapping secret information into latent feature spaces, particularly in complex image distributions.

Finally, hybrid approaches combine multiple architectures, such as diffusion models integrated with transformers or multi-image fusion strategies. These methods achieve the most balanced performance across all evaluation metrics, including robustness, payload capacity, and visual fidelity. However, their increased architectural complexity introduces significant computational overhead, which remains a key barrier to real-world deployment.

2.5 Evaluation Metrics

In order to ensure a fair, consistent, and systematic comparison across different CDSM approaches, a unified set of evaluation metrics is adopted. These metrics capture the key performance dimensions relevant to deep learning-based information hiding systems, including correctness of message recovery, embedding efficiency, robustness against attacks, perceptual image quality, and computational efficiency.

2.5.1 Accuracy: Accuracy refers to the correctness of secret message recovery at the receiver side [20]. It is typically measured as the percentage of correctly decoded or reconstructed messages from the transmitted or generated image representations. Higher accuracy indicates more reliable semantic mapping between the secret information and the visual features.

2.5.2 Payload Capacity: Payload capacity [24] represents the amount of secret information that can be encoded per image, typically expressed in bits per image (bits/image). In coverless and semantic-based approaches, payload capacity depends on the richness of the semantic representation space and the efficiency of the mapping function. A higher payload indicates improved information density without compromising security.

2.5.3 Robustness: Robustness measures the resistance of a method against intentional and unintentional distortions [21]. This includes common image processing and adversarial attacks such as JPEG compression, Gaussian noise, cropping, resizing, and steganalysis detection models. A robust system maintains stable message recovery performance even under such perturbations, reflecting its suitability for real-world deployment.

2.5.4 Visual Quality: Visual quality evaluates the perceptual realism and fidelity of generated or retrieved images [20]. In generative-based and diffusion-based methods, it is commonly quantified using metrics such as Peak Signal-to-Noise Ratio (PSNR), Structural Similarity Index Measure (SSIM), or Fréchet Inception Distance (FID), depending on the availability of ground truth references. High visual quality ensures that the carrier image remains indistinguishable from natural images, reducing the risk of detection.

2.5.5 Computational Complexity: Computational complexity refers to the resource consumption and scalability of the proposed method, including time complexity, memory usage, and training/inference cost [11]. This metric is particularly important for deep learning-based models such as GANs, diffusion models, and Vision Transformers, which often

International Research Journal of Innovations in Engineering and Technology (IRJIET)
ISSN (online): 2581-3048
Volume 10, Issue 5, pp 61-70, May-2026
https://doi.org/10.47001/IRJIET/2026.105009

require significant computational resources. Efficient models achieve a balance between performance and computational cost, enabling practical deployment in real-time or large-scale systems.

Based on these evaluation metrics, as shown in Figure (2), we are able to provide a standardized framework for assessing and comparing different CDSM approaches. They also enable a comprehensive analysis of their performance trade-offs across security, efficiency, and visual fidelity dimensions.

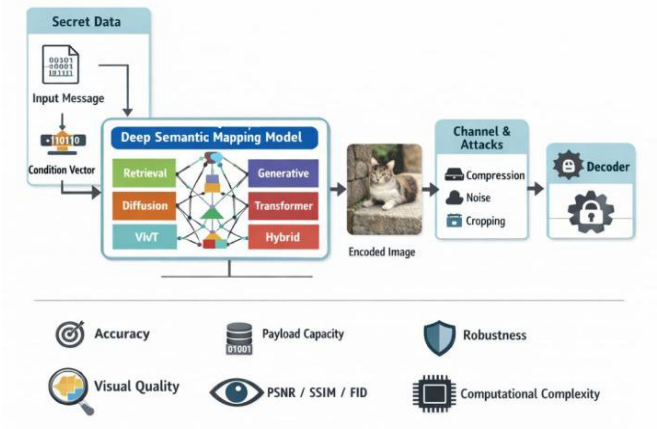

Figure 2: Integrated evaluation metrics framework and system architecture for the standardized assessment of CDSM approaches, emphasizing trade-offs in security, computational cost, and visual fidelity

The system architecture of CDSM, as demonstrated in Figure (2), follows a structured pipeline that ensures secure and indirect transmission of secret information. Initially, the secret message is transformed into a semantic condition vector, which serves as the input to the deep semantic mapping model. Depending on the selected architecture—retrieval-based, generative-based, diffusion-based, transformer-based, or hybrid—the model either retrieves a semantically similar image or generates a new image that encodes the hidden message within its latent representation. The resulting encoded image is then transmitted through a potentially hostile communication channel, where it may be subjected to compression, noise injection, or cropping attacks.

Despite these distortions, the semantic integrity of the encoded information is preserved due to the robustness of deep feature representations.

Finally, the receiver employs a decoding mechanism to extract the semantic features from the received image and reconstruct the original secret message. This pipeline highlights the fundamental advantage of CDSM systems: information is never explicitly embedded but implicitly represented in semantic space, significantly reducing detectability by traditional steganalysis methods. Furthermore, Table (2) presents a system architecture of CDSM-based steganography.

Table 2: System architecture of CDSM-based steganography

Stage	Input	Process	Output	Function
Secret Data Encoding	Raw message (text/binary)	Semantic conversion into latent vector	Condition vector (semantic embedding)	Converts message into model-compatible representation
Deep Semantic Mapping	Condition vector	Processing via retrieval/ generative/ diffusion / transformer models	Encoded semantic representation	Maps message into image feature space
Image Generation / Selection	Semantic representation	Image synthesis or retrieval	Encoded image	Produces carrier image containing semantic information
Transmission Channel	Encoded image	Exposure to attacks (compression, noise, cropping)	Distorted image	Simulates real-world communication environment
Decoder Module	Distorted/received image	Feature extraction + reverse mapping	Recovered message	Extracts original secret information

2.6 Comparative Analysis Procedure

The comparative analysis in this study is conducted by systematically mapping each reviewed CDSM method to the proposed taxonomy and evaluating its performance according to the defined evaluation metrics. Thus, this approach ensures a structured and uniform comparison across heterogeneous methods originating from different architectural paradigms, including retrieval-based, generative-based, diffusion-based, and transformer-based frameworks.

Since many existing studies do not report complete or directly comparable numerical results, a normalized qualitative scale is adopted to standardize the evaluation. Each method is assessed using a four-level ordinal scale: low, medium, high, and very high, depending on its reported performance trends and relative position within the literature. This normalization enables consistent comparison across studies with differing experimental setups, datasets, and evaluation protocols. The comparative analysis particularly focuses on the following key dimensions:

- **Evolution of Model Complexity Over Time:** The study examines how CDSM methods have evolved from simple retrieval-based systems to highly complex deep generative and transformer-based architectures, highlighting the increasing role of deep learning in semantic representation learning.
- **Trade-off Between Payload Capacity and Robustness:** A critical analysis is performed to investigate the inherent trade-off between embedding capacity and resistance to attacks. In general, higher payload capacity

ISSN (online): 2581-3048
Volume 10, Issue 5, pp 61-70, May-2026
https://doi.org/10.47001/IRJIET/2026.105009

may reduce robustness, whereas more robust systems often limit information density.

- **Impact of Deep Learning Architectures on Semantic Encoding Efficiency:** The influence of different deep learning models—such as CNNs, GANs, diffusion models, and Vision Transformers—is analyzed in terms of their ability to encode, preserve, and reconstruct semantic information effectively within latent feature spaces.

Altogether, this comparative framework is capable of providing a consistent and interpretable methodology for evaluating diverse CDSM approaches. It also enables clearer insights into their strengths, limitations, and suitability for secure information hiding applications.

2.7 Research Framework Overview

The overall methodological framework of this study is designed to ensure a structured, systematic, and reproducible analysis framework of CDSM techniques within the domain of deep learning-based information hiding. The framework integrates systematic survey principles with taxonomy construction and comparative evaluation to provide a comprehensive understanding of the field. Therefore, the proposed research workflow can be summarized into three main stages:

- **Input Stage:** This stage involves the collection of relevant scholarly publications related to CDSM, coverless steganography, and deep learning-based information hiding. Sources include peer-reviewed journal articles and conference papers retrieved from major scientific databases. The selected studies span recent advancements in retrieval-based, generative-based, diffusion-based, and transformer-based approaches.
- **Processing Stage:** In this stage, the collected literature undergoes a structured refinement process consisting of:
  - **Filtering:** Removal of irrelevant, duplicate, or non-peer-reviewed studies.
  - **Classification:** Mapping each study to the proposed taxonomy categories.
  - **Feature Extraction:** Identification of key attributes such as model type, architecture, dataset usage, and evaluation metrics. This stage ensures that only relevant and high-quality studies are included in the comparative analysis.
- **Output Stage:** The final stage produces the main outcomes of the study, including: a structured taxonomy of CDSM methods, a comparative evaluation based on unified performance metrics, and identification of key research gaps and limitations in existing approaches

Based on this structured workflow, we are capable of ensuring a comprehensive, systematic, and reproducible survey-based analysis of the current state of Coverless Deep Semantic Mapping techniques. In addition, it provides a solid foundation for future research in secure deep learning-based information hiding systems.

III. COMPARATIVE ANALYSIS RESULTS

In this section of our study, we present the results of the comparative evaluation of existing CDSM methods based on the proposed taxonomy and unified evaluation metrics. Each method is assessed across five key dimensions: accuracy, payload capacity, robustness, visual quality, and computational complexity. The analysis enables a structured comparison of heterogeneous approaches under a consistent evaluation framework.

The results reveal a clear evolutionary trend in coverless steganography, where performance improvements are strongly associated with the progressive adoption of advanced deep learning architectures. This evolution reflects a shift from handcrafted feature-based techniques toward data-driven semantic representation learning in latent spaces. Retrieval-based methods exhibit relatively limited payload capacity and moderate robustness, primarily due to their dependence on static feature descriptors and large-scale image database indexing. Although these methods avoid direct modification of carrier images, their scalability and efficiency are constrained by retrieval overhead and limited semantic expressiveness. In contrast, generative-based approaches based on Generative Adversarial Networks (GANs) significantly enhance both scalability and payload capacity by enabling direct synthesis of images conditioned on latent semantic representations. However, these methods may suffer from training instability and occasional visual artifacts, which can affect perceptual consistency.

Diffusion-based models demonstrate further improvements by generating high-fidelity images through iterative denoising processes guided by semantic embeddings. These models achieve superior robustness against common signal processing attacks such as JPEG compression, Gaussian noise, and cropping, while maintaining strong resistance to steganalysis techniques. Transformer-based approaches, particularly those utilizing Vision Transformers (ViT), provide enhanced global semantic understanding through self-attention mechanisms. This capability improves the accuracy of mapping between secret information and visual representations, especially in complex and high-dimensional image spaces. Hybrid models, which integrate multiple architectures such as diffusion processes combined with multi-image fusion or transformer-enhanced generative frameworks,

ISSN (online): 2581-3048
Volume 10, Issue 5, pp 61-70, May-2026
https://doi.org/10.47001/IRJIET/2026.105009

achieve the highest overall performance in terms of robustness, visual quality, and semantic consistency.

However, this performance gain is accompanied by increased computational complexity and resource requirements, which may limit their applicability in real-time or resource-constrained environments. The results of a comparative summary are presented in Table (3). Consequently, this comparative framework provides a unified benchmark perspective for future CDSM research and highlights key architectural trade-offs that must be balanced in practical steganographic system design.

Table 3: The comparative summary of CDSM approaches

Method Type	Accuracy	Payload	Robustness	Visual Quality	Complexity
Retrieval-based	Medium	Low	Medium	High	High
GAN-based	High	High	Medium	Medium	Medium
Diffusion-based	Very High	Medium	Very High	Very High	High
Transformer-based	High	Medium	High	High	Very High
Hybrid models	Very High	Very High	Very High	Very High	Very High

Overall, the comparative analysis confirms that deep learning-based semantic mapping approaches significantly outperform traditional coverless steganography techniques in terms of security, robustness, and representational efficiency. Nevertheless, the trade-offs between computational cost, payload capacity, and robustness remain a critical challenge for future research in this domain.

IV. DISCUSSION

In this study, we presented a comprehensive comparative evaluation of CDSM methods using a unified taxonomy and standardized evaluation metrics, enabling a consistent cross-paradigm analysis of heterogeneous approaches. The results clearly demonstrate that the evolution of CDSM is fundamentally driven by the transition from retrieval-based frameworks to deep generative and transformer-based architectures, reflecting a broader paradigm shift toward semantic-aware representation learning in high-dimensional latent spaces. From a performance perspective, retrieval-based methods, while inherently secure due to their non-modification of carrier media, exhibit structural limitations in scalability, payload capacity, and semantic flexibility. Their reliance on predefined feature descriptors and large-scale image repositories introduces computational bottlenecks and restricts adaptability to dynamic or previously unseen semantic distributions.

Consequently, their applicability in modern high-capacity covert communication systems remains constrained. The introduction of generative models, particularly those based on Generative Adversarial Networks (GANs), represents a critical inflection point in the development of CDSM. These approaches significantly improve payload capacity and scalability by enabling the synthesis of semantically meaningful images directly from latent codes. However, GAN-based methods are inherently sensitive to training instability, mode collapse, and distributional inconsistencies, which can compromise visual fidelity and reduce robustness under adversarial or noisy conditions. Diffusion-based models further advance the state-of-the-art by leveraging iterative denoising processes that enhance both visual quality and robustness. Their probabilistic formulation enables stable training dynamics and improved resistance to common distortions such as compression, noise injection, and geometric transformations. This robustness is particularly relevant in adversarial environments where steganographic systems are subject to active detection and signal degradation. Transformer-based architectures, especially Vision Transformers (ViT), contribute an additional layer of sophistication by enabling global semantic reasoning through self-attention mechanisms. This facilitates more accurate alignment between hidden information and visual representations, particularly in complex or high-entropy image domains. However, the computational overhead associated with self-attention operations introduces significant latency and resource demands, limiting their practicality in real-time or embedded systems.

Hybrid models, which integrate diffusion processes, transformer-based encoders, and multi-image fusion strategies, achieve the highest overall performance across evaluation dimensions. These architectures effectively combine complementary strengths, resulting in superior robustness, perceptual quality, and semantic consistency. Nonetheless, such gains are accompanied by increased computational complexity, memory requirements, and system design overhead, highlighting a persistent trade-off between performance and efficiency.

In addition, there are several key insights that contribute to a deeper understanding of CDSM system design:

a) **Paradigm Shift Toward Latent Semantic Encoding:** There is a clear transition from explicit feature matching to implicit semantic encoding in latent spaces, enabling more flexible, scalable, and high-capacity information embedding [25].
b) **Trade-off Between Payload Capacity and Robustness:** While generative and hybrid models significantly increase payload capacity, maintaining robustness

ISSN (online): 2581-3048
Volume 10, Issue 5, pp 61-70, May-2026
https://doi.org/10.47001/IRJIET/2026.105009

against distortions and steganalysis remains a non-trivial challenge, particularly under real-world transmission conditions [26].

c) **Robustness as a Function of Generative Stability:** Model robustness is closely linked to the stability of the underlying generative process. Diffusion models outperform GANs in this regard due to their probabilistic and iterative refinement mechanisms [27].

d) **Computational Complexity as a Limiting Factor:** Despite performance gains, advanced architectures (e.g., transformers and hybrid models) impose substantial computational costs, which may hinder deployment in resource-constrained or real-time environments [11].

e) **Importance of Global Semantic Awareness:** Transformer-based approaches demonstrate that capturing long-range dependencies and global context significantly improves semantic fidelity and decoding accuracy in coverless steganography systems [28].

Despite significant progress, several open challenges and research opportunities remain in the field of CDSM:

a) **Lightweight and Efficient Model Design:** Future work should focus on developing computationally efficient architectures, such as model compression, knowledge distillation, and edge-optimized diffusion or transformer variants, to enable real-time deployment.

b) **Robustness Against Advanced Steganalysis:** There is a need for designing adversarially robust CDSM systems capable of resisting both classical and deep learning-based steganalysis techniques, potentially through adversarial training and game-theoretic optimization.

c) **Adaptive and Context-Aware Semantic Mapping:** Incorporating context-aware and dynamic semantic encoding mechanisms can improve adaptability to varying communication environments and enhance resilience against distribution shifts.

d) **Multi-Modal and Cross-Domain CDSM:** Extending CDSM beyond images to multi-modal domains (e.g., text, audio, video) represents a promising direction for increasing embedding capacity and application versatility.

e) **Explainability and Interpretability:** As deep models become more complex, incorporating Explainable AI (XAI) techniques is essential to improve transparency, trustworthiness, and forensic analysis of steganographic systems.

f) **Benchmarking and Standardization:** The lack of standardized datasets and evaluation protocols remains a critical gap. Establishing public benchmarks and reproducible evaluation frameworks will facilitate fair comparison and accelerate research progress.

V. CONCLUSION

This paper presented a comprehensive and systematic review of CDSM methods within the domain of deep learning-based steganography, supported by a structured taxonomy and a unified evaluation framework that enabled consistent comparison across retrieval-based, generative-based, diffusion-based, transformer-based, and hybrid approaches. The analysis revealed a clear evolutionary trajectory from traditional retrieval-driven frameworks toward advanced deep learning architectures that leverage latent semantic representation learning, resulting in significant improvements in robustness, payload capacity, and visual fidelity, as well as enhanced resistance to modern steganalysis techniques.

In particular, diffusion-based and transformer-based models demonstrated superior performance due to their ability to capture complex semantic relationships and generate high-quality, semantically consistent visual representations, while hybrid architectures achieved state-of-the-art results by integrating complementary strengths of multiple paradigms.

Nevertheless, several limitations must be acknowledged, including the reliance on reported results from heterogeneous studies with varying experimental setups, the absence of standardized benchmarking datasets and evaluation protocols, and the use of qualitative normalization scales that may obscure fine-grained performance differences. Furthermore, most existing approaches are evaluated under controlled conditions, with limited consideration of real-world deployment constraints such as computational efficiency, latency, and adversarial communication environments, and the focus remains largely restricted to image-based applications. To address these challenges, future research should prioritize the development of standardized and reproducible benchmarking frameworks, the design of lightweight and computationally efficient models suitable for real-time and edge deployment, and the incorporation of adversarially robust training strategies to defend against increasingly sophisticated steganalysis techniques.

Additionally, advancing adaptive and context-aware semantic encoding mechanisms, extending CDSM to multi-modal and cross-domain scenarios, integrating explainable artificial intelligence for improved transparency and interpretability, and aligning these systems with secure communication paradigms will be essential for enabling practical, scalable, and trustworthy covert communication systems. Overall, while deep semantic mapping has significantly advanced the field of coverless steganography, achieving an optimal balance between robustness, payload capacity, and computational efficiency remains a fundamental challenge that will shape future research directions.

ISSN (online): 2581-3048
Volume 10, Issue 5, pp 61-70, May-2026
https://doi.org/10.47001/IRJIET/2026.105009

CONFLICTS OF INTEREST

The authors declare that there are no conflicts of interest regarding this study.

FUNDING

The authors received no specific funding for this study.

ACKNOWLEDGMENT

We would like to express sincere gratitude to **Assistant Prof. Dr. Yaseen Hikmat Ismaiel** for his invaluable supervision, continuous guidance, and insightful feedback throughout the course of this research. We also acknowledge the support of the Department of Computer Science, College of Computer Science and Mathematics, University of Mosul.

REFERENCES

[1] Y. A. Hamza and W. M. Abduallah, "A Secured Method of Reversible Data Hiding," *2020 International Conference on Computer Science and Software Engineering (CSASE), Duhok, Iraq,* 2020, pp. 61-66, doi: 10.1109/CSASE48920.2020.9142098.

[2] Y. H. Ismael, "Improved security using two levels of steganography," *AIP Conference Proceedings*, vol. 3264, no. 1, p. 030010, Mar. 2025, doi: 10.1063/5.0259914.

[3] J. Fridrich, "Steganography in Digital Media: Principles, Algorithms, and Applications," *IEEE Signal Processing Magazine*, vol. 36, no. 5, pp. 12–20, 2019.

[4] B. Li, M. Wang, J. Huang, and X. Li, "A New Cost Function for Spatial Image Steganography," *IEEE Transactions on Information Forensics and Security*, vol. 14, no. 2, pp. 512–526, 2019.

[5] Y. Qian, J. Dong, W. Wang, and T. Tan, "Deep Learning for Steganalysis via Convolutional Neural Networks," *IEEE Transactions on Information Forensics and Security*, vol. 15, pp. 124–135, 2020.

[6] D. Hu, L. Wang, W. Jiang, S. Zheng, and B. Li, "A Novel Image Steganography Method via Deep Convolutional Generative Adversarial Networks," *IEEE Access*, vol. 8, pp. 148948–148963, 2020.

[7] X. Zhang, F. Peng, and M. Long, "Robust Coverless Image Steganography Based on Image Retrieval and CNN Features," *Multimedia Tools and Applications*, vol. 79, pp. 25991–26011, 2020.

[8] T. Karras, S. Laine, M. Aittala, J. Hellsten, J. Lehtinen, and T. Aila, "Analyzing and Improving the Image Quality of StyleGAN," *in Proc. CVPR*, 2020.

A. Jolicoeur-Martineau, "The relativistic discriminator: a key element missing from standard GAN," *International Conference on Learning Representations (ICLR)*, 2020.

[9] J. Ho, A. Jain, and P. Abbeel, "Denoising Diffusion Probabilistic Models," *in Advances in Neural Information Processing Systems (NeurIPS)*, 2020.

A. Dosovitskiy et al., "An Image is Worth 16x16 Words: Transformers for Image Recognition at Scale," *in Proc. ICLR*, 2021.

B. Saharia et al., "Image Super-Resolution via Iterative Refinement," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2022.

[10] H. Wu, Y. Shi, and H. Wang, "Deep Learning-Based Robust Image Steganography Using GANs," *IEEE Access*, vol. 9, pp. 39821–39832, 2021.

[11] P. Dhariwal and A. Nichol, "Diffusion Models Beat GANs on Image Synthesis," *in NeurIPS*, 2021.

[12] Z. Zhang, Y. Chen, and F. Huang, "Coverless Image Steganography Based on Deep Learning," *Signal Processing*, vol. 187, 2021.

[13] S. Tan and B. Li, "Stacked Convolutional Auto-Encoders for Steganalysis of Digital Images," *IEEE Transactions on Information Forensics and Security*, vol. 16, pp. 454–467, 2021.

[14] R. Rombach, A. Blattmann, D. Lorenz, P. Esser, and B. Ommer, "High-Resolution Image Synthesis with Latent Diffusion Models," *in CVPR*, 2022.

[15] Y. Chen, Z. He, and L. Wang, "Vision Transformer-Based Secure Image Representation for Information Hiding," Pattern Recognition Letters, vol. 167, pp. 10–18, 2023.

[16] K. He, X. Zhang, S. Ren, and J. Sun, "Deep Residual Learning for Image Recognition," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2022.

[17] Y. Luo, J. Qin, and Y. Xiang, "Coverless Information Hiding Based on Generative Models," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 32, no. 6, pp. 3456–3469, 2022.

[18] J. Wang, L. Liu, and S. Xiang, "A Survey on Coverless Information Hiding," *ACM Computing Surveys*, vol. 55, no. 3, pp. 1–36, 2023.

[19] H. Zhang, Y. Li, and X. Zhang, "Diffusion-Based Image Steganography with High Robustness," *IEEE Signal Processing Letters*, vol. 30, pp. 450–454, 2023.

[20] X. Liu, Q. Wu, and W. Zhang, "Hybrid Deep Learning Framework for Secure Image Steganography," *IEEE Access*, vol. 11, pp. 98765–98780, 2023.

[21] Y. A. Hamza, N. E. Tewfiq, and M. Q. Ahmed, "An enhanced approach of image steganographic using discrete shearlet transform and secret sharing," *Baghdad Science Journal*, vol. 19, no. 1, Art. no. 2, 2022, doi: 10.21123/bsj.2022.19.1.0197.

ISSN (online): 2581-3048
Volume 10, Issue 5, pp 61-70, May-2026
https://doi.org/10.47001/IRJIET/2026.105009

[22] OpenAI, "DALL·E 2: Hierarchical Text-Conditional Image Generation with CLIP Latents," 2022.
[23] Stability AI, "Stable Diffusion: High-Resolution Image Synthesis with Latent Diffusion Models," 2023.
[24] K. Nichol and P. Dhariwal, "Improved Denoising Diffusion Probabilistic Models," *in ICML*, 2021.
[25] Y. LeCun, Y. Bengio, and G. Hinton, "Deep Learning," Nature, vol. 521, pp. 436–444, *updated perspectives* 2020.

AUTHORS BIOGRAPHY

Assistant Professor Dr. Yaseen Hikmat Ismaiel is a faculty member at the College of Computer Science and Mathematics, University of Mosul. His research focuses on information security, particularly steganography and secure data hiding.

Ali Farag Sultan is an MSc student in the Department of Computer Science. His research interests focus on cybersecurity, data protection, and secure communication systems.

Sadiq Sardar Sadiq is an MSc student in the Department of Computer Science. His academic interests include network security, intrusion detection systems, and data analytics.

Qutaiba Salim Murad is an MSc student in the Department of Computer Science. His research focuses on artificial intelligence, data mining, and intelligent systems.

Ahmed Waad Mohammed is an MSc student in the Department of Computer Science. His research interests include software engineering, information security, and distributed systems.

Citation of this Article:

Yaseen Hikmat Ismaiel, Ali Farag Sultan, Sadiq Sardar Sadiq, Qutaiba Salim Murad, & Ahmed Waad Mohammed. (2026). A Taxonomy-Driven Survey of Deep Learning-Based Semantic Mapping Methods for Coverless Steganography. *International Research Journal of Innovations in Engineering and Technology - IRJIET*, 10(5), 61-70. Article DOI https://doi.org/10.47001/IRJIET/2026.105009

****

