A Survey of Image Steganography and Generative Models in Artificial Intelligence

www.surveyx.cn

Abstract

Image steganography and generative models, particularly Generative Adversarial Networks (GANs), represent a convergence of advanced computational techniques aimed at secure communication, content generation, and data manipulation. This survey explores the integration of artificial intelligence (AI) in these domains, highlighting the transformative impact of AI-driven methodologies on traditional steganographic practices. Traditional methods, such as Least Significant Bit (LSB) modification, face limitations in payload capacity and detectability, prompting the evolution toward AI-enhanced techniques. GANs have introduced sophisticated data embedding strategies, enhancing imperceptibility and robustness against detection. The survey underscores AI's dual role in secure communication and content generation, with applications extending to healthcare, digital media, and beyond. Moreover, it examines the security and privacy implications of AI-generated content (AIGC), emphasizing the need for robust countermeasures against potential misuse. The integration of AI with communication networks presents opportunities for personalized content delivery, though challenges in data privacy and computational efficiency persist. The survey concludes by identifying future directions, including the development of adaptive Digital Rights Management (DRM) systems and the exploration of federated learning for privacy-preserving generative models. As AI technologies continue to evolve, their role in image steganography and generative modeling is expected to expand, offering innovative solutions to contemporary challenges in secure communication and digital content creation.

1 Introduction

1.1 Overview of Image Steganography and Generative Models

Image steganography is a pivotal technique for secure communication, allowing the concealment of information within digital images while maintaining visual integrity [1]. Traditional methods, such as the Least Significant Bit (LSB) technique, have been foundational but are limited by payload capacity and susceptibility to detection [2]. Recent advancements utilize deep neural networks to enhance steganographic capacity, enabling the embedding of multiple high-resolution images and improving both capacity and security [3].

Generative models, particularly Generative Adversarial Networks (GANs), have transformed image steganography by introducing advanced data embedding techniques. Originally designed for realistic image generation, GANs have been adapted to enhance the imperceptibility and robustness of steganographic methods against detection [4]. The combination of GANs with Convolutional Neural Networks (CNNs) has fortified steganographic systems, addressing challenges from adversarial perturbations [2].

The application of generative models extends beyond image steganography to various fields, including identity-preserving face generation and improved social media bot detection through enhanced classification accuracy [5]. Their relevance in industrial time series data and the 6G-enabled In-

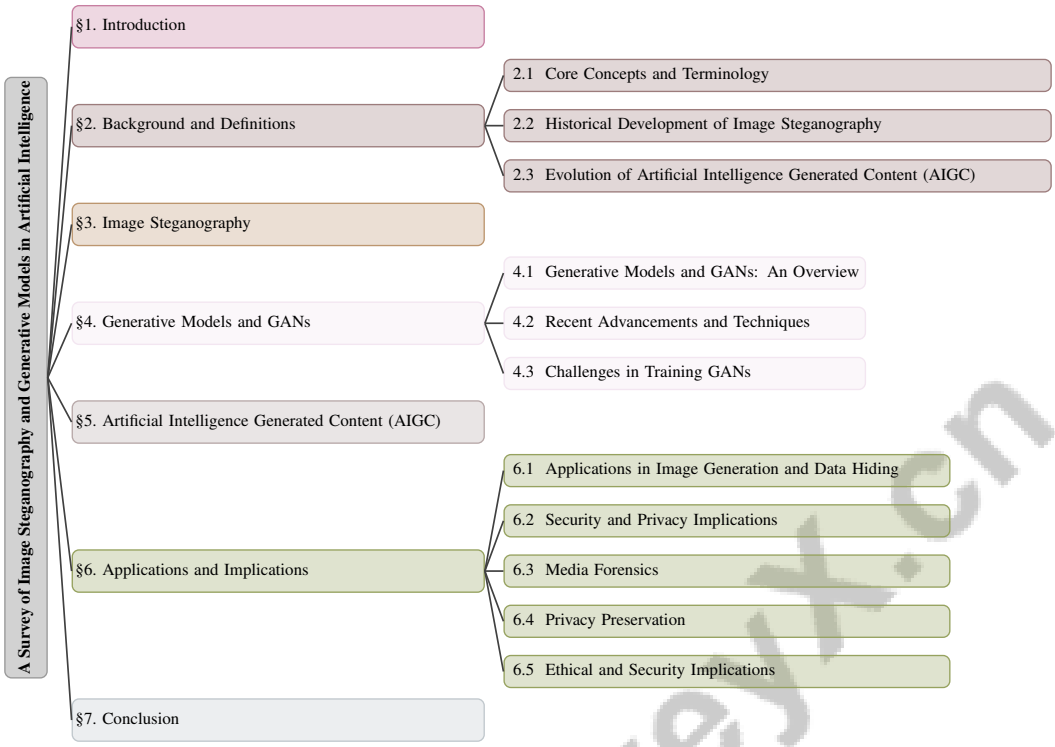


Figure 1: chapter structure

ternet of Things (IoT) underscores the necessity for robust cybersecurity in complex technological environments [6].

Innovative approaches in image steganography, such as coverless image steganography (CIS), enhance imperceptibility by eliminating the need for a cover image, presenting a novel paradigm in data hiding [4]. Additionally, the integration of generative models in visual compression has facilitated ultra-low bitrate communication, enabling intelligent machine analysis and broadening the application scope of steganography [3].

The intersection of image steganography and generative models represents a vibrant research area, promising innovative solutions to the challenges of data hiding and manipulation in the digital age. The ongoing evolution of these technologies is expected to expand steganography's applications, offering significant implications for secure communication and content generation [7].

1.2 Significance of AI in Secure Communication and Content Generation

Artificial Intelligence (AI) has emerged as a transformative force in secure communication and content generation, enhancing traditional methodologies significantly. In secure communication, AI-driven techniques have improved the robustness and capacity of image steganography. The integration of deep learning models, particularly Convolutional Neural Networks (CNNs), has been key in identifying less detectable regions for message embedding, thereby bolstering steganographic security [2]. This enhancement is crucial for covert communication, especially in environments where information integrity and confidentiality are paramount.

AI's influence on content generation is equally profound, with generative models like GANs revolutionizing digital media creation. These models enable the generation of high-quality, realistic content tailored to specific needs. The conditional GAN framework exemplifies this capability, allowing precise control over content generation, thereby expanding opportunities for creative expression and innovation in digital storytelling [8]. The integration of AI-generated content (AIGC) across various applications highlights its potential to transform narrative forms and enhance user engagement.

Moreover, the deployment of AIGC in mobile edge networks illustrates AI's dual role in content creation and privacy preservation. These applications leverage advanced AI technologies to provide

real-time, personalized services, enhancing user experiences while implementing robust privacy safeguards. This balance between innovation and security in AI, particularly regarding AIGC, presents remarkable opportunities for content creation alongside significant challenges related to security, privacy, and ethical considerations [9, 10]. The implications of AIGC in cyberspace, especially concerning security and privacy, necessitate robust countermeasures to address potential challenges in black and shadow internet industries.

AI plays a multifaceted role in secure communication and content generation, significantly enhancing security measures, operational efficiency, and creative possibilities. The advent of AIGC, powered by advanced algorithms, has accelerated the speed and quality of content creation, allowing rapid production of human-like text, images, and multimedia. However, this evolution raises critical issues regarding security, privacy, ethics, and trustworthiness, necessitating robust frameworks like TrustGAIN to ensure safe and fair AIGC services. As researchers continue to explore AIGC's capabilities and limitations, they are also identifying strategies to address these challenges, guiding the development of more intelligent and secure communication networks [9, 11, 12]. The transformative impact of AI across industries underscores the need for ongoing research and ethical considerations to ensure responsible and secure deployment of AI technologies.

1.3 Innovative Data Manipulation through AI

Artificial Intelligence (AI) has catalyzed innovative data manipulation, introducing methodologies that enhance the efficiency and security of information processing. A notable advancement is the STEGANOGAN method, achieving a payload of 4.4 bits per pixel while remaining undetectable by steganalysis tools, exemplifying an advanced approach to data embedding and manipulation [13]. This method highlights AI's ability to transcend traditional data hiding techniques.

The integration of Large Language Models (LLMs) and Artificial Intelligence Generated Content (AIGC) with communication networks marks a new frontier in innovative data manipulation. This synergy fosters the development of dynamic and adaptive communication systems, enhancing data processing and transmission across digital platforms [14]. Such advancements underscore AI's transformative role in reshaping data handling processes.

Innovative approaches, including Gaussian conditional generative models for image steganography, demonstrate significant improvements over static resource allocation by dynamically allocating resources based on data size and complexity [15]. This dynamic allocation is crucial for optimizing data embedding efficiency, illustrating AI's potential to refine traditional methodologies.

In secure communication, AI-driven techniques, such as specially trained networks that convert images from distinct domains into public and private keys, offer robust security mechanisms [16]. This exemplifies AI's capability to enhance cryptographic techniques through innovative data manipulation strategies.

Generative Adversarial Networks (GANs) also play a pivotal role in innovative data manipulation. For instance, a generator-focused GAN architecture has been developed to preserve essential features of original signatures, emphasizing the retention of critical data characteristics while enabling creative transformations [17]. Additionally, GANs have been applied to automatically generate digital artworks classified as NFT-style, illustrating the intersection of AI and digital art creation [18].

Further evidence of AI's innovative data manipulation capabilities is seen in non-adversarial image synthesis models like GLANN, which effectively sample from a learned latent space without adversarial training, maintaining high-quality image synthesis [19]. Similarly, style transfer techniques have been utilized to disguise embedded information, enhancing the security of steganographic methods by reducing detectability [20].

Current research highlights significant advancements in AIGC technology, enhancing efficiency and enabling new forms of artistic expression [21]. Collectively, these advancements illustrate AI's transformative impact on data manipulation, offering new possibilities for secure communication, content creation, and cross-cultural exchange. As AI continues to evolve, its role in data manipulation is expected to expand, driving further innovations across various domains.

1.4 Structure of the Survey

This survey is meticulously organized to provide a comprehensive analysis of image steganography and generative models in artificial intelligence. The paper is structured into seven main sections, each addressing critical aspects of the topic.

The introductory section establishes a foundation by providing an overview of image steganography and generative models, emphasizing their significance in secure communication and content generation. This section also explores AI's innovative capabilities in data manipulation, offering insights into the transformative impact of these technologies.

The second section delves into background and definitions, offering foundational understanding of core concepts such as image steganography, AIGC, GANs, and generative models. It discusses the historical development and evolution of these technologies, providing context for their current applications.

The third section shifts focus to image steganography, examining both traditional techniques and AI-enhanced methodologies. It underscores the pivotal role of GANs in advancing steganographic practices, highlighting the integration of generative models in this domain.

The fourth section is dedicated to generative models and GANs, presenting an overview of their development and applications. It reviews recent advancements and techniques while addressing challenges encountered in training GANs, providing a balanced perspective on their capabilities and limitations.

The fifth section explores the role of Artificial Intelligence Generated Content (AIGC) in content creation and manipulation, emphasizing its integration with communication networks. This section illustrates AI's dual role in enhancing both creative expression and secure communication.

Practical applications and implications are analyzed in the sixth section, discussing the use of these technologies in secure communication, media forensics, and privacy preservation. It also considers ethical and security implications, offering critical evaluation of potential risks and benefits.

The survey concludes with a summary of key findings, reflecting on future directions and potential advancements in the field. This discussion emphasizes innovative methodologies being developed in image steganography and generative models within AI, particularly highlighting the News Image Steganography (NIS) architecture, which uncovers inconsistencies in news imagery to enhance fake news detection, and the Warfare framework, which addresses vulnerabilities in watermarking AI-generated content. Furthermore, it outlines significant opportunities for future research aimed at improving content attribution and verification, underscoring the continuous advancement and adaptation of these technologies [22, 23].The following sections are organized as shown in Figure 1.

2 Background and Definitions

2.1 Core Concepts and Terminology

Image steganography is a vital technique for secure communication, embedding secret data within carrier images to evade unauthorized detection [1]. This approach maintains the visual integrity of the cover image while concealing the embedded information [4]. Traditional methods like the Least Significant Bit (LSB) technique have been enhanced by deep learning, improving security and payload capacity [5].

Generative Adversarial Networks (GANs), comprising a generator and a discriminator, exemplify deep generative models. The generator creates data that mimics real inputs, while the discriminator evaluates its authenticity [24]. This adversarial setup facilitates high-quality synthetic data production, applicable in forgery detection and realistic image generation [5]. GANs, along with diffusion models, normalizing flows, and variational autoencoders, are crucial for generating realistic samples across diverse data modalities [4].

Artificial Intelligence Generated Content (AIGC) involves content produced via AI technologies, including generative models and large language models (LLMs), significantly enhancing semantic communication and data transmission [21]. AIGC is categorized by intent, modalities, and generation methods, with applications in digital media and copyright protection through watermarking [4].

Steganalysis, the detection of hidden messages in images, is crucial in cybersecurity, focusing on risks from stegomalware—malicious software concealed within multimedia files [1]. The optimal cover image choice is essential for effective steganography, ensuring message concealment without noticeable changes [5].

Deep Generative Models (DGMs) and Large Generative Models (LGMs) are integral to generative AI, impacting fields from industrial time series analysis to privacy-preserving data synthesis [21]. These models address complexities in AI-driven content generation and manipulation, offering innovative solutions to data security and integrity challenges.

The discussed concepts and terminologies form the foundation of advancements in secure communication and content generation, highlighting the dynamic interplay between AI technologies and their applications. The evolution of AIGC technologies enhances data security and integrity, introducing innovative methodologies for secure communication and content generation. Frameworks like TrustGAIN ensure robust and fair AIGC services, while initiatives such as News Image Steganography improve media inconsistency detection, enhancing fake news identification accuracy. As AIGC evolves, it presents opportunities and challenges in enhancing digital content credibility and reliability across domains [22, 10, 9, 11, 12].

2.2 Historical Development of Image Steganography

The evolution of image steganography has progressively enhanced techniques to improve security, capacity, and message imperceptibility within digital images. Early methods like the Least Significant Bit (LSB) technique embedded information in the least significant bits of pixel values, maintaining visual integrity but limited by low payload and detection vulnerability [25].

As digital communication demand grew, traditional methods' limitations prompted more sophisticated approaches, including error-correcting codes like BCH codes, enhancing resilience against errors and distortions [26]. The challenge of classifying cover and stego images persists, complicating secure data transmission [27].

The integration of AI and deep learning in steganography marks a significant advancement, enabling adaptive algorithms that optimize data embedding based on cover image characteristics, enhancing robustness and imperceptibility [28]. Generative models, particularly GANs, revolutionize high-quality cover image synthesis and adversarial example generation, challenging traditional steganalysis [29].

The shift from analytical to generative AI, focusing on content creation, expands steganographic applications [30]. Despite advancements, challenges in standardizing methodologies and knowledge fragmentation hinder innovative techniques' widespread adoption [31]. The proliferation of generative data raises security and privacy concerns, necessitating robust countermeasures against breaches and misinformation [32].

The historical trajectory of image steganography reflects efforts to enhance secure communication by balancing imperceptibility with detection and payload capacity challenges. Integrating generative AI technologies, particularly GANs, significantly bolsters steganography capabilities, enabling advanced data concealment techniques within images. These innovations facilitate covert communication while improving steganographic content's perceptual quality, allowing high-capacity data transmission without compromising carrier images' visual integrity. Recent advancements, like the News Image Steganography architecture, leverage GANs to identify news image inconsistencies, aiding fake news detection and enhancing content protection [22, 23, 33, 13, 28].

2.3 Evolution of Artificial Intelligence Generated Content (AIGC)

The evolution of Artificial Intelligence Generated Content (AIGC) is marked by significant advancements in AI, particularly generative models. Initially grounded in rule-based and statistical models, AIGC development has been enhanced by deep and transfer learning, allowing high-quality, realistic, and contextually relevant content generation across media forms [34]. This evolution is evident in text-to-image generation, where large generative models (LGMs) create diverse, visually compelling images [35].

AIGC's influence spans industries, including digital media, healthcare, and entertainment. In medical imaging, GANs are pivotal in data augmentation, segmentation, and super-resolution, notably in combating COVID-19 [36]. These applications demonstrate AIGC's versatility in enhancing medical data quality and availability, supporting diagnostics and research.

Despite advancements, AIGC faces challenges in training deep generative models to approximate complex, high-dimensional probability distributions from limited samples [37]. High computational demands and slow sampling speeds of generative models, like diffusion models, present obstacles [34]. AIGC's integration into digital risk management systems underscores its dual role in empowering and challenging security frameworks, as it can be exploited for fraudulent activities, necessitating robust countermeasures [10].

AIGC's cultural implications are profound, influencing new paradigms and reshaping societal perceptions. Developing benchmarks for deepfake detection models, differentiating real from GAN and diffusion model-generated images, is crucial for digital media integrity [38]. The slower progress of generative AI models in 3D compared to 2D and text fields highlights the need for continued innovation [6].

AIGC's ongoing evolution transforms digital content creation by introducing innovative possibilities in text, images, audio, and video production. While offering opportunities for creativity and efficiency, it raises challenges related to privacy, bias, misinformation, intellectual property, and ethics. Recent research emphasizes responsible exploration and implementation of AIGC, addressing limitations in complex character animations and emotional expression, still reliant on human creativity. As the field progresses, addressing these challenges through research and ethical considerations is essential to ensure AIGC's responsible and beneficial societal integration [21, 39, 12]. Harnessing AIGC's full potential while mitigating risks is pivotal for its responsible and sustainable societal integration.

In recent years, the field of image steganography has undergone significant transformation, driven by advancements in technology and methodologies. This evolution is particularly evident in the classification of techniques employed within the discipline. Figure 2 illustrates the hierarchical classification of image steganography techniques, highlighting traditional methods alongside AI-enhanced advancements and the role of generative models. The figure organizes key concepts, such as spatial and transform domain methods, while also illustrating the integration of artificial intelligence through Generative Adversarial Networks (GANs) and Convolutional Neural Networks (CNNs). Furthermore, it showcases the innovative use of generative models, emphasizing the evolution and diversification of steganographic practices. By examining these classifications, we can better understand the trajectory of the field and the implications of these advancements for future research.

3 Image Steganography

3.1 Traditional Image Steganography Techniques

Traditional image steganography techniques are divided into spatial and transform domain methods, assessed for their robustness, capacity, and complexity [40]. Spatial domain methods, such as Least Significant Bit (LSB) modification and Pixel Value Differencing (PVD), involve direct pixel value alterations for data embedding, offering simplicity but often facing challenges with low embedding rates and high detectability [41, 42]. LSB substitutes the least significant bits of pixel values with message bits, maintaining visual integrity but remaining vulnerable to predictable pattern detection [41]. Conversely, PVD utilizes pixel value differences for adaptive embedding capacity, though it struggles with robustness against image manipulations [40].

Transform domain methods, such as Discrete Cosine Transform (DCT) and Discrete Wavelet Transform (DWT), embed information within frequency components, offering enhanced robustness against lossy compression and signal processing, suitable for applications prioritizing image quality and security [43]. However, these methods involve higher computational complexity and require sophisticated algorithms for effective message retrieval [43].

The evolution of traditional steganography highlights the ongoing challenge of overcoming limitations posed by lossy channels while innovating to enhance capacity and security. The integration of AI-generated content into steganographic practices introduces new opportunities and challenges, emphasizing the necessity for transparency and accountability in developing robust systems [39, 44].

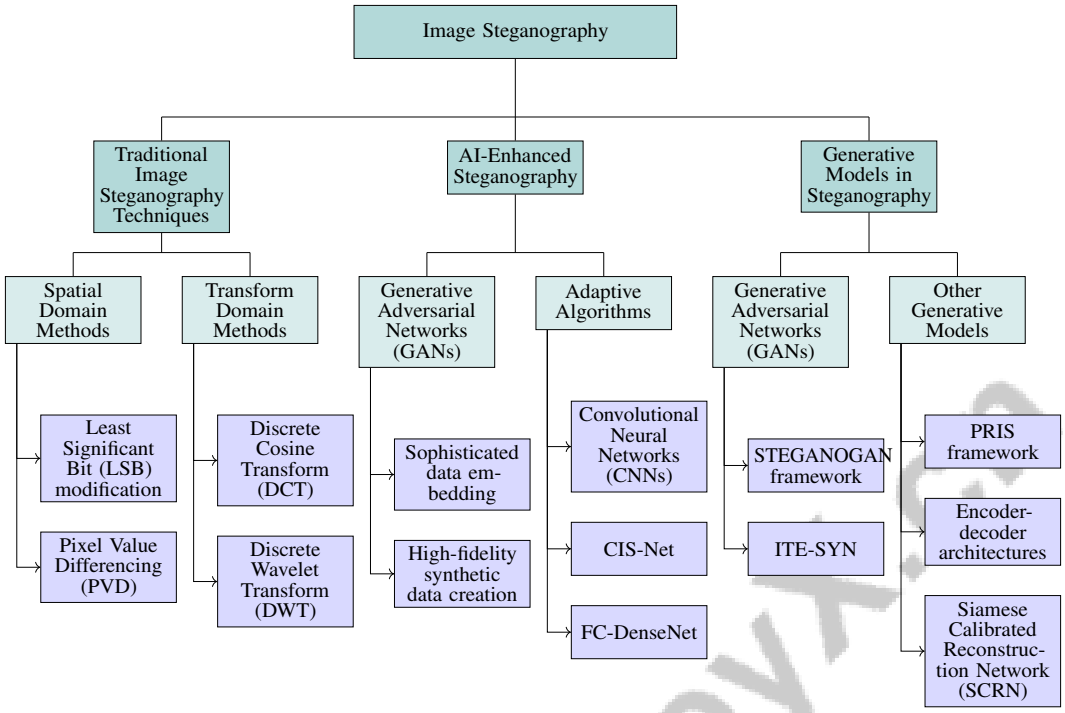


Figure 2: This figure illustrates the hierarchical classification of image steganography techniques, highlighting traditional methods, AI-enhanced advancements, and the role of generative models. It organizes key concepts like spatial and transform domain methods, the integration of AI through GANs and CNNs, and the innovative use of generative models, showcasing the evolution and diversification in steganographic practices.

3.2 AI-Enhanced Steganography

AI-enhanced steganography marks a significant advancement in data hiding, leveraging AI and generative models to improve security, capacity, and imperceptibility of hidden messages. Generative Adversarial Networks (GANs) play a central role, enabling sophisticated data embedding methodologies that enhance robustness against detection while preserving high perceptual quality [4]. GANs facilitate high-fidelity synthetic data creation, crucial for secure embedding and transmission in applications like healthcare [1].

Adaptive algorithms, particularly convolutional neural networks (CNNs), strengthen resistance to steganalysis, despite challenges in ensuring robustness across diverse scenarios [2]. CIS-Net, a CNN designed for image steganalysis, exemplifies this by suppressing cover image content while enhancing message detection [27]. Recent advancements focus on improving the perceptual quality of AI-generated images in steganography. The FC-DenseNet method allows for high-capacity information hiding with minimal visual distortion [1]. Methods like Model Rewriting enable users to specify desired outputs for generative models, optimizing weights while maintaining existing rules [45]. Enhancement modules and gradient approximation functions illustrate the dynamic interplay between AI-driven and traditional methods, significantly improving robustness [2].

Advanced architectures like StyleGAN2, which suppress coarse-scale noise to enhance content control and synthesis quality, highlight AI-enhanced steganography's versatility [7]. Additionally, GANs have been applied in radar signal generation, demonstrating broader applicability beyond traditional image steganography [24].

These advancements underscore AI-enhanced steganography's transformative role in addressing data hiding challenges, driving research and development in this dynamic field, emphasizing further exploration and refinement of AI-driven methodologies [5].

3.3 Generative Models in Steganography

Generative models, especially Generative Adversarial Networks (GANs), have revolutionized image steganography by enhancing capacity, security, and imperceptibility of embedded data. GANs modify cover images for effective secret message embedding, providing robust frameworks against detection [28]. The STEGANOGAN framework exemplifies this, achieving high payload capacities while maintaining image quality through adversarial training [13].

ITE-SYN integrates adversarial perturbations and synchronized modification directions, enhancing resilience against both conventional and CNN-based steganalysis [46]. This demonstrates adversarial strategies' effectiveness in bolstering steganographic security.

Other generative models contribute to the field's development. The PRIS framework employs enhancement modules and a three-step training strategy, showcasing invertible generative models' potential in steganography [25]. Deep learning techniques, particularly CNNs, have automated and enhanced steganography methods, highlighting generative models' role in advancing data hiding techniques [47].

Encoder-decoder architectures in style-based image steganography preprocess secret information and embed it into the latent representation of content images, generating stylized stego images [20]. This method exemplifies style transfer techniques' integration to improve imperceptibility.

Innovative approaches, like using multiple decoders for each secret image, allow retrieving multiple images from a single encoded cover, demonstrating versatility and capacity enhancements achievable through generative models [48]. Specific patterns governed by algorithms like the Canny algorithm further complicate hidden data detection and recovery, illustrating advanced techniques' complexity and effectiveness [49].

Novel methods operating without a discriminator, utilizing guiding components to enhance security, highlight generative models' evolving landscape in steganography [50]. These advancements underscore the dynamic interplay between generative models and traditional techniques, offering innovative solutions to data hiding challenges.

Generative models, particularly GANs, have introduced numerous innovative techniques significantly enhancing steganography's capabilities. By harnessing advanced deep learning frameworks, these methods provide effective solutions to secure communication and content protection challenges. For instance, the Siamese Calibrated Reconstruction Network (SCRN) enhances AI-generated text detection robustness against adversarial attacks, achieving notable accuracy improvements. The Warfare framework addresses vulnerabilities in watermarking techniques for AI-generated content, enabling effective content attribution and verification while mitigating unauthorized commercialization risks. Additionally, the TrustGAIN framework emphasizes robustness, security, and fairness in AI-generated content services, proposing innovative detection methods to ensure network service integrity. Collectively, these advancements pave the way for future innovations, enhancing intelligent communication systems' reliability and safety [51, 23, 11].

4 Generative Models and GANs

4.1 Generative Models and GANs: An Overview

Generative models, especially Generative Adversarial Networks (GANs), have become pivotal in artificial intelligence, enabling the generation of high-quality data across various domains. GANs employ a dual-network architecture comprising a generator and a discriminator, where the generator creates data samples and the discriminator evaluates them to distinguish between real and generated data. This adversarial framework enhances the realism of outputs and fosters innovation in diverse applications [7].

GANs have demonstrated versatility across multiple fields. In digital content creation, modifications to StyleGAN2 allow for precise image feature control via latent variables, enhancing content synthesis [7]. In steganography, techniques like FC-DenseNet Steganography use fully convolutional networks to embed secret images, maintaining visual quality [1]. Generative models like GSK exemplify the potential of GANs to redefine traditional steganography by generating secret messages without

embedding them [4]. Additionally, GANs have been adapted for one-dimensional radar signal data, showcasing their adaptability to new domains [24].

The evolution of generative models is further enriched by frameworks that categorize deep learning development into stages, reflecting the progression from task-specific to general-task applications [6]. This structured approach underscores the expanding capabilities of generative models in addressing complex challenges across various fields.

The ongoing evolution of generative models, with GANs at the forefront, is driving significant advancements in artificial intelligence. These models are expanding applications from secure communication to digital content creation. As research evolves, AI-generated content (AIGC) is poised to revolutionize sectors such as content creation and digital risk management, highlighting the profound implications of AIGC technologies like ChatGPT and Stable Diffusion. This transformation necessitates a reevaluation of risk management strategies to address the complexities introduced by these powerful tools [10, 12].

4.2 Recent Advancements and Techniques

Recent advancements in generative models, particularly GANs, have enhanced their capacity to produce diverse and high-quality outputs. The integration of adversarial training with variational autoencoders, such as in -GAN, leverages the strengths of both to create stable and effective models [52]. This is crucial for applications demanding high fidelity and diversity in generated data.

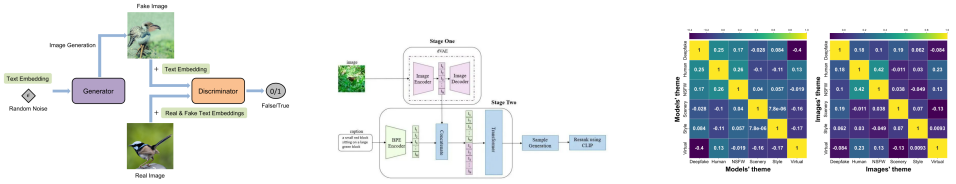
Innovations like CycleGANs have improved unsupervised image-to-image translation, enhancing GAN robustness in scenarios with scarce labeled data [53]. The 'many-discriminator, one-generator' architecture addresses mode collapse, enhancing output diversity [54]. COEGAN employs neuroevolution to optimize GAN architectures, improving training stability [55].

Dynamic memory mechanisms in models like DM-GAN refine image generation, producing high-quality images by iteratively improving initial outputs [56]. In text-to-image synthesis, StyleGAN advancements generate high-quality, identity-preserving images, crucial for digital media and entertainment sectors [57, 58].

In anomaly detection, methods like Discriminator Driven Latent Sampling (DDLS) enhance sampling efficiency from the GAN latent space, correcting generator biases [59]. Recent research has also advanced generative diffusion models, demonstrating effectiveness in image synthesis and natural language processing [60].

Generative convolution layers, assessed by metrics like Inception Score (IS) and Fréchet Inception Distance (FID), show significant image quality improvements [61]. Experiments with encoding multiple secret images within a single cover image highlight increased steganographic capacity [48].

These advancements in generative models, particularly GANs, facilitate applications in secure communication, digital content creation, and media manipulation detection, addressing misinformation concerns in social media. Ongoing research expands GAN efficacy and scope, with new evaluation metrics assessing model performance, underscoring the field's dynamic nature and potential to reshape digital interactions and security measures [62, 63, 64, 65, 28].



(a) Generative Adversarial Network (GAN) for Image-to-Text Translation[66]

(b) Stage One and Stage Two of a Machine Learning Model for Image Captioning and Reranking[67]

(c) Comparison of Image Themes Across Models and Subjects[68]

Figure 3: Examples of Recent Advancements and Techniques

AI-generated, for reference only.

As shown in Figure 3, the exploration of generative models, particularly GANs, has led to significant advancements in machine learning, as illustrated by recent techniques in image-to-text translation, image captioning, and thematic analysis of generated images. Figure 3 provides a visual overview of these advancements. The first subfigure showcases a GAN designed for image-to-text translation, highlighting its core components: the generator, discriminator, and text embedding. This setup enables the generation of images from random noise and text inputs, with the discriminator distinguishing between real and fake text embeddings. The second subfigure presents a two-stage machine learning model for image captioning and reranking, where an image is first encoded into a latent space using a Deep Variational Autoencoder (dVAE), which subsequently facilitates the generation of descriptive captions. Lastly, the third subfigure offers a comparative analysis of image themes across various models and subjects through a heat map, illustrating the thematic alignment or divergence, with values ranging from -1 to 1 to indicate the degree of similarity. Together, these examples underscore the dynamic progress in generative modeling techniques and their applications in diverse domains [66, 67, 68].

4.3 Challenges in Training GANs

Training Generative Adversarial Networks (GANs) presents inherent challenges due to their adversarial architecture, comprising a generator and a discriminator. Achieving stable training is a primary difficulty, as the adversarial nature can lead to oscillations and convergence failures [69]. This instability is often exacerbated by mode collapse, where the generator produces limited and similar outputs.

The issue of mode collapse is compounded by challenges in hyperparameter tuning, critical for effective GAN training [37]. The assumption of an optimal adversary before each generator update is often impractical, leading to ineffective learning signals [70]. Robust training methodologies are needed to accommodate the dynamic nature of adversarial learning.

Moreover, the computational resources required for centralized GAN training present challenges, including privacy risks associated with centralized data processing [71]. The complexity of data distributions, particularly with complex or poorly defined modes, further complicates GAN training, as seen in methods like MIC-GANs [72].

GAN effectiveness is also limited by the difficulty in generating images that accurately reflect complex textual descriptions or adhere to specific styles, a challenge persisting despite advancements in image synthesis techniques [64]. Existing metrics like Inception Score (IS) and Fréchet Inception Distance (FID) often fall short in evaluating GAN performance, necessitating novel approaches like the Likeness Score (LS) for more direct image analysis [73].

Adversarial attacks add complexity, as demonstrated in benchmark tests of models like Deep Convolutional GAN (DCGAN) and Auxiliary Classifier GAN (ACGAN), which evaluate generative model robustness against such threats [29]. Methods like Discriminator Driven Latent Sampling (DDLS) offer improvements in sample quality, addressing issues like mode dropping [59].

Overall, GAN training challenges are multifaceted, involving stability, mode collapse, computational demands, and evaluation metrics. Continuous research and innovation are essential to improve GAN robustness and applicability across various fields, including natural language processing, architectural design, and 3D content generation. This effort will enhance GAN performance and address key limitations, ensuring effectiveness in critical applications like deep fake detection and visual credibility assessment in social media [74, 70, 63, 12, 75].

5 Artificial Intelligence Generated Content (AIGC)

The rise of Artificial Intelligence Generated Content (AIGC) has profoundly influenced various sectors by transforming content creation, particularly in digital media and healthcare. AIGC enhances creative workflows and user engagement while addressing deployment challenges. By examining its development and integration into digital storytelling, we can understand its capabilities and limitations, stressing the need for responsible use to mitigate risks such as privacy concerns, bias, and intellectual property issues [21, 39, 12].

5.1 Role of AIGC in Content Creation

AIGC leverages advanced generative models, such as Generative Adversarial Networks (GANs), to produce high-quality outputs across various domains. In digital media, modified GANs like StyleGAN2 synthesize complex visual content with enhanced control over identity features, providing a robust foundation for artist-friendly digital creation tools [7]. This advancement fosters innovative artistic expressions and expands digital storytelling's scope [21].

In healthcare, AIGC generates synthetic datasets, such as Electronic Health Records (EHR), addressing data scarcity and improving diagnostic processes. By producing high-quality data, AIGC enhances diagnostic accuracy and research capabilities through advanced algorithms and extensive datasets [57, 76, 12, 21, 77]. This augmentation is crucial for advancing healthcare technologies and ensuring reliable diagnostic outcomes.

AIGC democratizes content creation in creative industries, enabling novice users to engage through intuitive interfaces and targeted model edits. This accessibility fosters personalized and diverse outputs, catering to the demand for customized digital experiences. By integrating AIGC into creative workflows, productivity is enhanced, allowing creators unprecedented control over content generation. While AIGC streamlines digital storytelling across various media, it raises ethical concerns regarding privacy, bias, and misinformation. Effective AIGC utilization requires a balanced approach to maximize benefits while addressing limitations [21, 39, 78, 12].

AIGC faces challenges such as vulnerability to adversarial attacks, which can create indistinguishable out-domain examples. This underscores the need for robust defense mechanisms and evaluation frameworks prioritizing AIGC outputs' security, fairness, and integrity, especially regarding adversarial threats, privacy breaches, and media forgery risks [32, 11]. As AIGC technologies advance, they promise to reshape content creation while addressing emerging challenges in digital media, healthcare, and beyond.

5.2 Integration with Communication Networks

Integrating AIGC with communication networks significantly alters how digital content is generated, transmitted, and consumed. Efficient wireless and semantic communication (SemCom) technologies are crucial for implementing AIGC services, ensuring seamless integration into existing infrastructures [79].

A primary challenge in this integration is the need for large-scale and diverse datasets, particularly in 3D content generation, where existing architectures often prove inefficient and costly [74]. The high training costs of generative models further complicate scalability, necessitating efficient solutions to support AIGC within communication networks [12].

Data privacy and security are critical concerns in AIGC integration with communication networks. The evolving roles of educators and algorithm transparency significantly influence AI-generated content's trustworthiness in educational and professional contexts [76]. Advanced cryptographic protocols, such as those proposed by TrustGAIN, are essential for protecting user data and ensuring the integrity of transmitted content [11].

The Artifact Purification Network (APN) offers a promising method for detecting AI-generated content, enhancing security and preventing harmful or misleading information [80]. Future research may explore integrating additional noise variables into generative models to improve robustness across various domains [59].

AIGC's potential to transform communication networks is significant, paving the way for more personalized and dynamic content delivery. Innovative methods that separately encode visual and textual features, using prompts to influence quality assessments of AIGC content, exemplify advancements enhancing AI-generated outputs' relevance and quality [57].

While integrating AIGC with communication networks offers substantial opportunities for enhancing content creation and cross-cultural communication, it presents challenges related to security, privacy, and ethical considerations. Ongoing research and innovation are vital to address these issues, develop effective regulatory frameworks, and ensure secure AIGC technologies implementation, maximizing benefits while mitigating societal risks [9, 14, 78, 81]. As the field evolves, developing efficient,

AI-generated, for reference only.

secure, and trustworthy communication frameworks will be essential to fully harness AIGC's potential in the digital age.

6 Applications and Implications

6.1 Applications in Image Generation and Data Hiding

The fusion of generative models with AI technologies has broadened the scope of image generation and data hiding, providing innovative solutions across various fields. Techniques like GINR-Stega demonstrate adaptability across media types while maintaining image quality during information embedding, crucial for secure communication and content protection [33]. In image generation, CycleGAN exemplifies its dual utility in both image synthesis and data hiding, effectively translating images between domains without paired examples, showcasing its versatility in scenarios with limited labeled data [53]. Recent advancements in multi-image steganography, allowing for the embedding of multiple images within a single cover image, have significantly improved embedding capacity without degrading quality, essential for high data throughput applications [48].

GAN-based methods generally outperform traditional techniques in effectiveness, security, and embedding capacity, leveraging deep learning to enhance the fidelity and security of steganographic processes, as evidenced by superior PSNR and SSIM metrics [28]. Techniques like VEM, which offer flexible variational distributions to account for pixel dependencies, yield high-quality images even at high compression rates, further emphasizing their potential [82]. Diffusion models also play a transformative role in generating high-quality samples across numerous applications, enhancing the fidelity and diversity of AI-generated content [60]. Innovative applications extend to digital art, where approaches for NFT art significantly reduce the time and effort required for artists, enabling enhancements to GAN-generated artworks [18]. The MPG framework exemplifies the application of generative models in producing high-quality, diverse images that accurately reflect specified attributes [8]. Additionally, the PRIS method enhances robustness against distortions, improving the reliability of data hiding techniques [25].

The integration of generative models and AI in image generation and data hiding is advancing innovation, as seen in sophisticated techniques such as text-to-image generation and 3D content creation. These developments not only improve image fidelity from textual descriptions but also address complex challenges in fields such as virtual simulations and automatic 3D content generation. As large models evolve, their applications are expected to extend into more intricate tasks, offering transformative solutions that could redefine creative processes and productivity in the era of AI-generated content (AIGC) [74, 66]. As research progresses, these applications will expand, providing innovative solutions while ensuring robust security and high-quality outputs.

6.2 Security and Privacy Implications

AI integration into data hiding and content generation presents complex security and privacy implications. Generative models like GANs enhance the robustness and imperceptibility of hidden messages but pose challenges in maintaining data integrity and preventing misuse. A significant concern is the risk of generative models replicating sensitive training data, leading to privacy breaches, especially in healthcare, where synthetic Electronic Health Records (EHR) data must balance utility with privacy protection [2]. Adaptive methods in steganography, such as those discussed by Duan et al., enhance security by dynamically adjusting embedding strategies to minimize detectability and maximize data protection. Techniques like StegColNet improve detection accuracy by capturing diverse features across multiple color spaces, strengthening steganographic systems. Moreover, methods that disentangle private and non-private information in the latent space allow for natural image reconstruction while safeguarding sensitive data [4].

Despite advancements, the potential misuse of generative data for malicious purposes remains a significant concern, necessitating the evolution of Digital Rights Management (DRM) systems to address unique challenges posed by AIGC technologies. Models like CIS-Net can suppress cover image content without losing relevant information, emphasizing the need for robust classification accuracy to prevent unauthorized access. Additionally, image steganography based on style transfer offers high message recovery accuracy and enhances the artistic quality of stego images, improving resistance to detection by steganalysis networks [7].

The ethical and security issues inherent in current AIGC models, particularly in light of evolving regulatory requirements, underscore the need for compliance and vigilance to ensure responsible technology deployment. Balancing the benefits of AI advancements with robust security measures is essential, emphasizing adaptable frameworks to protect against potential threats [2]. The security and privacy implications of AI in data hiding and content generation are multifaceted, necessitating ongoing innovation and proactive measures to tackle emerging challenges such as watermark vulnerability, privacy leaks, and misinformation risks. As AIGC becomes prevalent, issues like unauthorized commercialization and media forgery highlight the urgent need for effective regulatory frameworks and advanced watermarking techniques to ensure content attribution and verification. Continuous assessment of security and privacy threats, along with the development of robust countermeasures, is crucial to maintaining trust in generative technologies [9, 32, 23, 83]. Responsible deployment necessitates careful balancing of AI capabilities with comprehensive security measures.

6.3 Media Forensics

AI's role in media forensics has become increasingly critical as AI-generated content proliferates. Techniques such as the Artifact Purification Network (APN) effectively detect AI-generated images, playing a crucial role in distinguishing authentic from synthetic content [80]. This capability is vital for maintaining digital media integrity, particularly as deepfakes and other synthetic media become more sophisticated. Developing benchmarks for deepfake detection, as highlighted by Guarnera et al., provides a robust framework for forensic analysis, enabling the reconstruction of multimedia data history [38]. These benchmarks are integral to forensic processes, offering systematic approaches to evaluate and enhance detection methodologies. By standardizing performance assessment for deepfake detection models, benchmarks facilitate identification of weaknesses and areas for improvement, strengthening the media forensic landscape.

Integrating AI technologies into media forensics not only improves manipulated content detection but also enhances the ability to trace origins and modifications of digital media. This capability is essential for legal and regulatory purposes, ensuring reliable use of digital evidence in judicial contexts. As AI-generated content becomes more pervasive, the role of AI in media forensics will expand, offering innovative solutions to challenges posed by synthetic media. The rapid advancement of AIGC technologies is transforming various sectors, including the digital shadow industry, creating unique challenges for forensic analyses. This ongoing evolution necessitates sustained research and development efforts to enhance forensic methodologies' reliability and accuracy, effectively addressing complexities introduced by AIGC that can generate realistic, personalized content potentially exploited for fraudulent activities. Adapting forensic practices to counter emerging threats is imperative to maintain integrity in the digital age [10, 39, 12].

6.4 Privacy Preservation

AI technologies in data hiding and content generation significantly enhance privacy preservation by securing sensitive information. AI-driven methods, particularly those utilizing GANs, create synthetic datasets that mimic real-world data without exposing actual sensitive information, crucial in sectors like healthcare where synthetic EHR data can be analyzed and shared without compromising patient privacy [2]. In image steganography, AI enables sophisticated embedding strategies that minimize hidden message detectability, preserving transmitted data confidentiality. Techniques like StegColNet enhance detection accuracy by leveraging features from multiple color spaces, showcasing AI's potential to improve steganographic system security while maintaining embedded information privacy [4].

Moreover, AI's capability to disentangle private and non-private information within generative models' latent space allows natural image reconstruction while protecting sensitive data. This separation is vital for maintaining privacy while generating high-quality synthetic content for various applications without risking unauthorized exposure of confidential data [7]. Despite advancements, the potential misuse of AI-generated data for malicious purposes poses significant privacy risks, necessitating robust DRM systems to address challenges presented by AIGC. Ensuring the ethical and secure deployment of AI technologies in privacy preservation involves balancing AI advancements' benefits with comprehensive security measures against potential threats [2].

AI's role in privacy preservation is complex, providing innovative solutions that enhance data security and protect sensitive information confidentiality. This includes advanced techniques like blockchain and privacy computing, which can integrate with AIGC tools to address security and privacy challenges. Ongoing research highlights countermeasures for risks associated with generative data, such as privacy leaks and media forgery, emphasizing trustworthiness in AIGC. AI technologies are reshaping content creation while prompting critical discussions on ethical, legal, and security implications, fostering a more secure digital environment [9, 10, 32, 81]. Continued research and development are essential to address emerging challenges and ensure responsible AI use in privacy-preserving applications.

6.5 Ethical and Security Implications

The integration of AI in generative content and data hiding technologies presents significant ethical and security challenges necessitating a comprehensive approach. A primary ethical concern involves biases in AI-generated content due to imbalanced training datasets, particularly problematic in creative domains where a lack of diversity can exacerbate existing inequalities. Addressing these biases requires frameworks incorporating privacy computing and blockchain technologies [32]. Ethical considerations extend to GAN use in cybersecurity, where reliance on fixed neural networks for securing AI systems underscores the need for continuous evaluation and enhancement against adversarial threats [10].

In content generation, potential critical detail loss during AI processes raises safety concerns, especially in applications like autonomous driving [8]. Challenges of computational efficiency and extensive training data for text-to-image generation highlight ethical and security implications, as resource-intensive processes may lead to environmental and economic concerns [35]. Ethical implications of synthetic data use in AI applications, particularly regarding hallucination phenomena in large vision-language models (LVLMs), underscore complexities in ensuring data integrity and reliability [84].

The proliferation of deepfakes and misinformation complicates the ethical landscape, as these technologies can create deceptive media content. The rapid advancement of deepfake technologies often outpaces detection methods, and insufficient regulatory frameworks exacerbate ethical concerns. Current detection tools for AI-generated images lack reliability for practical use, necessitating significant improvements in robustness and generalization [38]. Ethical implications of using AI in adversarial training are also considered, particularly the balance between improving detection and risks associated with generating adversarial samples [10].

Regulatory frameworks play a crucial role in addressing these challenges. The regulatory landscape surrounding AIGC watermarking emphasizes the need for clear guidelines and standards across jurisdictions. Future research should focus on developing robust detection technologies, exploring multimodal detection, and addressing ethical implications of AI-generated content [32]. Automating cost function design through deep learning is also discussed as an ethical and security implication of advanced AI techniques in steganography [45].

Addressing the ethical and security implications of AI in generative content and data hiding requires a multifaceted approach, including improving data diversity, enhancing model robustness, and developing comprehensive evaluation frameworks to ensure responsible AI technology deployment [32].

7 Conclusion

7.1 Innovative Approaches and Future Directions

The trajectory of image steganography and generative models is poised for significant advancements, driven by innovative approaches that emphasize robustness, applicability, and ethical dimensions. In the realm of image steganography, enhancing resilience against evolving forensic detection methods is paramount. Future research should prioritize the development of intricate embedding patterns and effective compensation techniques to advance steganographic practices. Additionally, optimizing feature selection and exploring alternative color spaces or hybrid models are promising strategies to enhance detection capabilities.

For generative models, key areas of focus include stabilizing GAN training and exploring multi-object generation. The introduction of novel architectures, such as Capsule Networks, holds potential for improving image content understanding and facilitating more stable GAN training. Furthermore, expanding fingerprinting schemes to include other generative models and strengthening defenses against model extraction attacks are crucial research avenues that reflect the innovative trends in the field.

The integration of generative models with advanced techniques like federated learning offers a promising research direction, optimizing resource allocation and addressing privacy concerns to ensure efficiency and security. Additionally, exploring new artistic workflows and establishing ethical standards for Artificial Intelligence Generated Content (AIGC) are vital for the responsible and sustainable deployment of these technologies.

Developing adaptive Digital Rights Management (DRM) systems that leverage real-time data analysis and machine learning is essential for mitigating AIGC-driven fraud, thereby enhancing the security and applicability of generative models across various domains. Moreover, improving Generative Steganography with Kerckhoffs' principle by applying it to more complex datasets and optimizing architectures to minimize key size represents a critical direction for future research.

The future of image steganography and generative models is abundant with opportunities for innovation. By addressing current limitations and exploring novel methodologies, these fields are positioned to evolve, offering transformative solutions to complex challenges across diverse industries.

References

[1] Duan Xintao and Liu Nao. Hide the image in fc-densenets to another image, 2019.

[2] Mehdi Sharifzadeh, Chirag Agarwal, Mohammed Aloraini, and Dan Schonfeld. Convolutional neural network steganalysis's application to steganography, 2017.

[3] Tianxiao Han, Jiancheng Tang, Qianqian Yang, Yiping Duan, Zhaoyang Zhang, and Zhiguo Shi. Generative model based highly efficient semantic communication approach for image transmission, 2022.

[4] Yan Ke, Minqing Zhang, Jia Liu, Tingting Su, and Xiaoyuan Yang. Generative steganography with kerckhoffs' principle, 2021.

[5] Guanlin Li, Guowen Xu, Han Qiu, Shangwei Guo, Run Wang, Jiwei Li, Tianwei Zhang, and Rongxing Lu. Fingerprinting image-to-image generative adversarial networks, 2024.

[6] Chenghao Li and Chaoning Zhang. When chatgpt for computer vision will come? from 2d to 3d, 2023.

[7] Vaibhav Vavilala and David Forsyth. Controlled gan-based creature synthesis via a challenging game art dataset – addressing the noise-latent trade-off, 2021.

[8] Fangda Han, Guoyao Hao, Ricardo Guerrero, and Vladimir Pavlovic. Mpg: A multi-ingredient pizza image generator with conditional stylegans, 2021.

[9] Yuntao Wang, Yanghe Pan, Miao Yan, Zhou Su, and Tom H. Luan. A survey on chatgpt: Ai-generated contents, challenges, and solutions, 2023.

[10] Qichao Wang, Huan Ma, Wentao Wei, Hangyu Li, Liang Chen, Peilin Zhao, Binwen Zhao, Bo Hu, Shu Zhang, Zibin Zheng, and Bingzhe Wu. Attention paper: How generative ai reshapes digital shadow industry?, 2023.

[11] Siyuan Li, Xi Lin, Yaju Liu, Xiang Chen, and Jianhua Li. Trustworthy ai-generative content for intelligent network service: Robustness, security, and fairness, 2025.

[12] Chengzhang Zhu, Luobin Cui, Ying Tang, and Jiacun Wang. The evolution and future perspectives of artificial intelligence generated content, 2024.

[13] Kevin Alex Zhang, Alfredo Cuesta-Infante, Lei Xu, and Kalyan Veeramachaneni. Steganogan: High capacity image steganography with gans, 2019.

[14] Jie Guo, Meiting Wang, Hang Yin, Bin Song, Yuhao Chi, Fei Richard Yu, and Chau Yuen. Large language models and artificial intelligence generated content technologies meet communication networks, 2024.

[15] Wenkang Su, Jiangqun Ni, Yuanfeng Pan, Xianglei Hu, and Yun-Qing Shi. Image steganography using gaussian markov random field model, 2019.

[16] Ziqiang Zheng, Hongzhi Liu, Zhibin Yu, Haiyong Zheng, Yang Wu, Yang Yang, and Jianbo Shi. Encryptgan: Image steganography with domain transform, 2019.

[17] Haadia Amjad, Kilian Goeller, Steffen Seitz, Carsten Knoll, Naseer Bajwa, Ronald Tetzlaff, and Muhammad Imran Malik. Block induced signature generative adversarial network (bisgan): Signature spoofing using gans and their evaluation, 2024.

[18] Sakib Shahriar and Kadhim Hayawi. Nftgan: Non-fungible token art generation using generative adversarial networks, 2021.

[19] Yedid Hoshen and Jitendra Malik. Non-adversarial image synthesis with generative latent nearest neighbors, 2018.

[20] Donghui Hu, Yu Zhang, Cong Yu, Jian Wang, and Yaofei Wang. Image steganography based on style transfer, 2022.

[21] Rongzhang Gu, Hui Li, Changyue Su, and Wayne Wu. Innovative digital storytelling with aigc: Exploration and discussion of recent advances, 2023.

[22] Jizhe Zhou, Chi-Man Pun, and Yu Tong. News image steganography: A novel architecture facilitates the fake news identification, 2021.

[23] Guanlin Li, Yifei Chen, Jie Zhang, Shangwei Guo, Han Qiu, Guoyin Wang, Jiwei Li, and Tianwei Zhang. Warfare:breaking the watermark protection of ai-generated content, 2025.

[24] Thomas Truong and Svetlana Yanushkevich. Generative adversarial network for radar signal generation, 2020.

[25] Hang Yang, Yitian Xu, Xuhua Liu, and Xiaodong Ma. Pris: Practical robust invertible network for image steganography, 2023.

[26] Debajit Sensarma and Samar Sen Sarma. Data hiding using graphical code based steganography technique, 2015.

[27] Songtao Wu, Sheng hua Zhong, Yan Liu, and Mengyuan Liu. Cis-net: A novel cnn model for spatial image steganalysis via cover image suppression, 2019.

[28] Jia Liu, Yan Ke, Yu Lei, Zhuo Zhang, Jun Li, Peng Luo, Minqing Zhang, and Xiaoyuan Yang. Recent advances of image steganography with generative adversarial networks, 2019.

[29] Dario Pasquini, Marco Mingione, and Massimo Bernaschi. Adversarial out-domain examples for generative models, 2019.

[30] Chaoning Zhang, Chenshuang Zhang, Sheng Zheng, Yu Qiao, Chenghao Li, Mengchun Zhang, Sumit Kumar Dam, Chu Myaet Thwal, Ye Lin Tun, Le Luang Huy, Donguk kim, Sung-Ho Bae, Lik-Hang Lee, Yang Yang, Heng Tao Shen, In So Kweon, and Choong Seon Hong. A complete survey on generative ai (aigc): Is chatgpt from gpt-4 to gpt-5 all you need?, 2023.

[31] H. B. Bahar and Ali Aboutalebi. Image steganography, a new approach for transferring security information, 2008.

[32] Tao Wang, Yushu Zhang, Shuren Qi, Ruoyu Zhao, Zhihua Xia, and Jian Weng. Security and privacy on generative data in aigc: A survey, 2024.

[33] Zhong Yangjie, Liu Jia, Ke Yan, and Liu Meiqi. Image steganography based on generative implicit neural representation, 2024.

[34] Chaoning Zhang, Chenshuang Zhang, Sheng Zheng, Yu Qiao, Chenghao Li, Mengchun Zhang, Sumit Kumar Dam, Chu Myaet Thwal, Ye Lin Tun, Le Luang Huy, et al. A complete survey on generative ai (aigc): Is chatgpt from gpt-4 to gpt-5 all you need? *arXiv preprint arXiv:2303.11717*, 2023.

[35] Minrui Xu, Hongyang Du, Dusit Niyato, Jiawen Kang, Zehui Xiong, Shiwen Mao, Zhu Han, Abbas Jamalipour, Dong In Kim, Xuemin Shen, Victor C. M. Leung, and H. Vincent Poor. Unleashing the power of edge-cloud generative ai in mobile networks: A survey of aigc services, 2023.

[36] Hazrat Ali and Zubair Shah. Combating covid-19 using generative adversarial networks and artificial intelligence for medical images: A scoping review, 2022.

[37] Lars Ruthotto and Eldad Haber. An introduction to deep generative modeling. *GAMM-Mitteilungen*, 44(2):e202100008, 2021.

[38] Luca Guarnera, Oliver Giudice, and Sebastiano Battiato. Level up the deepfake detection: a method to effectively discriminate images generated by gan architectures and diffusion models, 2023.

[39] Chen Chen, Jie Fu, and Lingjuan Lyu. A pathway towards responsible ai generated content, 2023.

[40] Ramita Maharjan, Ajay Kumar Shrestha, and Rejina Basnet. Image steganography: Protection of digital properties against eavesdropping, 2019.

[41] Idakwo M. A., Muazu M. B., Adedokun A. E., and Sadiq B. O. An extensive survey of digital image steganography: State of the art, 2024.

[42] Soumendu Chakraborty, Anand Singh Jalal, and Charul Bhatnagar. Lsb based non blind predictive edge adaptive image steganography, 2022.

[43] Xiaolong Duan, Bin Li, Zhaoxia Yin, Xinpeng Zhang, and Bin Luo. Robust image steganography against lossy jpeg compression based on embedding domain selection and adaptive error correction, 2023.

[44] Yiluo Wei and Gareth Tyson. Understanding the impact of ai generated content on social media: The pixiv case, 2024.

[45] David Bau, Steven Liu, Tongzhou Wang, Jun-Yan Zhu, and Antonio Torralba. Rewriting a deep generative model. In *Computer Vision–ECCV 2020: 16th European Conference, Glasgow, UK, August 23–28, 2020, Proceedings, Part I 16*, pages 351–369. Springer, 2020.

[46] Xinghong Qin, Shunquan Tan, Bin Li, Weixuan Tang, and Jiwu Huang. Image steganography based on iteratively adversarial samples of a synchronized-directions sub-image, 2021.

[47] High-capacity image steganograph.

[48] Abhishek Das, Japsimar Singh Wahi, Mansi Anand, and Yugant Rana. Multi-image steganography using deep neural networks, 2021.

[49] Youssef Bassil. Image steganography based on a parameterized canny edge detection algorithm, 2012.

[50] Miaoxin Ye, Dongxia Huang, Kangkang Wei, and Weiqi Luo. A novel residual-guided learning method for image steganography, 2023.

[51] Guanhua Huang, Yuchen Zhang, Zhe Li, Yongjian You, Mingze Wang, and Zhouwang Yang. Are ai-generated text detectors robust to adversarial perturbations?, 2024.

[52] Mihaela Rosca, Balaji Lakshminarayanan, David Warde-Farley, and Shakir Mohamed. Variational approaches for auto-encoding generative adversarial networks. *arXiv preprint arXiv:1706.04987*, 2017.

[53] Nibraas Khan, Ruj Haan, George Boktor, Michael McComas, and Ramin Daneshi. Steganography gan: Cracking steganography with cycle generative adversarial networks, 2020.

[54] Anant Shukla, Martin Jurecek, and Mark Stamp. Social media bot detection using dropout-gan, 2023.

[55] Victor Costa, Nuno Lourenço, João Correia, and Penousal Machado. Coegan: Evaluating the coevolution effect in generative adversarial networks, 2019.

[56] Minfeng Zhu, Pingbo Pan, Wei Chen, and Yi Yang. Dm-gan: Dynamic memory generative adversarial networks for text-to-image synthesis, 2019.

[57] Xi Fang, Weigang Wang, Xiaoxin Lv, and Jun Yan. Pcqa: A strong baseline for aigc quality assessment based on prompt condition, 2024.

[58] Jiquan Yuan, Xinyan Cao, Jinming Che, Qinyuan Wang, Sen Liang, Wei Ren, Jinlong Lin, and Xixin Cao. Tier: Text-image encoder-based regression for aigc image quality assessment, 2024.

[59] Tong Che, Ruixiang Zhang, Jascha Sohl-Dickstein, Hugo Larochelle, Liam Paull, Yuan Cao, and Yoshua Bengio. Your gan is secretly an energy-based model and you should use discriminator driven latent sampling, 2021.

[60] Hanqun Cao, Cheng Tan, Zhangyang Gao, Yilun Xu, Guangyong Chen, Pheng-Ann Heng, and Stan Z. Li. A survey on generative diffusion model, 2023.

[61] Seung Park and Yong-Goo Shin. Generative convolution layer for image generation, 2021.

[62] Diego Gragnaniello, Davide Cozzolino, Francesco Marra, Giovanni Poggi, and Luisa Verdoliva. Are gan generated images easy to detect? a critical analysis of the state-of-the-art, 2021.

[63] Preeti Sharma, Manoj Kumar, Hitesh Kumar Sharma, and Soly Mathew Biju. Generative adversarial networks (gans): introduction, taxonomy, variants, limitations, and applications. *Multimedia Tools and Applications*, pages 1–48, 2024.

[64] He Huang, Philip S. Yu, and Changhu Wang. An introduction to image synthesis with generative adversarial nets, 2018.

[65] Hrishikesh Sharma. A chronological survey of theoretical advancements in generative adversarial networks for computer vision, 2023.

[66] Fengxiang Bie, Yibo Yang, Zhongzhu Zhou, Adam Ghanem, Minjia Zhang, Zhewei Yao, Xiaoxia Wu, Connor Holmes, Pareesa Golnari, David A. Clifton, Yuxiong He, Dacheng Tao, and Shuaiwen Leon Song. Renaissance: A survey into ai text-to-image generation in the era of large model, 2023.

[67] Nonghai Zhang and Hao Tang. Text-to-image synthesis: A decade survey, 2024.

[68] Yiluo Wei, Yiming Zhu, Pan Hui, and Gareth Tyson. Exploring the use of abusive generative ai models on civitai, 2024.

[69] Abdul Jabbar, Xi Li, and Bourahla Omar. A survey on generative adversarial networks: Variants, applications, and training. *ACM Computing Surveys (CSUR)*, 54(8):1–49, 2021.

[70] Andreas Munk, William Harvey, and Frank Wood. Assisting the adversary to improve gan training, 2020.

[71] Xumin Huang, Peichun Li, Hongyang Du, Jiawen Kang, Dusit Niyato, Dong In Kim, and Yuan Wu. Federated learning-empowered ai-generated content in wireless networks, 2023.

[72] Hui Ying, He Wang, Tianjia Shao, Yin Yang, and Kun Zhou. Unsupervised image generation with infinite generative adversarial networks, 2021.

[73] Shuyue Guan and Murray Loew. A novel measure to evaluate generative adversarial networks based on direct analysis of generated images, 2021.

[74] Ke Zhao and Andreas Larsen. Challenges and opportunities in 3d content generation, 2024.

[75] Victor Costa, Nuno Lourenço, João Correia, and Penousal Machado. Exploring the evolution of gans through quality diversity, 2020.

[76] Shi Yang, Siyuan Yang, Chaoran Tong, et al. In-depth application of artificial intelligence-generated content aigc large model in higher education. *Adult and Higher Education*, 5(19):9–16, 2023.

[77] Bowen Qu, Xiaoyu Liang, Shangkun Sun, and Wei Gao. Exploring aigc video quality: A focus on visual harmony, video-text consistency and domain distribution gap, 2024.

[78] HU Zi-yang. Aigc related context: A new communication culture for human. *Journal of Literature and Art Studies*, 14(10):921–931, 2024.

[79] Guangyuan Liu, Hongyang Du, Dusit Niyato, Jiawen Kang, Zehui Xiong, Dong In Kim, Xuemin, and Shen. Semantic communications for artificial intelligence generated content (aigc) toward effective content creation, 2024.

[80] Zheling Meng, Bo Peng, Jing Dong, and Tieniu Tan. Artifact feature purification for cross-domain detection of ai-generated images, 2024.

[81] Chuan Chen, Zhenpeng Wu, Yanyi Lai, Wenlin Ou, Tianchi Liao, and Zibin Zheng. Challenges and remedies to privacy and security in aigc: Exploring the potential of privacy computing, blockchain, and beyond, 2023.

[82] Minsoo Kang, Hyewon Yoo, Eunhee Kang, Sehwan Ki, Hyong-Euk Lee, and Bohyung Han. Information-theoretic gan compression with variational energy-based model, 2023.

[83] Kui Ren, Ziqi Yang, Li Lu, Jian Liu, Yiming Li, Jie Wan, Xiaodi Zhao, Xianheng Feng, and Shuo Shao. Sok: On the role and future of aigc watermarking in the era of gen-ai, 2024.

[84] Ghadeer Ghosheh, Jin Li, and Tingting Zhu. A review of generative adversarial networks for electronic health records: applications, evaluation measures and data sources. *arXiv preprint arXiv:2203.07018*, 2022.

Disclaimer:

SurveyX is an AI-powered system designed to automate the generation of surveys. While it aims to produce high-quality, coherent, and comprehensive surveys with accurate citations, the final output is derived from the AI's synthesis of pre-processed materials, which may contain limitations or inaccuracies. As such, the generated content should not be used for academic publication or formal submissions and must be independently reviewed and verified. The developers of SurveyX do not assume responsibility for any errors or consequences arising from the use of the generated surveys.