CTICTR
CURRENT TRENDS IN INFORMATION COMMUNICATION TECHNOLOGY RESEARCH

BABCOCK UNIVERSITY

ISSN: 2814-1709

CTICTR 1(2): 1 – 19 (December 2022)

Reversible Image Steganography: A Review of Existing GAN-based Models

Uchenna Nzenwata[a], Oludele Awodele[a,a], Sunday Idowu[b,b], Afolashade Kuyoro[a,c]

[a]*Computer Science Department, Babcock University, Ilishan-Remo, Ogun State, Nigeria*
[b]*Software Engineering Department, Babcock University, Ilishan-Remo, Ogun State, Nigeria*
[a]*awodeleo@babcock.edu.ng*
[b]*idowus@babcock.edu.ng*
[c]*kuyoros@babcock.edu.ng*

*Correspondence Email: *nzenwatau@babcock.edu.ng*

Abstract

Image steganography is the method of concealing information, which can be text, image, or video, under a cover image in a way that it is invisible to the human eyes. Payload capacity, security and robustness are the features of steganography. A high-quality image steganography ensures a large payload capacity, superior security, and robustness against adversarial attacks. Image steganography is considered to be reversible if the concealed information's original state can be restored with little or no loss in the pixel values and textures. The main goal of this paper is to look at some of the typical techniques and models that are used to build image steganography systems by adopting the Preview, Question, Read, Summarize (PQRS) method to explain how the payload capacity, security, and robustness of these models may be quantified for future research and recommend GAN-based deep learning methods for improving existing reversible image steganography models.

***Keywords*: Image steganography, GAN-based models, Payload capacity, PSNR, Reversibility, SSIM**

1. Introduction

The methods of confidential communication and data preservation had evolved over the years [1]. Before the existence of computing devices, confidential communication and data

[a] Uchenna Nzenwata

preservation were preferably done via a person's physical presence – running errands either by employing one-to-one, many-to-one, one-to-many, or many-to-many information sharing schemes. In any case, the integrity of these methods is frequently questioned as a result of unanticipated human limitations, which endangers confidential communication and secret data storage [2],[3]. Many solutions, including cryptography, encryption, watermarking, and steganography, have been developed to address this issue by guaranteeing that data and information are shared in a secure way [4][5].

In [6], Cryptography and encryption follow similar technique, where the information is first converted to cipher form before it can be communicated via secured channels. Watermarking and steganography on the other hand [7], operate by concealing or embedding data in a cover media. Watermarking is commonly used to conceal information in a patented product in order to ensure the product's authenticity and copyright. According to [8], steganography conceals information in such a way that it goes undetected. This paper therefore defines image steganography as the use of image medium for hiding information in such a manner that it remains unsuspicious to information hijackers. Therefore, reversible image steganography is a method used for hiding covert messages under image covers with the intent of recovering the original status of the covert messages, and it has been widely used in confidential communication and data storage [9].

Despite the numerous uses discovered for image steganography, it has drawbacks such as poor payload capacity, security, and robustness. Several studies [10-13], have found that attempts to address these concerns either result in a trade-off between payload capacity and security, or in difficulties due to the complexities associated with the methodology used to construct image steganography systems.

2. Concept of Steganography

Information hiding is sometimes used interchangeably with “Steganography”, which conceals information leaving no suspicion. In Figure 1, Steganography is clearly represented as one of the branches of Information hiding. In [14], steganography is identified as a branch of information hiding which hides the existence of the secret data in a cover medium.

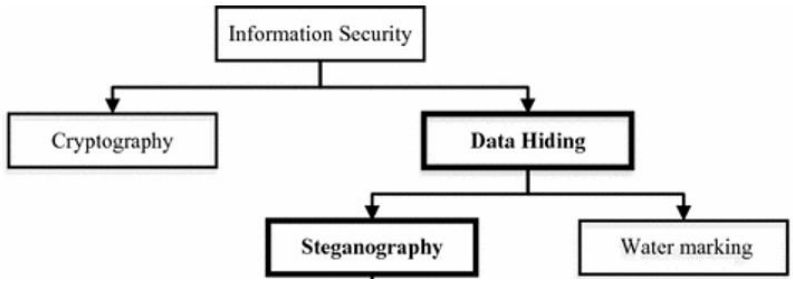

Figure 1: Steganography: A branch in Data Hiding [14]

In [15], the difference between Steganography, Watermarking and Cryptography was emphasized. It was stated that these tools employed similar ideas, but have different approaches and purposes in cyber security. Also, Steganography and Watermarking embed information onto a cover medium, while Cryptography obscures information. In a similar study [16], the individual roles of Steganography, Watermarking and Cryptography in information security was explained. [16] stated that with steganography, information security is catered for by the hiding of secret data. While watermarking caters for information security via embedding data, which finds its application around copyright protection and barcode

systems; Table 1 compares the characteristics of Steganography, Watermarking, and Cryptography.

Table 1: Comaprison of Information Hiding Techniques [17]

Characteristics	Steganography	Watermarking	Cryptography
Goal	Preserve the confidential data from the detection	Preserve the authenticity of the cover media	Obfuscate the content or the data
Cover Selection	Free cover selection	Restriction	N/A
Challenges	Imperceptibility, Security and Capacity	Robustness	Robustness
Key	Optional	Optional	Compulsory
Output	Stego-file	Watermarked-file	Cipher-text
Visibility	Certainly not	Sometimes	Always
The System is Invalid if	Detected	Removed or replaced	De-ciphered
Attack	Steganalysis	Any image analysis	Cryptanalysis

3. Steganography and Steganalysis

Figure 2 shows a steganographic process, where secret messages ($\boldsymbol{S_m}$) are encoded within the properties of the cover medium ($\boldsymbol{C_m}$), by employing a suitable encoding algorithmic function $\boldsymbol{f}$ and a stego key ($\boldsymbol{S_k}$). The output is usually a stego object ($\boldsymbol{S_o}$). The stego object is transmitted through a secured transmission channel from the sender to the receiver. The receiver will use the provided stego key to extract the secret message from the stego object. The decoding process takes place at the receiver's end, which is termed steganalysis. Steganalysis is the counter part of steganography. In [18], it is defined as the art or science of discovering hidden data in cover object.

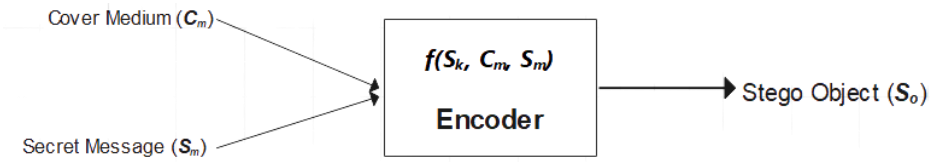

Figure 2: Steganography – Encoding Block Concept

Steganalysis process always takes place at the decoding phase. In Figure 3, the stego object and the stego key serve as the inputs to the decoder. The steganalysis outputs are the cover medium and the secret message. Depending on the approach used, it either results into lossy or lossless steganography. In the case of reversibility, the extracted secret message maintains similar status as the retrieved secret message. In Figure 4, the encoding and the decoding phases are combined to show a basic steganographic conceptual framework.

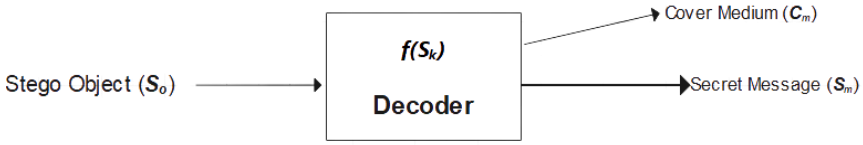

Figure 3: Steganalysis – Decoding Block Concept

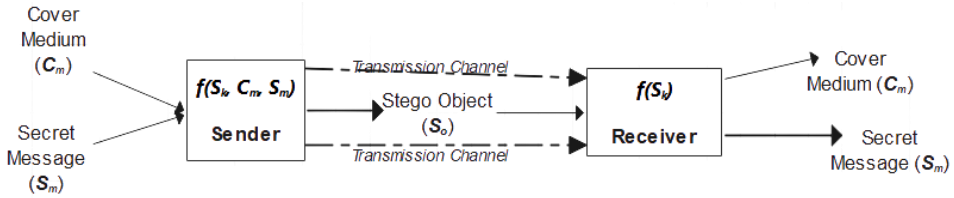

Figure 4: Basic Steganography Concept

4. Image Steganography

Image steganography is the process of concealing data by using the cover object as an image. Because the digital representation of an image contains a large number of bits, digital images are frequently employed as a cover source in digital steganography [19].

4.1 Image Steganography based on Image format

There are known image steganography based on image formats, owing to the fact that images can be represented in a variety of formats as shown in figure 5. In [20-22], the most prevalent image formats were identified, and they are the Joint Photographic Exchange Group (JPEG), Bit Map (BMP), Graphics Interchange Format (GIF), Portable Network Graphics (PNG), and Tag Image File Format (TIFF).

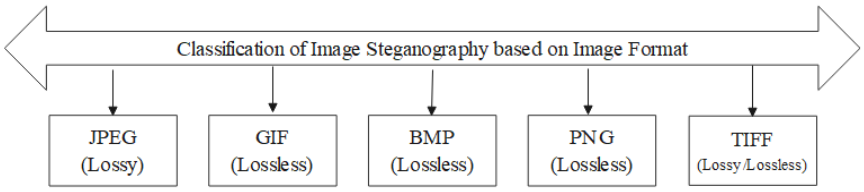

Figure 5: Classification of Image Steganography based on Image Format [20, 22]

i. **Joint Photographic Exchange Group (JPEG) Image** is a common cover image format used in steganography. The compression approach is a lossy compression, which implies that some visual quality will be lost in the compression process. Regardless of its lossy compression, JPEG has a very good image quality with many redundant bits that provide enough possibility to hide a sizeable value of secret information. This study adapted the use of JPEG images in such a way that the image datasets are compressed using the downsizing arm of the Generative Adversarial Network (GAN). This was done to reduce the visual quality effect of the lossy compression.

ii. **Bit Map (BMP) Image files** are device-independent files based on the RGB colour model that are often used on Windows systems. The size and colour depth information are found in the header area. That is, the values of each pixel on a line are stored in the data area. The challenge with this image format is that during the process of embedding secret data, the outcome most times results in grey images.

iii. **Graphics Interchange Format (GIF) Image** colour of the pixel is referenced from a palette database of up to 256 unique colours translated to the 24-bit RGB colour space, and it was introduced by CompuServe in 1987. It supports up to 8 bits per

pixel. The 24-bit RGB value of a pixel can be changed by LSB embedding a GIF image, which can result in a change in the palette colour. Although this image format has no loss of quality, but it is simple to decode and secret data displayed.

iv. **Portable Network Graphics (PNG) Image format** was first used in 1999. PNG can store three sorts of images: true colour, grayscale, and palette-based ("8-bit"). JPEG only supports the first two, whereas GIF only supports the third (although it can fake grey scale by using a grey palette). PNG allows for various alpha channels (transparency). PNG usually comes with a very high visual quality. It was not considered suitable for this study because the image dimension is wide with large capacity, which affects the neural network training time because the image features cannot be down-sampled and it is lossless.

v. **Tag Image File Format (TIFF)** is a type of computer file that stores raster graphics and picture data. TIFFs, which are popular among photographers, are a convenient option to keep high-quality images before processing if you wish to avoid lossy file formats. Due to its large size and unprocessed format, TIFF is not a suitable image steganography tool.

5. Image Steganography Techniques

There are many techniques used in exploring steganography. According to [21], the approach used is influenced by the nature of the steganographic cover to be utilised. Based on the scope of this study, consideration is given to digital image steganography. Hence, this section explores different steganographic techniques used with digital images as cover media.

Depending on the application and phases involved in the embedding process, many approaches to image steganography have been presented in the past. As a result, these approaches were categorised according to [17] in Figure 6. These approaches are presented based on the type of cover image used (2D or 3D images), target application type, retrieval method (reversible or irreversible), nature of embedding process (spatial or transform domain), and adaptive steganography.

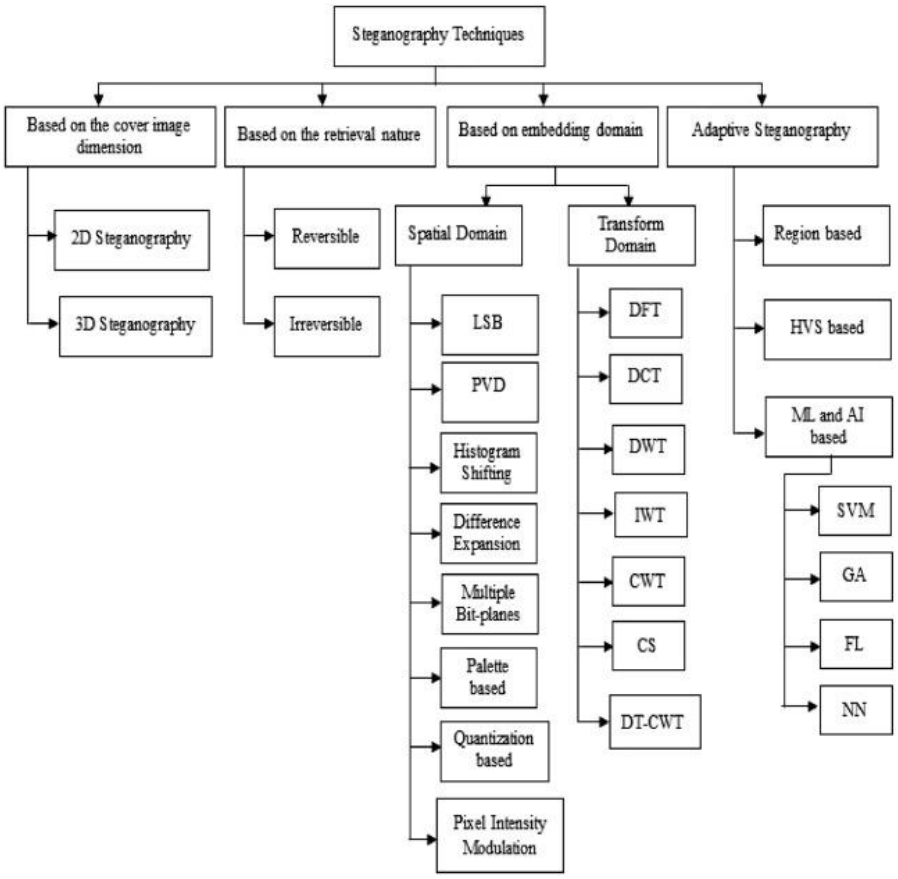

Figure 6: Classification of Image Steganography Techniques [17]

5.1 Based on Cover Image Dimension

Images are available in a variety of formats. It can be two-dimensional images like grayscale or binary, three-dimensional images like tri-color (Red, Green, Blue) RGB images, or multi-slice images like Magnetic Resonance Imaging (MRI) [23].

i. **In 2D Steganography**, the hidden data is embedded across the 2D plane of the cover image in 2D image steganography. The embedding can be done on either the spatial domain pixel values or the transform domain coefficient values. The fundamental advantage of 2D steganography is that popular public images may be utilised as cover media to assist convey hidden data in a private mode, which is why more study is being done in this field.
ii. **In 3D Steganography**, there are several ways to 3D steganography, including geometrical domain steganography [24], topological domain steganography, and representation domain steganography. The primary strategy in 3D steganography is to embed the secret bit streams in the vertices of a 3D cover image. When a greater payload capacity is required, the bits are embedded in the vertices of 3D geometrical models and are favoured over other models (topological and representation). 3D steganography, like 2D steganography, may be employed in both the spatial and transform domains. However, temporal complexity is an issue here since a large number of points must be addressed in 3D space for embedding.

5.2 Based on Retrieval Nature

Steganography can be carried out in two major modes, based on the retrieval nature: irreversible or non-reversible and reversible modes [16]. In image steganography, secret message is embedded into the bits of the cover image and the resulting stego-image is sent to the receiver, who extracts or retrieves the secret message either by irreversible form or reversible form. The later takes the cover image and the secret message into consideration for analysis, while the former does not consider the retrieval of the cover image, the only the secret message.

i. Irreversible Steganography

From the word 'Irreversible', it means the cover image cannot be recovered to its original state. Once the stego object ($\mathbf{S_o}$), herein, the object with the embedded secret message ($\mathbf{S_m}$) is sent and received, the cover medium ($\mathbf{C_m}$), which is the carrier of the secret message cannot be recovered in the process of retrieving the secret message. Most times, the cover media are distorted or destroyed. That is, the extraction process leaves the cover medium destroyed [25]. In other words, in an irreversible approach, the cover image bits are irreversibly altered, and recovering the altered cover image bits from the stego-image is very difficult. As a result, the approach is recommended for applications in which no relevant information in the cover image has to be conveyed to the recipient.

ii. Reversible Steganography

Unlike the irreversible steganography, both the cover medium and the secret message are retrievable without loss. The idea behind the reversible steganography is for a fragile authentication. Reversibility is more typically utilised in watermarking than in steganography. The cover data may be more significant than the secret image in this case, and any leakage or loss of these information would be vital, and methods are constructed in such a way that retrieval accuracy is high. Therefore, reversible steganography finds application areas in the medical [26] and military fields [27], where the original cover medium is as important as the secret message; any visible change on original cover medium during the transmission can alter the intelligence of the steganography technique and affect the overall results.

5.3 Based on Embedding Domain

The embedding technique, which is employed to hide secret data over the cover image, is the main support for steganographic methods. The spatial and transform domains are the options for embedding secret data across the cover image.

a. Spatial Domain Technique

This domain technique is the easiest and the simplest way of embedding secret data into digital images by modifying the pixel values of the cover image. The techniques under spatial domain technique encode the secret message bits directly or indirectly by using the cover image pixel intensity value levels. This is so, because these cover media have tremendous amount of wasted or redundant bits; it is these bits that the steganography program will take advantage of and use to hide another message, on the bit level, within the digital cover [28].

The primary steganographic schemes that fall under the spatial domain technique are as follows:

i. Least Significant Bit (LSB)

LSB technique is commonly used due to its simplicity in implementation which makes it efficient, but a recent study conducted by [29], established that there are lots of decoding algorithms available for its steganalysis by guessing the location where the secret message is embedded on the cover medium. The study gave a simplified demonstration to Least-Significant Bit (LSB) technique by using the following string of bytes to represent part of a digital image cover 10000100 10000110 100001001 10001101 01111001 01100101 01001010 00100110. Each byte is comprised of eight bits; these bits make up a color value in the cover image, with a shade of red, or blue, or green. The bits that make up the byte go from left to right in order of importance to the color value they are representing. They tried changing the first bit in the first string from a 1 (10000100) to a 0 (00000100) and it drastically change the color, as opposed to changing the last number from a 0 (10000100) to a 1 (10000101). Hence, it is that last bit that is considered the least significant, because changing its value has little effect on the information the byte is representing. LSB technique does well with digital media covers like pictures, audio files or video files.

LSB works well with both grayscale and color photos. However, the LSB approach has several limitations. Depending on the pixel, altering the LSB might have a dramatic effect on the pixel's attributes, making it appear out of place in the image and hence susceptible to detection. This issue has the potential to restrict the degree of swapped bits, and hence the size of the secret message. Another issue with this type of message concealment is the image's resistance to modification. If the image is cropped or rotated, the algorithm will be unable to determine which least-significant bits are part of the message and which are merely present.

ii. Pixel Value Differencing (PVD)

The secret message is encoded in this approach by examining the differences between the pixel values of two subsequent pixels. Many ways to PVD steganography have been presented by analysing the correlation of pixels, but the fundamental disadvantage of PVD approach is the absence of security, despite the fact that it delivers significantly high image perceptibility factor.

iii. Histogram Shifting

In histogram shifting, the main idea of the method is by shifting the histogram points of the cover image. The lowest and highest points in the cover image histogram are calculated, and the embedding process is then carried out by modifying these lowest and highest points. The approach improves imperceptibility and increases payload capacity. The fundamental benefit of histogram-based image hiding is that it provides reversible data hiding.

iv. Expansion based

According to this approach, the secret data is integrated across different pixel pairings. Different approaches are used to increase the difference values, and secret data bits are inserted across this enlarged difference range. The disadvantage of different expanding approaches is that they are confined to target applications where cover image is important and the communication channel is less vulnerable to intruder assaults.

v. Multiple Bit-planes based

This method is an extended version of the LSB method. It hides secret data bits using the bit planes of the cover image. If the cover image is properly chosen, the bit plane

selection is perfect, and hidden bits may be embedded with no loss of visual quality. The fundamental disadvantage of this technique is that the image imperceptibility may be lost if the bit slice is chosen incorrectly. The method also has the drawbacks associated with the average LSB methodology.

vi. Quantization based

The steganographic system in this category employs any type of compression encoding technology to conceal hidden data bits. Any common compression codec, such as JPEG, vector quantization, and so on, can be used as the encoding system. The hidden data is often broken into small blocks of data sub samples, and these little data fragments are placed with the encoded carrier images. The same coding is applied to the stego-image at the receiver, and the secret data is extracted via the inverse embedding method.

vii. Palette based

In this method, palette-based images are used as cover image. The image formats such as PNG, GIF, and TIFF are appropriate for this method. A secret key is utilised to produce pseudo random numbers, and the secret data bit that is chosen is embedded on a single cover pixel. Instead of the original colour, the colour in the palette with the same parity as the secret bit is utilised for the embedding operation.

viii. Pixel Intensity Modulation based

Pixel intensity modulation or adjustment-based steganographic systems was presented as a variation on the LSB steganographic type pixel intensity adjustment embedding. The secret data bits are embedded in the intensity adjustment between neighbouring pixels or blocks, depending on the nature of the embedding mechanism. Because of the indirect embedding process, these approaches assist to create higher quality stego-images when compared to LSB modification systems.

b. Transform Domain Technique

Transform domain techniques could also be called frequency domain techniques. The transform domain techniques are based on manipulating the orthogonal transform of the image rather than the image itself. Transformation domain methods are well suited for image processing that are based on frequency content. Because the typical transform domain allows action on the image's frequency content, high frequency features such as edges and other subtle information may be easily increased.

The secret bits are concealed behind the sub-band frequency coefficients in transform domain techniques. The embedding and decoding procedure in the transform domain are more sophisticated than the techniques employed in the time domain. This will increase the system's security. Another advantage is that most Frequency domain approaches are less sensitive to compression, cropping, scaling, and rotation attacks. As a result, transform-based solutions are more effective at preserving stego-image quality and are less detectable in an unprotected channel.

In the subject of image steganography, according to [30], many transform domain methods are utilised, the most prominent of which are the;

i. Discrete Fourier Transform (DFT)

DFT is a prominent transform method in signal processing. An image may be divided into appropriate frequency components using DFT (sines and cosines). These

transform coefficients may also be changed according to the secret information in picture steganography, making them an important tool in image data concealment. The modulated transform coefficients are then transferred to picture form and may be utilised to generate the stego-image. The stego-image is fragmented into frequency elements again at the receiver, and hidden data can be recovered from there.

ii. Discrete Cosine Transform (DCT)

To transfer an image from the spatial domain to the frequency domain, DCT is one of the most frequent and successful images transform techniques. The DCT coefficients are adjusted according to the secret data bits in basic DCT-based steganography. The image is divided into its appropriate high, medium, and low frequency components in DCT steganography. Low-frequency subbands contain the most critical information, whereas high-fidelity features are found in the High Frequency bands. The JPEG compression paradigm is used to hide secret data bits in DCT coefficients.

iii. Discrete Wavelet Transform (DWT)

The wavelet transform of digital pictures can be used to localise time and frequency. To discover the subband images, DWT uses a variety of filtering and down sampling techniques. DWT can was used for image steganographic applications based on the characteristics of distinct wavelets. The DWT of the cover image is usually used in DWT-based image steganography methods, and the coefficient space is adjusted according to the secret message bits. The lack of intelligent embedding methodologies in basic DWT systems is a hurdle to achieving optimum accuracy. As a result, various wavelet transform systems improvements appeared on the scene to increase the performance of embedding systems.

iv. Other variants of these fundamental transforms

Some DWT variations like Integer Wavelet Transform (IWT) is defined as a modified wavelet transform format in which image features can be conveyed in a variety of resolution levels; Complex Wavelet Transform (CWT) is regarded as an improved variant of the DWT; Dual-Tree Complex Wavelet Transform (DTCWT)-based steganography is comparable to CWT, except it employs dual tree complex wavelet decomposition.

5.4 Adaptive Steganography Technique

The adaptive steganography technique is a subset of the spatial and transforms method. It can be referred to as “Model-Based” or "Statistics-aware embedding and Masking”. Adaptive techniques may be divided into several categories based on their nature and manner of adaptiveness [17]. This contains;

i. Region-based Steganography

The embedding process of this method selects the most relevant areas or characteristics in the cover image. The region-based method concentrates the embedding on areas with significant texture and edge features rather than smooth sections. This helps to fight picture quality loss since changes in textured and edge regions are not as quickly noticed as changes in smooth regions. The major advantages of these sorts of systems are their robustness and better imperceptibility, whereas the main disadvantage is their low embedding rate.

ii. Human Vision System (HVS)-based steganography

HVS-based steganography makes use of a human's perception of an image. Human eyesight contains many dark spots and has difficulty distinguishing visual features in certain situations. When dense embedding is done over complex texture regions and edge areas in the pictures, HVS is insensitive to modest variations in intensity in a smooth area and even indistinguishable. As a result, these strategies often employ processes to recognize such target regions in images.

iii. Machine Learning and Artificial Intelligence Techniques based Steganography

Advanced machine learning algorithms are used in machine learning-based steganographic approaches to produce efficient steganography systems. Popular steganographic machine learning methods include any algorithms that seek to apply artificial intelligence approaches for pre-processing and/or defining embedding nature and/or restricting or extending the embedding stages or extraction process. Some of the algorithms are as follows:

a. **Support Vector Machine (SVM)**: SVM is a supervise machine learning algorithm that can be used as a classifier in several elements of image watermarking and steganography. [31] paved the ground for the use of SVM for optimising steganography-based approaches. They stated that SVM is well suited for locating embedding spots with high confidentiality and imperceptibility while minimising retrieval error through the use of quicker computing phases.
b. **Genetic Algorithm (GA):** GA was first used for selecting aptly the coefficients in image steganography. They adapted this idea from the fundamental intents of GA: selection, crossover and mutation, which are used for selecting the best offspring to pass through future generations. Their approach is still considered as one of the best means of selecting good region for embedding secret data due to its good payload capacity.
c. **Fuzzy Logic:** The fuzzy logic-based approach is more concerned with preserving visual quality in order to improve the imperceptibility of the steganography at the expense of increased modelling complexity. Fuzzy techniques have been implemented in a variety of ways. For example, a Fuzzy Inference System (FIS) with Human Vision System (HVS) is utilised for the decision-making phase from local statistical, texture, and brightness information-based feature vectors.
d. **Neural Network (NN):** The neural network-based steganography systems analyse the image details thoroughly to determine the system robustness and imperceptibility of the schemes. The majority of such techniques employ a back-propagation algorithm to determine the proper embedding sites, or it is included into the secret extraction process. In general, NN outperforms other categorization algorithms. Their decision-making process is built on non-linear and adaptive phases, making them helpful in a variety of settings. The 'Black Box' nature, on the other hand, is a key worry when using a NN approach. The fundamental reason for this impact is the relationship between weight changes in the training stage, and thus the key challenge in neural network-based systems is to assign parameters such as number of layers, number of neurons, and so on. Another disadvantage of these approaches is their time-consuming nature.

5.5 Direct Sequence Spread Spectrum Techniques

The information stream to be conveyed is broken into tiny parts in direct sequence spread spectrum. Each component is assigned to a different frequency channel of the spectrum. At the moment of transmission, the data signal is mixed with a higher data-rate bit sequence that splits the data according to a predefined spread ratio. Redundant data-rate bit sequence coding aids signal immunity and allows the original data to be restored if any of the data bits are destroyed during transmission.

5.6 Frequency Hoping Spread Spectrum Techniques

This method separates a large portion of the bandwidth spectrum into a variety of broadcast frequencies. In general, frequency-hopping devices utilize less power and are less expensive, while direct sequence spread-spectrum systems often perform better and are more dependable [32].

5.7 Statistical Techniques

Statistical techniques use a "1-bit" steganographic scheme. This technique is described as the technique that inserts only one bit of information in a digital carrier, resulting in a statistical change, although a little one. They further explained that the statistical change in the cover implies a "1," whereas a cover that remains static denotes a "0". The recipient's aptitude to discriminate between modified and unmodified covers is the foundation of this system.

5.8 Distortion Techniques

To hide information, this steganography technique alters the cover image. When the algorithm compares the modified and distorted cover image to the original cover image, the hidden message is extracted. This technique is used in lossy steganography.

5.9 Cover Generation Techniques

Cover generation techniques are perhaps the most distinct of the seven strategies. Normally, a cover item is used to conceal a hidden message, but that is not the case here. A cover generating method provides a cover only for the purpose of concealing information. Spam Mimic is a great example of a cover generating approach.

6. Methodology

Over 150 main articles providing expertise on image steganography were obtained for this study's review from reputable journals. These articles were analyzed and synthesized using Preview, Question, Read, Summarize (PQRS) method. This method was adopted so as to easily identify a number of articles that are closely related to the focus of reversible image steganography. Five (5) closely related literatures were characterized and result summary was obtained.

6.1 Characterization of Reversible Image Steganography Deep Learning Models

In image steganography, reversibility has been achieved utilising a variety of approaches found in spatial domain and transform domain techniques. Complex mathematical computations had proven to be a difficulty for these techniques [33]. To achieve reversible image steganography, deep learning algorithms based on the Generative Adversarial Network (GAN) have been employed effectively [34-36]. This is owing to the GAN-based algorithms endowed two game arms (generator and discriminator).

There are several existing deep learning models that have been employed to achieve image steganography, but just a few of these models can be adapted to a reversible image steganography, where the recovery of the cover images is as important as the recovery of the secret images with good payload and security features [35]. The fundamental building block for reversible image steganography using deep learning model is the Generative Adversarial Network (GAN) [36]. In [37], the structural analysis of GAN was done, and it was presented that the structure of a GAN completely corresponds to the structure of a steganography system as shown in Figures 7a and 7b, with the exception that the basic GAN discriminator will have to be modified in order to be able to do extraction on stego image. That is, GAN has two gaming arms, the generator and the discriminator which corresponds to the encoding phase and the decoding phase of a steganography system respectively with an extraction network. Also, GAN's generator and discriminator features enable image-to-image transformation, which makes it easy to be used for reversible image steganography [38]. Also, GAN-based models are said to resist steganalysis more effectively than other deep learning models, but are faced with payload challenges.

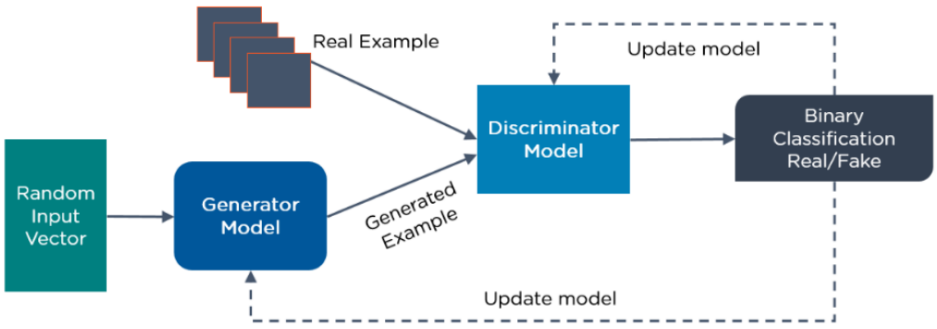

Figure 7a: Generative Adversarial Networks (GANs) [37]

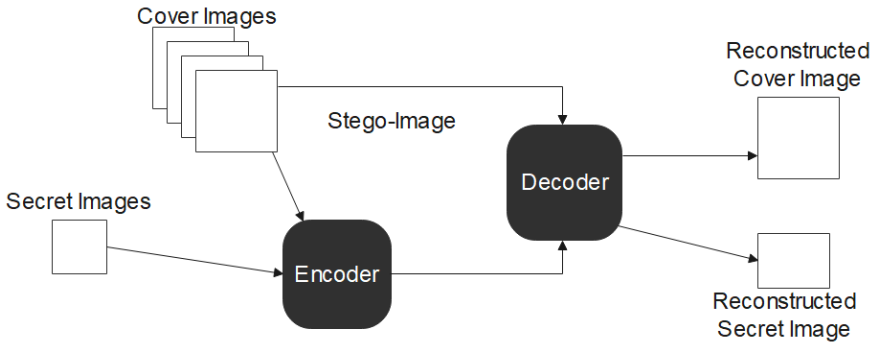

Figure 7b: GAN-based Reversible Image Steganography Correspondence [37]

As a result, related studies have developed and use variants of GAN-based models for images steganography. Some of these GAN-based variants are Steganographic Generative Adversarial Network (SGAN) [39], Generative Reversible Data Hiding (GRDH) [40], Modified Cycle-Consistent Generative Adversarial Network (Mod-CycleGAN) [41], Wasserstein Generative Adversarial Network Gradient Penalty (WGAN-GP) [42], GAN based steganography framework UT-GAN [43], and many others. Table 2 summarizes the features, strength and weakness of some the GAN-based models.

Table 2: Characterization of GAN-based Models

SN	Model	Author & Year	Features	Strength	Weakness
1.	SGAN	Chavdarova & Fleuret (2018)	The model has two GAN-base generator G_0 and G_1, and a "messenger" discriminator. Only one generator G_0 was trained and used for the purpose measuring the efficiency of SGAN against GAN. "messengers" discriminator was trained to improve the steganalysis of GAN.	The model shows an improved GAN model in terms of its accuracy.	The model shows poor security as the stego images show a visual effect distortion.
2	GRDH	Zhang et al. (2019)	It Uses the image-to-image translation cycle GAN model as the image generator, and uses the basic GAN as the discriminator.	The payload was good because a numpy tool was used to generate random bit which informs the choice of cover selection	No extraction model was trained, but uses key for extraction. Poor security.3
3	ModCycleGAN	Kuppusamy et al.(2020)	Uses the basic GAN generator for encoding; a modified discriminator for decoding, but the discriminator was not trained.	The security of the model is good and the use of a predefined pix-to-pix model at the encoding phase.	The payload capacity of the model is poor and it was an incomplete model.
4	WGAN-GP	Li et al. (2020)	The features are the generator network, discriminator network, and the extraction network using CNN. No modification was done on the encoder.	A reversibility steganography, with good security measure	Poor payload capacity.
5	UT-GAN	Li, Li, Tan and Li, (2021)	This incorporated a ThiNet model for	The security of the model was	The payload capacity is

			pruning a UT-GAN. ThiNet enables appropriate cover selection, which was trained alongside with the UT-GAN generator. The discriminator uses CNN model for the extraction.	encouraging and performs well when subjected to steganalysis.	poor resulting from the over parameterized ThiNet model.

6.2 Evaluation Metrics

As GANs do not have objective evaluation functions as supposedly stated in the paper [44], the proposed image quality metrics are as follows:

i. **Peak Signal-to-Noise Ratio (PSNR):** This metric is often used to compare the quality of the stego-image (C') to its matching cover image (C) by measuring the peak signal-to-noise ratio of two images, (C') and (C). The greater the PSNR in decibel (dB), the higher the visual quality. It is calculated using the following equation.

$$PSNR_{(C',C)} = 10log_{10}\left(\frac{Max^2C}{MSE}\right) \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots 1$$

Taking the square root of the equation 1,

$$PSNR_{(C',C)} = 20log_{10}\left(\frac{MaxC}{\sqrt{MSE}}\right) \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots 2$$

$$= 20log_{10}(Max^2C) - 20log_{10}(MSE_{C',C}) \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots 3$$

Where MSE is the Mean Squared Error, and it is given by,

$$MSE_{(C',C)} = \frac{1}{mn}\sum_{i=0}^{m-1}\sum_{j=0}^{n-1}(C(i,j) - C'(i,j))^2 \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots 4$$

C = the matrix data of our cover image
C' = the matrix data of our stego-image
m = the numbers of rows of pixels of the images and i represents the index of that row
n = the numbers of columns of pixels of the images and j represents the index of that column
maxC = the maximum signal value that exists in the cover image.
The PSNR values can also be used to ascertain the level of the system's security.

ii. **Structural Similarity Index (SSIM):** This can be used in case there is an undefined value of PSNR, where the MSE = 0, and this makes human perception not important. As such, SSIM is necessary, since it aids the Human Vision System (HVS). The SSI calculates the degree of similarity between two images. It is regarded as one of the most often used quality metrics, and it is connected to the single-scale measurement, which performs best when used at an appropriate scale. The best denoising approach is indicated by the highest SSIM.
SSIM was used to compare corresponding pixels and their neighborhoods in the cover and the stego-images, denoted by C and C', using three quantities: luminance (I), contrast (C), and structure (S). The equation is given thus:

$$I(C,C') = \frac{2\mu_C\mu_{C'} + k_1}{\mu_C^2\mu_{C'}^2 + k_1} \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots 5$$

$$C(C, C') = \frac{2\sigma_C \sigma_{C'} + k_2}{\sigma_C^2 \sigma_{C'}^2 + k_2} \text{.....6}$$

$$S(C, C') = \frac{\sigma_{CC'} + k_3}{\sigma_C \sigma_{C'} + k_3} \text{.....7}$$

Where the variables: $\mu_C, \mu_{C'}$, σ_C and $\sigma_{C'}$ are the mean and standard deviations of the pixel intensity in a small image patch centered on C or C'. The variable $\sigma_{CC'}$ represents the sample correlation coefficient between matching pixels in patches centered on C and C'. k_1, k_2, and $k_3$are minor values that were included for numerical stability.
To derive the SSIM equation, equations 5 – 7 were combined to produce equation 8.

$$SSIM(c, c') = [l(c, c')^{\alpha}] . [C(c, c')^{\beta}] . [S(c, c')^{\gamma}] \text{ ... ...8}$$

Where α, β and γ are the positive constants that must be greater than zero $(\alpha, \beta, \gamma > 0)$.

iii. Payload Capacity: In other to calculate the size of concealable secret message and maximize the embedding capacity of the steganography system, this study employed the bit per pixel (bpp) method as shown in equation 9.

$$BPP = \frac{Number\ of\ secret\ bits\ embeded}{Total\ pixel\ on\ the\ cover\ image} \text{ ..... 9}$$

7. Conclusion and Recommendation

The outcome of the PQRS methods helped to identify Five (5) closely related articles that are GAN-based image steganography. As evident in [34], GAN deep learning architecture can produce reversibility in image steganography by using the two gaming arms, the generator and the discriminator arms. Also, by implementing steganography using GAN makes the steganography process more resilient, and the resulting stego image is more hidden and secure without, but does not take care of the certainty in achieving good pay load capacity. To cater for the payload capacity [45] suggested that some sort of model need to be set in place for an appropriate cover selection which can be used for encoding purposes. Also, the characterized models show that in achieving steganography using GAN-base model, there must be an appropriate cover selection model, an encoder model, decoder model, which may be modified to solve the inadequacies in payload and security of image steganography systems, and with the gamming arms of GAN-base models, reversible image steganography can be achieved. It is therefore recommended that to achieve

References

[1] Gupta, O., & Goyal, N. (2021). The evolution of data gathering static and mobility models in underwater wireless sensor networks: A survey. *Journal of Ambient Intelligence and Humanized Computing*, 1-17.

[2] Rustad, M. L., & Koenig, T. H. (2019). Towards a global data privacy standard. *Fla. L. Rev.*, *71*, 365.

[3] Schwartz, P. M., & Peifer, K. N. (2017). Transatlantic Data Privacy Law. *Geo. LJ*, *106*, 115.

[4] Singh, A. K., Anand, A., Lv, Z., Ko, H., & Mohan, A. (2021). A survey on healthcare data: a security perspective. *ACM Transactions on Multimidia Computing Communications and Applications*, *17*(2s), 1-26.

[5] Basahel, A. M., Yamin, M., & Abi Sen, A. A. (2019). Enhancing security of transmitted data by improved steganography method. *IJcSNS*, *19*(4), 239-244.

[6] Mostafa, G., & Alexan, W. (2019, June). A high capacity double-layer gray code based security scheme for secure data embedding. In *2019 International Symposium on Networks, Computers and Communications (ISNCC)* (pp. 1-6). IEEE.

[7] Bhargava, S., & Mukhija, M. (2019). HIDE IMAGE AND TEXT USING LSB, DWT AND RSA BASED ON IMAGE STEGANOGRAPHY. *ICTACT Journal on Image & Video Processing*, *9*(3).

[8] Kose, J., Chia, O. B., & Baboolal, V. (2020). Review and Test of Steganography Techniques. *arXiv preprint arXiv:2012.08460*.

[9] Sahu, A. K., & Swain, G. (2020). Reversible image steganography using dual-layer LSB matching. *Sensing and Imaging*, *21*(1), 1-21.

[10] Kadhim, Inas Jawad, Prashan Premaratne, Peter James Vial, and Brendan Halloran. "Comprehensive survey of image steganography: Techniques, Evaluations, and trends in future research." *Neurocomputing* 335 (2019): 299-326.

[11] Hashim, M., MOHD RAHIM, M. S., & ALWAN, A. A. (2018). A REVIEW AND OPEN ISSUES OF MULTIFARIOUS IMAGE STEGANOGRAPHY TECHNIQUES IN SPATIAL DOMAIN. *Journal of Theoretical & Applied Information Technology*, *96*(4).

[12] Hussain, M., Wahab, A. W. A., Idris, Y. I. B., Ho, A. T., & Jung, K. H. (2018). Image steganography in spatial domain: A survey. *Signal Processing: Image Communication*, *65*, 46-66.

[13] Rashid, R. D., & Majeed, T. F. (2019, March). Edge based image steganography: problems and solution. In *2019 International Conference on Communications, Signal Processing, and their Applications (ICCSPA)* (pp. 1-5). IEEE.

[14] Thanikaiselvan, V., Shastri, S., & Ahmad, S. (2017). Information hiding: steganography. In *Intelligent Techniques in Signal Processing for Multimedia Security* (pp. 65-91). Springer, Cham.

[15] Agbaje, M., Awodele, O., & Ogbonna, C. (2015). Applications of Digital Watermarking to Cyber Security (Cyber Watermarking). *Proceedings of the 2015 InSITE Conference*, 001–011. https://doi.org/10.28945/2138

[16] Mohamed, K. S. (2020). Data Hiding: Steganography and Watermarking. In *New Frontiers in Cryptography* (pp. 89-98). Springer, Cham.

[17] Kadhim, I. J., Premaratne, P., Vial, P. J., & Halloran, B. (2019). Comprehensive survey of image steganography: Techniques, Evaluations, and trends in future research. *Neurocomputing*, *335*, 299–326. https://doi.org/10.1016/j.neucom.2018.06.075

[18] Wang, H., & Wang, S. (2004). Cyber warfare: steganography vs. steganalysis. *Communications of the ACM*, *47*(10), 76-82.

[19] Yahya, A. (2019). *Steganography Techniques for Digital Images*. Springer International Publishing.

[20] MBadr, S., Ismaial, G., & H. Khalil, A. (2014). A Review on Steganalysis Techniques: From Image Format Point of View. *International Journal of Computer Applications*, *102*(4), 11–19. https://doi.org/10.5120/17802-8617

[21] Menon, N. (2017, December). A survey on image steganography. In *2017 International Conference on Technological Advancements in Power and Energy (TAP Energy)* (pp. 1-5). IEEE.

[22] Ansari, A. S., Mohammadi, M. S., & Parvez, M. T. (2019). A comparative study of recent steganography techniques for multiple image formats. *International Journal of Computer Network and Information Security*, *11*(1), 11-25. https://DOI: 10.5815/ijcnis.2019.01.02

[23] Girdhar, A., & Kumar, V. (2018). Comprehensive survey of 3D image steganography techniques (vol 12, pg 1, 2018). *IET IMAGE PROCESSING*, *12*(4), 619-619.

[24] Jiang, R., Zhang, W., Hou, D., Wang, H., & Yu, N. (2018). Reversible data hiding for 3D mesh models with three-dimensional prediction-error histogram modification. *Multimedia Tools and Applications*, *77*(5), 5263-5280.

[25] Chakraborty, S., Jalal, A. S., & Bhatnagar, C. (2013). Secret image sharing using grayscale payload decomposition and irreversible image steganography. *Journal of information security and applications*, *18*(4), 180-192.

[26] Parah, S. A., Sheikh, J. A., Akhoon, J. A., & Loan, N. A. (2020). Electronic Health Record Hiding in Images for smart city applications: A computationally efficient and reversible information hiding technique for secure communication. *Future Generation Computer Systems*, *108*, 935-949.

[27] Li, T., Li, H., Hu, L., & Li, H. (2020). A Reversible Steganography Method with Statistical Features Maintained Based on the Difference Value. *IEEE Access*, *8*, 12845-12855.

[28] Koech, V. K. (2016). *Using Image Stenography Technique to Obscure Information From Unauthorized Users-a Case Study of Smartware Solutions Ltd* (Doctoral dissertation, University of Nairobi).

[29] Sasmal, M. M., & Mula, M. D. (2021, February). An Enhanced Method for Information Hiding Using LSB Steganography. In *Journal of Physics: Conference Series* (Vol. 1797, No. 1, p. 012015). IOP Publishing.

[30] Singh, S., & Siddiqui, T. J. (2018). Transform domain techniques for image steganography. In *Computer Vision: Concepts, Methodologies, Tools, and Applications* (pp. 170-186). IGI Global.

[31] Tanwar, R., & Malhotrab, S. (2017). Scope of Support Vector Machine in Steganography. In *IOP Conference Series: Materials Science and Engineering* (Vol. 225, No. 1, p. 012077). IOP Publishing.

[32] Adeboje, O. T., Gabriel, A. J., & Adetunmbi, A. O. (2020, July). *Development of an Audio Steganography System Using Discrete Cosine Transform and Spread Spectrum Techniques*. Paper presented at the International Conference on Computational Science and Its Applications (pp. 412-427). Springer, Cham. Springer. https://doi.org/10.1007/978-3-030-58817-5_31

[33] Kose, J., Chia, O. B., & Baboolal, V. (2020). Review and Test of Steganography Techniques. *arXiv preprint arXiv:2012.08460*.

[34] Subramanian, Nandhini, Omar Elharrouss, Somaya Al-Maadeed, and Ahmed Bouridane. "Image steganography: A review of the recent advances." *IEEE Access* (2021).

[35] Yuan, C., Wang, H., He, P., Luo, J., & Li, B. (2022). GAN-based image steganography for enhancing security via adversarial attack and pixel-wise deep fusion. *Multimedia Tools and Applications*, 1-21.

[36] Sahu, A. K., & Swain, G. (2020). Reversible image steganography using dual-layer LSB matching. *Sensing and Imaging*, *21*(1), 1-21.

[37] Meng, R., Cui, Q., & Yuan, C. (2018). A survey of image information hiding algorithms based on deep learning. *CMES - Computer Modeling in Engineering and Sciences*, *117*(3), 425–454. https://doi.org/10.31614/cmes.2018.04765

[38] Biswal, A. (2021). *Top 10 Deep Learning Algorithms You Should Know in 2021*. Retrieved June 16, 2021, from https://www.simplilearn.com/tutorials/deep-learning-tutorial/deep-learning-algorithm.

[39] Chavdarova, T., & Fleuret, F. (2018). Sgan: An alternative training of generative adversarial networks. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 9407-9415).

[40] Zhang, Z., Fu, G., Di, F., Li, C., & Liu, J. (2019). Generative Reversible Data Hiding by Image-to-Image Translation via GANs. *Security and Communication Networks*, *2019*. https://doi.org/10.1155/2019/4932782

[41] Kuppusamy, P. G., Ramya, K. C., Rani, S. S., Sivaram, M., & Dhasarathan, V. (2020). A novel approach based on modified cycle generative adversarial networks for image steganography. *Scalable Computing: Practice and Experience*, *21*(1), 63-72.

[42] Li, J., Niu, K., Liao, L., Wang, L., Liu, J., Lei, Y., & Zhang, M. (2020, July). A generative steganography method based on wgan-gp. In *International Conference on Artificial Intelligence and Security* (pp. 386-397). Springer, Singapore.

[43] Li, Q., Li, S., Tan, S., & Li, B. (2021, July). ThiNet Based Pruning Method for GAN Based Steganography Framework UT-GAN. In *2021 International Symposium on Signals, Circuits and Systems (ISSCS)* (pp. 1-4). IEEE.

[44] Borji, A. (2021). *Pros and Cons of GAN Evaluation Measures: New Developments*. 1–32. http://arxiv.org/abs/2103.09396

[45] Subhedar, M. S., & Mankar, V. H. (2014). ScienceDirect Current status and key issues in image steganography : A survey. *Computer Science Review*, 1–19. https://doi.org/10.1016/j.cosrev.2014.09.001.