IMAGE STEGANOGRAPHY AND STATISTICAL STEGANALYSIS

GROUP - 16

21BCY10210 - SWATI
21BCY10019 - SIDDHARTH DAYAL
21BCY10109 - RUCHI BHATTACHARJEE
21BCY10123 - HARSHITAA ASHISH

OUR TEAM

21BCY10210

SWATI

21BCY10019

SIDDHARTH DAYAL

21BCY10109

RUCHI BHATTACHARJEE

21BCY10123

HARSHITAA ASHISH

Guided by- Dr. Soma Saha

What is cryptography?

Cryptography is a technique for ensuring the privacy and security of files and communication by converting messages into an unreadable form for exchange between parties over an insecure channel.

There are 2 types of cryptography :

- Symmetric (Secret Key Cryptography)
- Asymmetric (Public Key Cryptography)

(a) Process of Encryption (At Sender Side)

(b) Process of Decryption (At Receiver Side)

Figure 1 Overview of Cryptology

Symmetric Key Cryptography

Secret key encryption, also known as symmetric-key encryption, uses a single key for both encryption and decryption. The sender encrypts the plain text message with the key and the receiver uses the same key to decrypt the encrypted message back to plain text.
Only authorized parties with knowledge of the key can perform encryption/decryption.

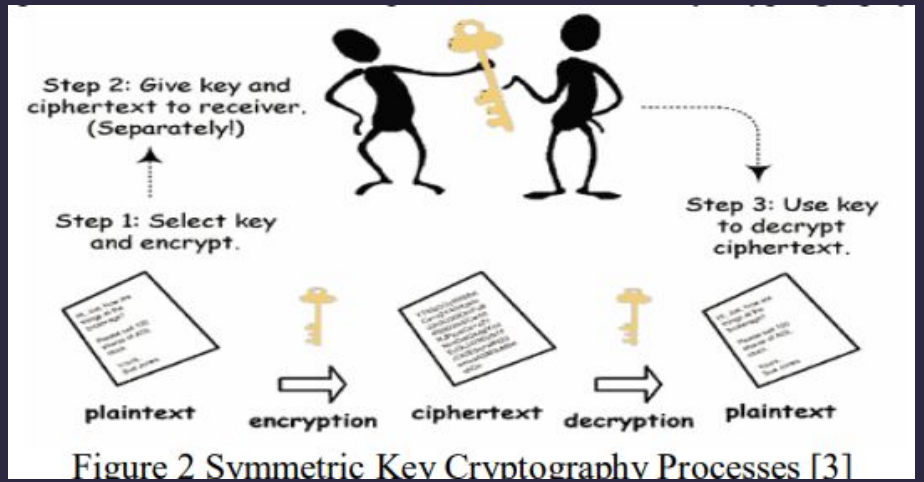


Figure 2 Symmetric Key Cryptography Processes [3]

Asymmetric Key Cryptography

Asymmetric cryptography, also known as public key cryptography, uses two mathematically related keys, one for encryption and one for decryption. The encryption key is public and the decryption key is kept secret and is known as the private key.
These keys cannot be derived from each other and both are required for the process to work.

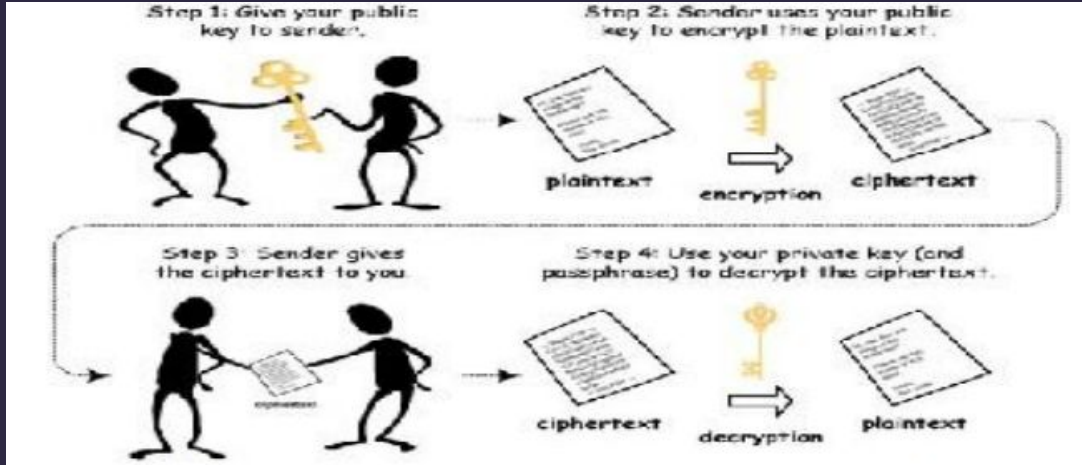


Figure 3 Symmetric Key Cryptography Processes [4]

What is STEGANOGRAPHY?

Steganography is a technique for hiding secret information within objects to conceal its existence.

It differs from cryptography, which focuses on making message contents unreadable, by aiming to keep the existence of the message hidden.

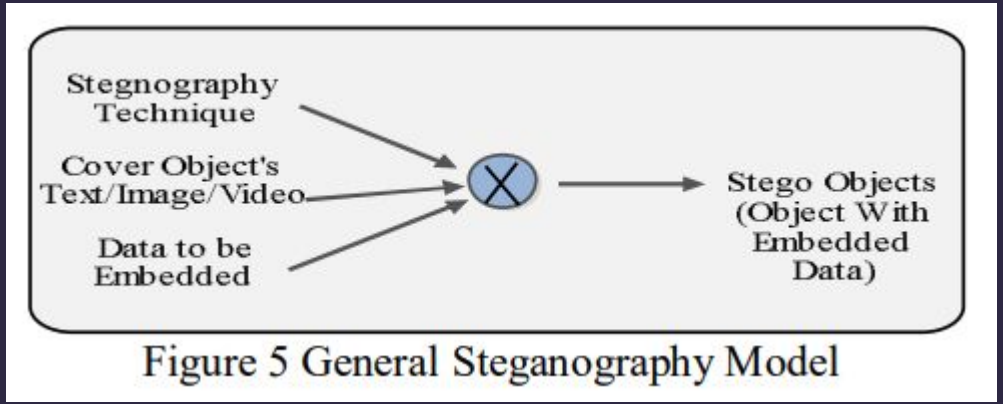


Figure 5 General Steganography Model

Types of STEGANOGRAPHY

Steganography we used in our project will be regarding image steganography.

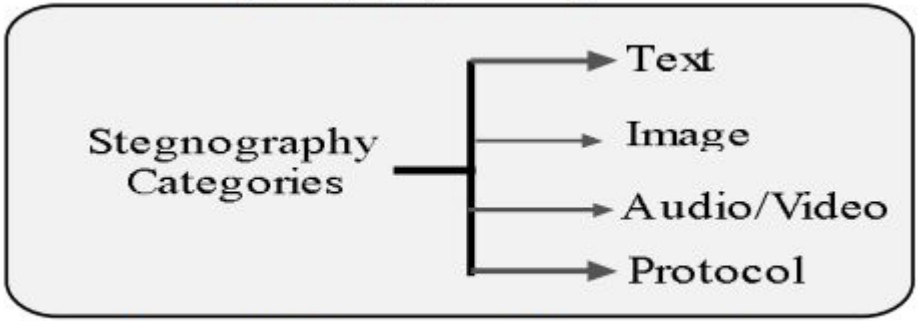


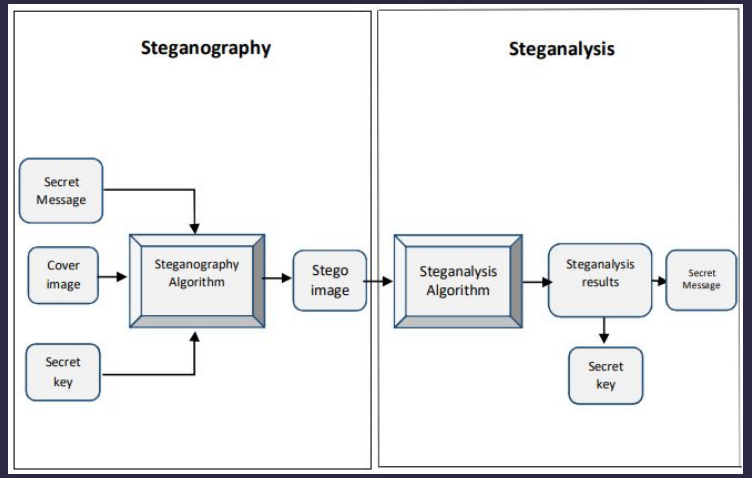
Figure 7 Categories of Steganography

PROBLEM STATEMENT

Steganography hides the very existence of a message, information can be hidden in carriers such as images, audio files, text files, videos and data transmissions.

Steganography is a technique for data hiding such that the existence of the secret message is concealed. No one is aware about the existence of the message except the intended recipients.

And ***steganalysis*** is to identify suspected information streams, determine whether or not they have hidden messages encoded into them, and, if possible, recover the hidden information.



Literature Review

Steganalysis

Steganalysis is the art of seeing the unseen; to separate stego objects and notstego-objects with practically no information about the steganography based algorithms. The objective of steganalysis is to gather any evidence about the presence of hidden data.
The documents without any hidden messages are called clean documents and the documents with hidden messages are named cover or stego documents.

Steganalysis is usually performed in one of two ways: signature analysis and blind detection. In signature analysis, the steganographic hiding method is known, which makes detection easier. Embedding algorithms always leave a particular signature, which can be tracked for detection.

Classes of Steganalysis

Targeted Steganalysis: A targeted steganalysis technique works on a specific type of steganography scheme and sometimes is limited to certain image formats.

Blind Steganalysis: A blind steganalysis technique is designed to work on all types of embedding techniques and image formats.

Quantitative Steganalysis: The quantitative steganalysis approach differs from the qualitative steganalysis in that it predicts the length of the message that has been hidden in the cover medium.

Forensic Steganalysis: the forensic steganalysis goes beyond the detection step of the classical steganalysis, obtaining the actual hidden message

Steganalysis Approaches

Visual attacks: by analyzing the images visually, when inspecting an image a compound with a known clean in the same image, to find out if there are differences .

Structural attacks: the process of embedding secret data in a cover medium may result in structural or format changes which can be detected at steganalysis stage. For example a change in compression or resolution of the cover image is an indication that the image was manipulated.

Statistical attacks: in this types of attacks the statistical analyses of the images by some mathematical formulas is applied and the detection of hidden data is, based on these statistical results

STEGANOGRAPHY

The science and art of covert communication.

Study of uncovering the steganographic process

STEGANALYSIS

IMAGE FORMATS

BMP image

The bitmap or BMP format is considered a simple image file format. BMP files are device-independent files most frequently used in Windows systems, and it is based on the RGB color model. Header region contains information and other details about size and color depth. Data region contains the values of each pixel. Files in the BMP format can be single channel or three channels color or grayscale. The bmp format allows for lossless compression but it is most often used with uncompressed images

Tiff image

TIFF (Tag Image File Format) is a common format for exchanging raster graphics (bitmap) images between applications programs, A TIFF file can be identified as a file with a ".tiff" or ".tif" file name suffix. TIFF format supports RGB, indexed color, and grayscale images with alpha channels and bitmap mode images without alpha channels. TIFF is a flexible bitmap image format supported by all paint, page layout, and image editing. TIFF documents have a maximum file size of 4 GB. TIFF image format allows for lossless compression.

JPEG image

Joint Photographic Experts Group (JPEG) format is commonly used to display photographs in HTML documents. JPEG format supports RGB, and grayscale color modes, and does not support transparency. The JPEG format retains all color information in an RGB image but compresses file size by selectively discarding data. A JPEG format is a commonly used method of lossy compression for digital images. A JPEG file is created by choosing a range of compression qualities. When a JPEG image is converted from another format to JPEG, image quality is required to be specified.

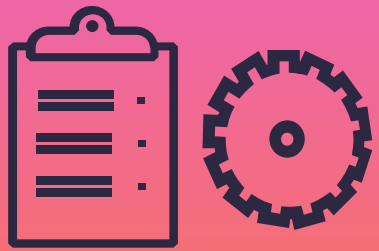
PNG image

Portable Network Graphics (PNG) format is a raster graphics file format that supports lossless data compression, it is expected to replace the Graphics Interchange Format (GIF) that is widely used on today on the Internet. PNG format supports RGB, grayscale, indexed color and bitmap mode images. 17 PNG preserves transparency in grayscale and RGB images. The PNG format was developed by an Internet commission expressly to be patent-free. PNG supports 24-bit images and produces background transparency without jagged edges; however, some web browsers do not support it.

Reasons for Choosing Steganalysis of Images

Image is the most available type of cover to hide a secret message over the internet. Also they can be used as carrier objects without raising much suspicion. Image files have a lot of capacity redundancy, which provides space for embedding. Therefore due to the wide use of images in information hiding, research work in steganalysis have addressed the problem of detecting hidden data inside various types of images.

METHODOLOGY APPROACH



We follow an experimental approach for achieving the objectives. Relevant data about secret and cover images will be analyzed as necessary to enhance the detection performance of the proposed model.

STATISTICAL FEATURES SELECTION

Feature Name	Feature Description
CC-LR	Correlation coefficient between LHB and RHB
CV-B	Coefficient of variation of full bytes
CV-R	Coefficient of variation of RHB
GLCM-B	Contrast, Correlation, Homogeneity Energy, of full bytes
GLCM-R	Contrast, Correlation, Homogeneity, Energy, of RHB
GLCM-3LSB	Contrast, Correlation, Homogeneity, Energy, of 3LSB
GLCM-4LSB	Contrast, Correlation, Homogeneity, Energy, of 4LSB
Entropy-B	Entropy of full bytes
Entropy-R	Entropy of RHB
Diff-R	Average of absolute difference between successive right half bytes
Skew-B	Skewness of full bytes
Skew-R	Skewness of RHB

OBJECTIVE

To realize the objectives of the proposed model, three processes are required:

1. Steganography, which involves embedding of secret data inside cover images
2. Feature extraction from clean and stego images
3. Training and testing of a classifier based on the selected features and the classifier

REQUIRED FUNCTIONALITIES

Steganography modules to embed a secret file inside an RGB image

Feature extraction from a batch of clean and stego images

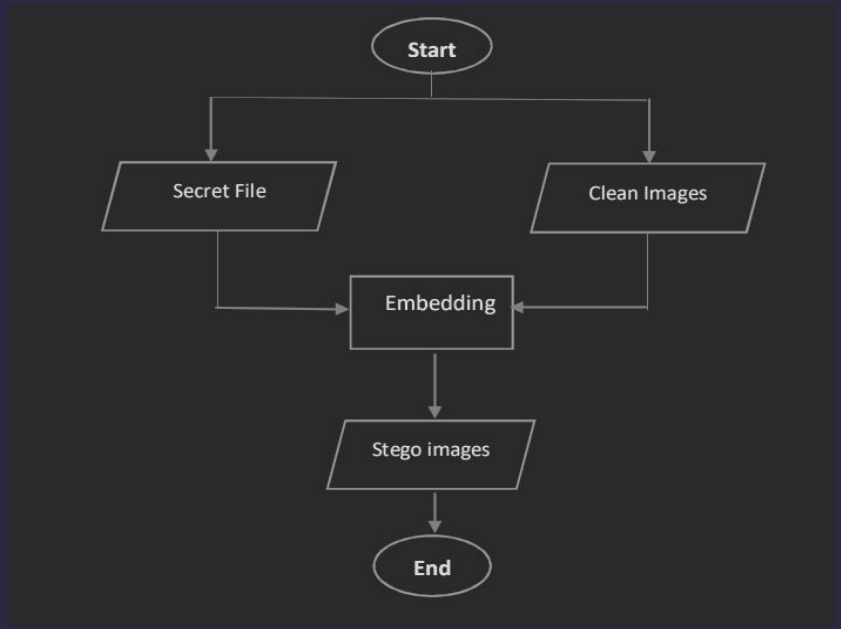
Batch classification of a group of test images against a training set

Single image classification, against a training set

PROPOSED SYSTEM

The implementation of the proposed system which includes the required functionalities

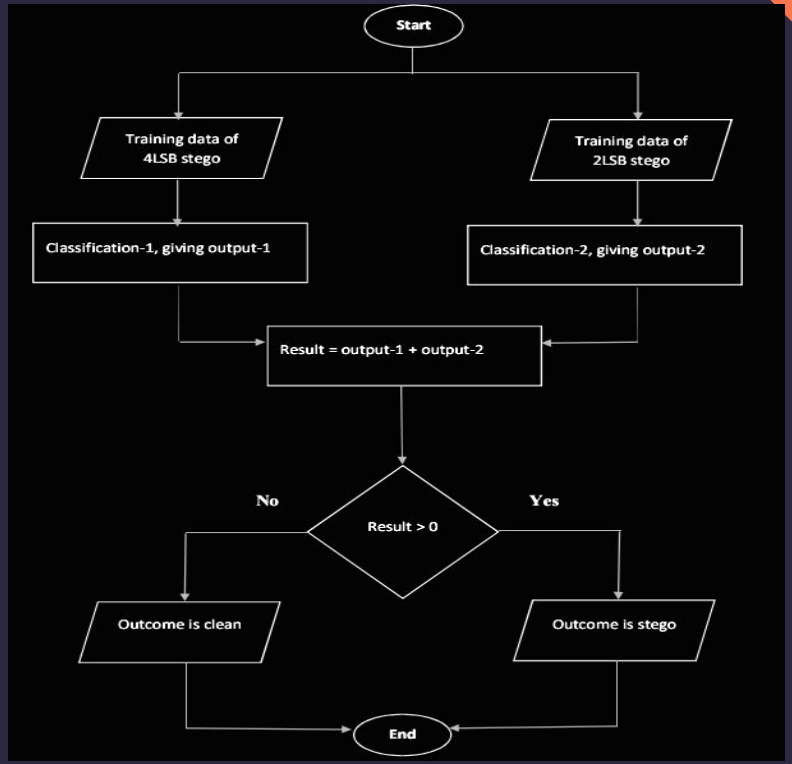
PHASE-1: EMBEDDING



PHASE-2: FEATURE EXTRACTION

$$CV = \frac{\text{Standard Deviation (n)}}{\text{Mean (n)}}$$

PHASE-3: SINGLE IMAGE CLASSIFICATION



PHASE-4: BATCH IMAGE CLASSIFICATION

In this phase, a batch of testing images are classified as in the single image classification phase. The feature set data of individual testing images are processed independently, giving the outcome for each image. The purpose of this phase is to simplify the process of classifying a large number of images.

EVALUATION METRICS

True Negative Rate (TN) : The ratio of true negative detections to the number of clean images.

True Positive Rate (TP) :The ratio of true positive detections to the number of stego images.

False Negative Rate (FN) : The ratio of false negative detection to the number of stego images.

False Positive Rate (FP) : The ratio of false positive detection to the number of clean images.

Detection Accuracy The ratio of correctly detected clean and stego images to the total number of clean and stego images represent the detection accuracy

Accuracy = (TN + TP) / (TN + TP + FP +FN)

FUTURE SCOPE OF STEGANOGRAPHY & STEGANALYSIS

In the near future, the most important use of steganographic techniques will probably be lying in the field of digital watermarking. Content providers are eager to protect their copyrighted works against illegal distribution and digital watermarks provide a way of tracking the owners of these materials. Steganography might also become limited under laws, since governments already claimed that criminals use these techniques to communicate.

- Hiding data on the network in case of a breach.
- Peer-to-peer private communications.
- Posting secret communications on the Web to avoid transmission.
- Embedding corrective audio or image data in case corrosion occurs from a poor connection or transmission.

THANK YOU