Statistical Steganalysis Detector Model for

8-bit Depth Images

<div dir="rtl">

نموذج إحصائي لكشف الإخفاء في الصور ذات عمق 8-بت

</div>

Prepared by

Zaid Hadi Al-Taie

Supervisor

Dr. Mudhafar Al-Jarrah

Thesis Submitted in Partial Fulfillment of the Requirements of the Degree of Master of Computer Science

Department of Computer Science

Faculty of Information Technology

Middle East University

January- 2017

AUTHORIZATION STATEMENT

I, Zaid Hadi Al-Taie, authorize the Middle East University to provide hard copies or soft copies of my thesis to libraries, institutions or individuals upon their request.

Name: Zaid Haid Al-Taie

Date: 6/2/2017

Signature:

تفويض

انا زيد هادي الطائي ، افوض جامعة الشرق الاوسط للدراسات العليا بتزويد نسخ من رسالتي للمكتبات المعنية والمؤسسات و الهيئات عند طلبها.

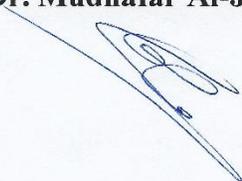الاسم : زيد هادي الطائي

التاريخ : 2017/2/6

التوقيع :

Examination Committee Decision

This is to certify that the thesis entitled "Statistical Steganalysis Detector Model for 8-bit Depth Images" was successfully defended and approved on 17/1/2017

Examination Committee Members Signature
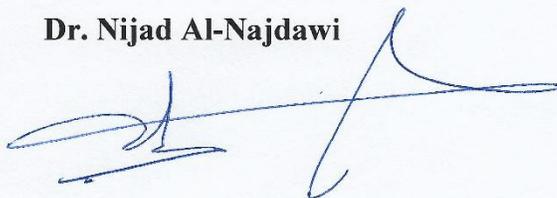
(Supervisor)

Dr. Mudhafar Al-Jarrah

(Head of the Committee and Internal Committee Members)

Dr. Abdelrahman Abuarqoub

(External Committee Members)

Dr. Nijad Al-Najdawi

قرار لجنة المناقشة

نوقشت هذه الرسالة وعنوانها "نموذج احصائي لكشف الاخفاء في الصور ذات عمق 8-بت"
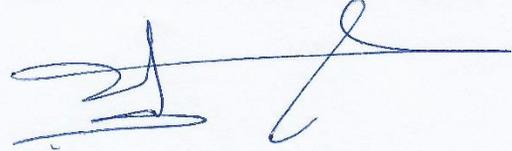واجيزت بتاريخ 2017/1/17

اعضاء لجنة المناقشة

الدكتور : مظفر الجراح (مشرفا)

الدكتور : عبد الرحمن ابو عرقوب (رئيسا)

الدكتور : نجاد النجداوي (عضوا خارجيا)

ACKNOWLEDGMENTS

I would like to thank God for helping me to achieve this study and proceed successfully.

I would like to express my great appreciation to my supervisor Dr. Mudhafar Al-Jarrah for his countless help, constructive suggestion and guidance in every step of my research. I have benefited a lot from him since the beginning of this research.

I would like to express my gratitude to my family for providing me with continuous support throughout the years of my study.

Dedication

(وقل رب زدني علمًا) طه الاية (111)

This thesis is dedicate to my Mother, Father, Sister and Brother God's mercy.

Table of Contents

List of Tables

List of Figure

List of Acronym

RGB	Red Green Blue
LSB	Least Significant Bit
MSB	Most Significant Bit
SVM	Support Vector Machine
GLCM	Gray Level Co-occurrence Matrix
CGCM	Colors Gradient Co-occurrence Matrix
MLP	Multilayer Perceptron
LHB	Left Half Byte
RHB	Right Half Byte
ESS	Experimental Steganalysis System
BPP	Bit Per Pixel
TNR	True Negative Rate
TPR	True Positive Rate
FNR	False Negative Rate
FPR	False Positive Rate
2LSB	Two Least Significant Bit
3LSB	Three Least Significant Bit
4LSB	Four Least Significant Bit
CV	Coefficient of Variation
DA	Discriminant Analysis

Statistical Steganalysis Detector Model for

8-bit Depth Images

By

Zaid Hadi Al-Taie

Supervisor

Dr. Mudhafar Al-Jarrah

Abstract

This thesis aims to develop a statistical model for steganalysis to enhance the detection of the existence of hidden data inside 8-bit depth gray-scale BMP images. The proposed model is based on enhancing the image texture features through analyzing both full-bytes and parts of bytes of an image. It is known that most steganography techniques embed the bits of a secret message within the right half of a cover image's bytes, the least significant half of a byte, to avoid obvious visual distortion. Therefore, the focus of the steganalysis process in this work is the right half part of each byte of an image under investigation. The selected feature set is based on the gray level co-occurrence model, including contrast, homogeneity, correlation, and energy. Additional features include: entropy, coefficient of variation of the image's right half-bytes, correlation coefficient between left and right half-bytes, and the average of difference between the intensity of right half -bytes in successive pixels.

The work involved implementation of the proposed model in MATLAB, which consisted of modules for feature extraction, training and testing using the two-category discriminant analysis classifier, and the batch classification of a set of test images. Testing of the detection accuracy of the proposed model was carried out in three stages. First, a dataset of 180 mixed-source images were analyzed using 3-fold cross validation. In the second testing stage, the 3-fold cross validation was applied using a dataset of 1500 images from a single public dataset. The third stage realized a large-scale field test of the proposed model, using a public dataset of 5000 images for testing and 1500 images for training. The training dataset was independent from the testing dataset, it was randomly selected from another part of the dataset that was not included in the testing dataset. The steganalyzer combined two training feature datasets, to deal with the two embedding methods that were used to generate the stego images. The average of the detection accuracy ranged from 97.50% to 98.73% in the validation test and 97.82% to 98.28% in the field test.

Keywords: Steganography, Statistical Steganalysis, LSB, Discriminant Analysis, Classifier, Feature Extraction, Detection Accuracy

نموذج إحصائي لكشف الإخفاء في الصور ذات عمق 8 بت

إعداد

زيد هادي الطائي

إشراف

الدكتور مظفر الجراح

ملخص تنفيذي

تهدف هذه الرسالة إلى تطوير نموذج إحصائي تحليلي بهدف تحسين عملية كشف إخفاء بيانات داخل ملفات صور بصيغة بي. إم. بي (BMP) ذات عمق 8 بت وتدرج رمادي. ويستند النموذج المقترح على تعزيز تحليل تركيب الصورة من خلال تحليل كامل للبايت بالإضافة الى تحليل أجزاء من بايت تلك الصورة. ومن المعروف أن معظم تقنيات إخفاء المعلومات تعتمد على تضمين وحدات البت الخاصة بالرسالة السرية داخل النصف الأيمن من البايت الخاص بالصورة المستخدمة كغطاء وذلك لتجنب إحداث تشوه بصري واضح، لأن هذا النصف من البايت يعد اقل وزنا. وعليه فإن التحليل الإحصائي في هذا البحث يركز على النصف الأيمن من كل بايت من وحدات بايت الصورة قيد التحقيق. وتعتمد مجموعة الخصائص المختارة على مستوى توارد التدرج الرمادي في النموذج، بما في ذلك مستويات التباين، والتجانس، والارتباط، والطاقة. وتتضمن الميزات الإضافية: العشوائية ومعامل الاختلاف في وحدات أنصاف البايت للصورة، ومعامل الارتباط بين وحدات بايت النصف الأيمن والنصف الأيسر، ومتوسط الفرق في كثافة وحدات البايت في البكسلات المتعاقبة.

وقد تطلب العمل تنفيذ النموذج المقترح في برنامج ماتلاب (MATLAB)، واشتمل على: استخراج الخصائص بالصور المستخدمة للغرض التدريب والاختبار، وتوليد مجموعات من بيانات تلك الخصائص، وتصنيف الكمية المعدة من صور الاختبار باستخدام مصنف تحليلي تمايزي ذي صنفين. وقد أجري اختبار دقة أداء الكشف على ثلاث مراحل. الاولى، مجموعة صور من مصادر مختلفة عددها 180 صورة تم تحليل خصائصها باستخدام تدقيق ثلاثي. الثانية، استخدمت طريقة التحقيق الثلاثي على مجموعى من 1500 صورة. الثالثة، تم إخضاع النموذج المقترح إلى اختبار مستقل وواسع النطاق على مجموعة صور مكونة من 5000 صورة لاختبار مدى دقة الكشف. وكانت مجموعة بيانات التدريب مستقلة عن مجموعة بيانات الاختبار، وقد اختيرت بشكل عشوائي من جزء آخر من مجموعة البيانات التي لم تكن مدرجة في بيانات الاختبار. وقد جمع نموذج تحليل الاخفاء المقترح بين نموذجين من مجموعة الملامح ومجموعتين من بيانات التدريب للتعامل مع أسلوبين من أساليب تضمين البيانات. وقد تراوح متوسط دقة الكشف بين 97.50% الى 98.73% وتراوحت دقة التحقق من 97.82% الى 98.28%.

الكلمات المفتاحية: نموذج إحصائي، كشف الإخفاء، عمق 8 بت.

Chapter 1
Introduction

1.1 Background

Recently, there was a big increase in the use of social media and information technology over the Internet and mobile networks. Users are exchanging a large amount of documents for business and leisure. Many of the documents that are exchanged over the internet are being used as covers for concealing secret documents of legal and illegal pursues. The hiding of secret document for good purposes such as to avoid hackers and criminals is ethically acceptable. However, information hiding is also being used for illegal purposes by criminals, terrorists and in insider's threats, which needs to be monitored and detected. This study deals with steganalysis to detect the presence of embedded data inside images.

Steganography: The art of hiding the very presence of communication by embedding secret messages into innocuous-looking cover documents, such as digital images (Anderson & Petitcolas, 1998). Detection of steganography, estimation of message length, and its extraction belongs to the field of steganalysis. Figure 1.1 shows the Steganography and Steganalysis diagram.
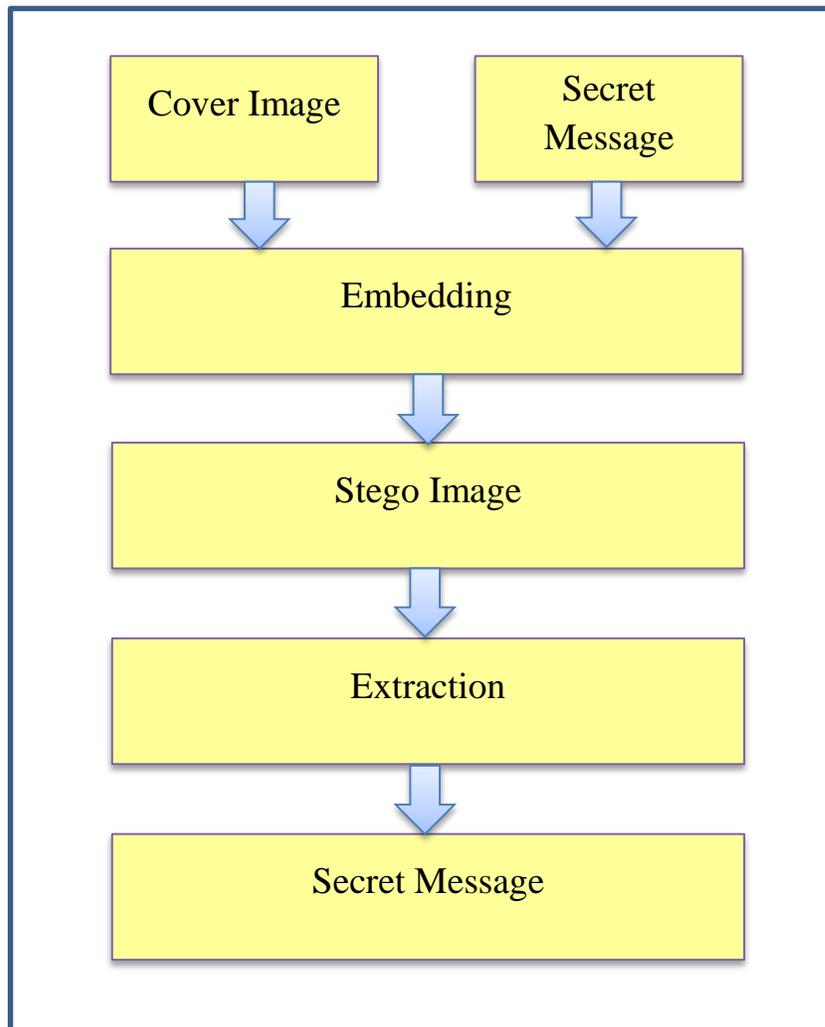
Figure 1.1 Steps in Steganography.

Steganalysis: Recently steganalysis received a great deal of attention both from law enforcement, the media, and the multimedia research fields. Steganalysis is fundamentally a problem of classifying samples as either simple covers or stego-objects. Besides that, steganalysis is the discovery of the existence of hidden information; therefore, like

cryptography and cryptanalysis, the goal of steganalysis is to discover hidden information and to break the security of its carriers. (Codr, 2009).
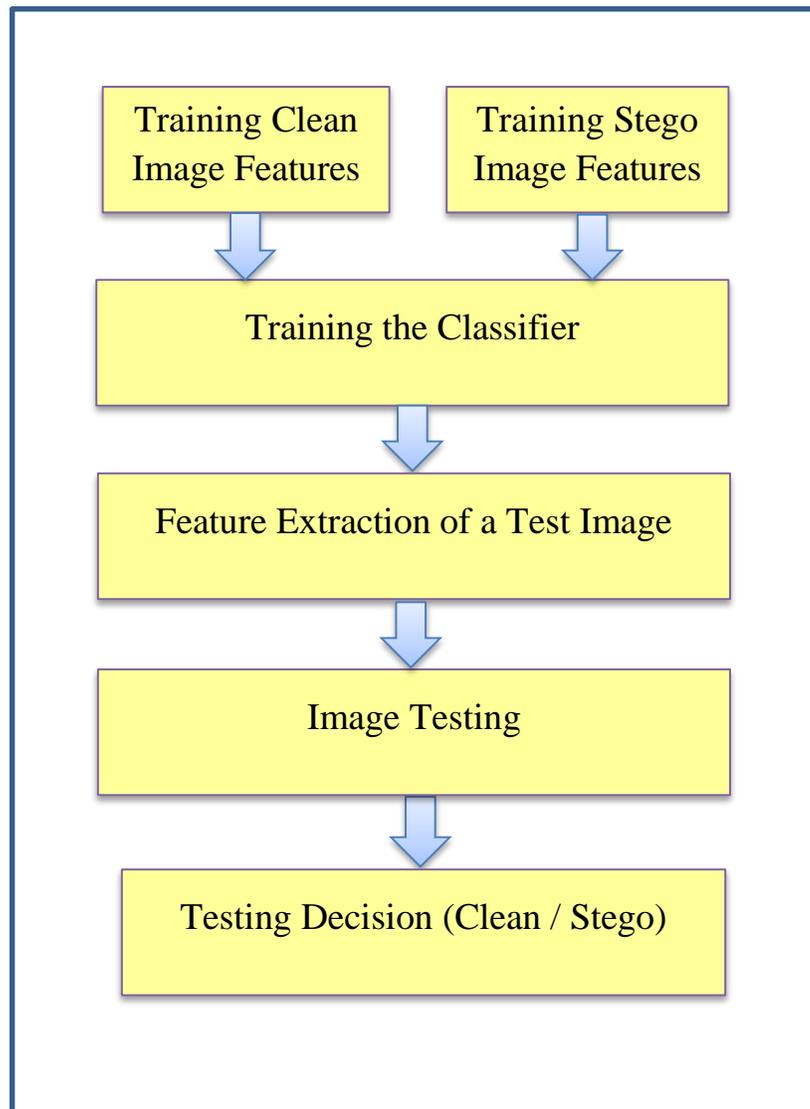
Figure 1.2 Steganalysis Steps

1.2 Problem Statement

The problem addressed in this study is the detection of the existence of hidden messages inside 8-bit images (grey-scale BMP format). The choice of 8-bit images is twofold: it is widely used in steganalysis research and it is simpler to use this single channel format for evaluation of a model before the solution is generalized for color images.

Machine learning approach will be used which is considered an essential method for detecting the presence of a certain pattern of hidden messages inside cover images.

The proposed work in this thesis is to present a scheme to detect the existence of a message inside gray-scale image. Although there are quite a lot of researches in the steganalysis field, we believe that more experimental work is needed to guide us in evaluating and selecting more effective steganalysis methods and features.

1.3 Aim and List of Objectives

The main aim of this research is to enhance the detection performance of the steganalysis method by introducing a new statistical model that will be able to detect the presence of hidden data in cover images. This approach is based on extracting image features from a dataset of clean and stego images. In this approach, a classifier detector will be trained on the extracted features of the dataset to be able to detect a stego image based on the profile data of clean and stego images learned during the training phase.

1.4 Motivation

The growth of using multimedia and transmission of data over different communication networks increases the need of security such as steganography and cryptography. On the other hand, this also increases the need to discover the hidden data based on these images of transmitted data over networks this field is the steganalysis. Both Steganography and Cryptography are excellent means by which to accomplish this but neither technology alone is perfect and both can be broken. It is for this intention that most experts would suggest using both to add multiple layers of security. The watermarking and fingerprinting are other technologies that are closely related to steganography (Nagaraj, 2013).

1.5 Research Questions

In this research work we will try to answer the following questions:

1. Is it possible to develop a steganalysis model that can detect hidden messages (without extracting it) using the statistical approach?
2. What are the features that will be used in the detection process?
3. What are the results of the detection accuracy when using an implementation of the proposed model on a set of stego/clean images?

1.6 Significance of Work

The outcome of this research is expected to enhance the detection capability of steganalysis methods in uncovering the existence of hidden secret messages inside cover images, which should help authorities and business management to detect bad intentioned hiding of secret messages. It is envisaged that the outcome of enhancing the steganalysis in 8-bit depth images can be generalized to images of other formats.

Chapter 2
Review of Literature and Related Studies

2.1 Steganography Concept

Steganography, like cryptology, is intended to add a layer of security to communications so that unauthorized users don't know what embedded in an image. However, unlike cryptology, steganography is not meant to obscure the message, but to obscure the fact that there is a message at all. Attacks against cryptography take what is known to be an encrypted message and attempt to decrypt the message. Attacks against steganography take what seems to be an ordinary image, text, multimedia file, or other document and determine whether or not there is another message hidden within (Codr, 2009).

Steganography and cryptography are strongest when combined. A message sent in secret (steganography) in an encrypted form (cryptography) is much more secure than a "plain text" message sent by secret means or a clearly sent encrypted message. There are some cases in which steganography can take the place of cryptography; for instance, German bans on encrypting radio communications were recently countered by applying steganography to radio communications (Westfield, 2006). Besides that, digital images are the most widespread carrier medium (Westfield, 1999) Figure 2.1 shows the stenographic system. Generally, however, steganography is not intended to replace cryptography (Johnson, 1995) but to provide an alternative method of data protection that does not invite attacks.

Steganography, watermarking and cryptography are interlinked techniques, they share the same aim of protecting secret information that is transmitted over communication channels. The first two methods are quite similar as they both embed data inside images. Figure 2.2 and Table 2.1 illustrate the different embodiment disciplines of information hiding (Cheddad, Condell, Curran & McKevitt, 2010).
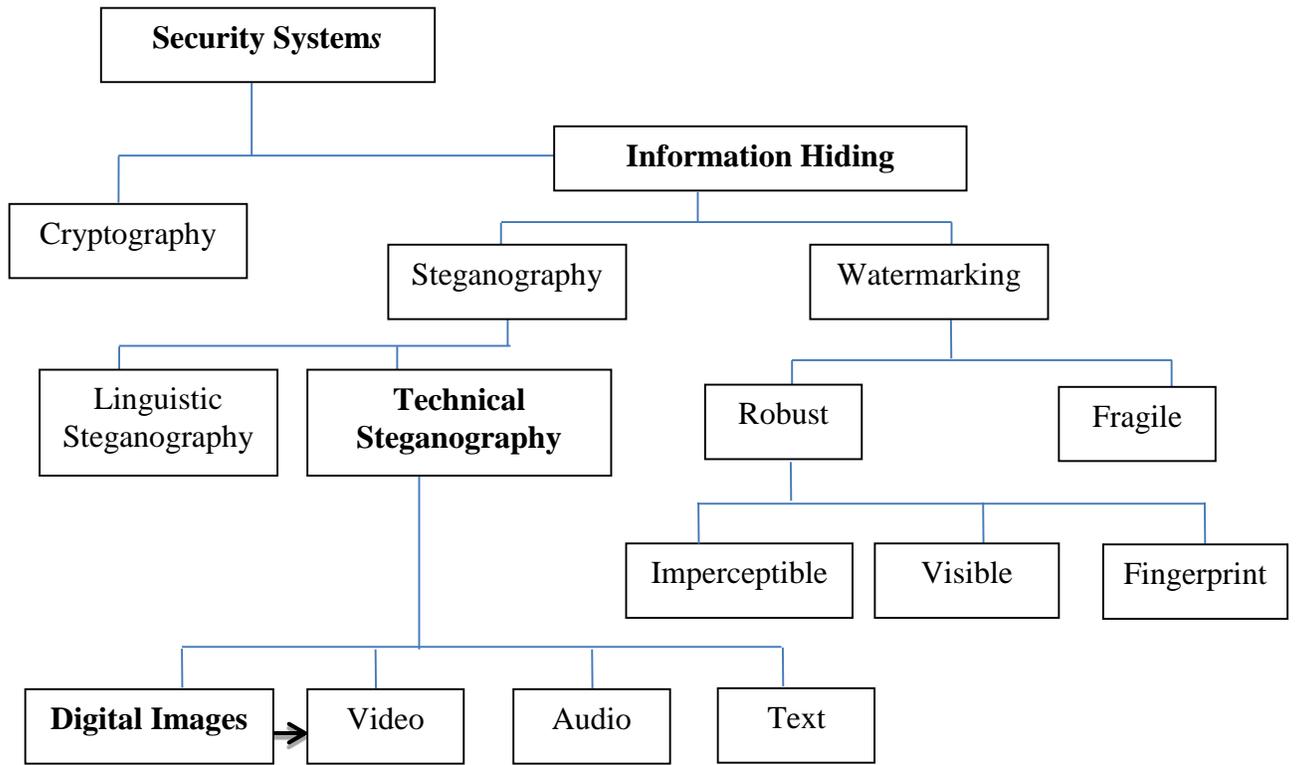
Figure 2.1 Different embodiment disciplines of information hiding. (Cheddad, et. al, 2010)

Table 2.1 Comparison of different embodiment disciplines of information hiding (Cheddad,

et. al, 2010)

Criterion/Method	Steganography	Watermarking	Encryption
Carrier	Any digital media	Mostly image/audio files	Usually text-based with some extensions to image files
Secret data	Payload	Watermark	Plain text
Key	Optional		Necessary
Input files	At least two unless in self – embedding		One
Detection	Blind	Usually informative (i.e. original cover or watermark is needed for recovery)	Blind
Authentication	Full retrieval of data	Usually achieved by cross correlation	Full retrieval of data
Objective	Secret communication	Copyright preserving	Data protection
Result	Stego-file	Watermarked-file	Cipher- text
Concern	Delectability capacity	Robustness	robustness
Type of attack	Steganalysis	Image processing	Cryptanalysis
Visibility	Never	Sometimes	Always
Fails when	It is detected	It is removed replaced	De-ciphered
Relation to cover	Not necessarily related to the cover. The message is more important than the cover	Usually becomes an attribute of the cover image. The cover is more important than the message	N/A
Flexibility	Free choose any suitable cover	Cover choice is restricted	N/A
History	Very ancient except its digital version	Modern era	Modern era

The steganography method embeds secret data in a cover image by replacing bits from the cover image with bits of the secret data in such a way that the modification of the cover image will not be visually perceptible. The embedding process may depend on a secret stego key. The stego key is used to control the embedding process, such as the selection of pixels location within the cover image for embedding. Also, cryptography can be combined with steganography to strengthen the security of hidden data.

2.1.1 Types of Steganography Techniques

Almost all digital file formats can be used for steganography, but the formats that are more suitable are those with a high degree of redundancy. Redundancy can be defined as the bits of an object that provide accuracy far greater than necessary for the object's use and display. The redundant bits of an object are those bits that can be altered without the alteration being detected easily. Image and audio files especially comply with this requirement, while research has also uncovered other file formats that can be used for information hiding. Figure 2.3 shows the four main categories of file formats that can be used for steganography. Morkel, Eloff, and Olivier, (2005).

Types of Steganography:

| Text | Images | Audio/Video | Protocol |
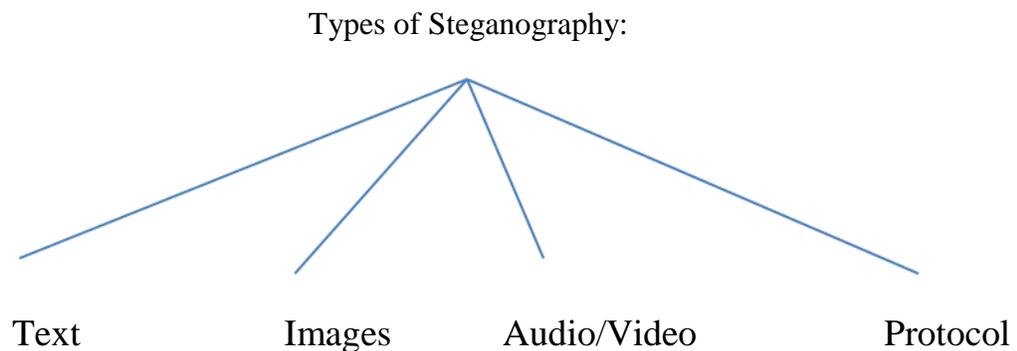
Figure 2.2 Main categories of file formats that can be used for steganography

Text Steganography: Hiding text information in other text files is historically the earliest method of steganography. An obvious text-in-text steganography method was to hide a secret message in every letter of every word of a text message. Text steganography to embed digital files is not used very often since text files have a very small amount of redundant data. It is only since the beginning of the Internet era that other digital file formats were used as cover media instead of text files.

Image Steganography: Images are the most popular cover objects used for steganography. In the domain of digital images many different image file formats exist, most of them for specific applications. For these different image file formats, different steganographic algorithms exist. To a computer, an image is a collection of numbers that constitute different light intensities in different areas of the image. This numeric representation forms a grid and the individual points are referred to as pixels. Most images on the Internet consists of a rectangular map of the image's pixels (represented as bits) where each pixel is located and its color. These pixels are displayed horizontally row by row Morkel, et al (2005). The secret message is embedded in a digital image through an embedding algorithm using a confidential key. The receiver must have the same key and use the extracting algorithm to extract the message.

Audio and Video Steganography: Techniques used to transmit hidden information by modifying an audio or video signal in an imperceptible manner. The secret message is embedded into digitized signal that results from slight altering of the binary sequence of the original file.

Regardless of the type of steganography or the media used in steganography technique, it represents the science of hiding some secret text or audio information in a host message. The host message before steganography and stego message after steganography has the same characteristics.

2.2 Steganalysis Concept

The security of a steganographic system is defined by its strength to defeat detection. The effort to detect the presence of steganography is called *steganalysis*.

The steganalyst (i.e., the warden in Simmons' anecdote) is assumed to control the transmission channel and watch out for suspicious material. A steganalysis method is considered as successful, and the respective steganographic system as 'broken', if the steganalyst's decision problem can be solved with higher probability than random guessing. (Principles of Modern Steganography and Steganalysis).

Steganalysis is the art and science of detecting secret messages hidden using steganography Fridrich, Goljan, and Du (2001), Johnson and Jajodia (1998). The goal of steganalysis is to collect sufficient evidence about the presence of embedded message and to break the security of its carrier. Steganalysis finds its use in computer forensics, cyber warfare, tracking the criminal activities over the internet and gathering evidence for investigations particularly in the case of anti-social elements (Fridrich, et al., 2001). Apart from its law enforcement and anti-social significance steganalysis also has a peaceful application—improving the security of steganographic tools by evaluating and identifying their weaknesses.

Different robust steganalysis techniques have been proposed in the literature. In this research, a various proposed approaches discussed and classifying in the next chapter. Different steganalysis techniques are shown in Figure 2.4 Arooj and Mir (2010).

```
                          ┌──────────────────┐
                          │   STEGANALYSIS   │
                          └──────────────────┘
                   ┌──────────────┴──────────────┐
            ┌─────────────┐               ┌─────────────┐
            │  SIGNATURE  │               │ STATISTICAL │
            └─────────────┘               └─────────────┘
          ┌───────┴───────┐           ┌────────┴────────┐
    ┌──────────┐  ┌───────────┐  ┌──────────┐  ┌───────────┐
    │ SPECIFIC │  │ UNIVERSAL │  │ SPECIFIC │  │ UNIVERSAL │
    └──────────┘  └───────────┘  └──────────┘  └───────────┘
```

LSB Embedding	LSB Matching	Spread Spectrum	BPCS

JPEG Compression	Transform Domain	Additive Noise	Others
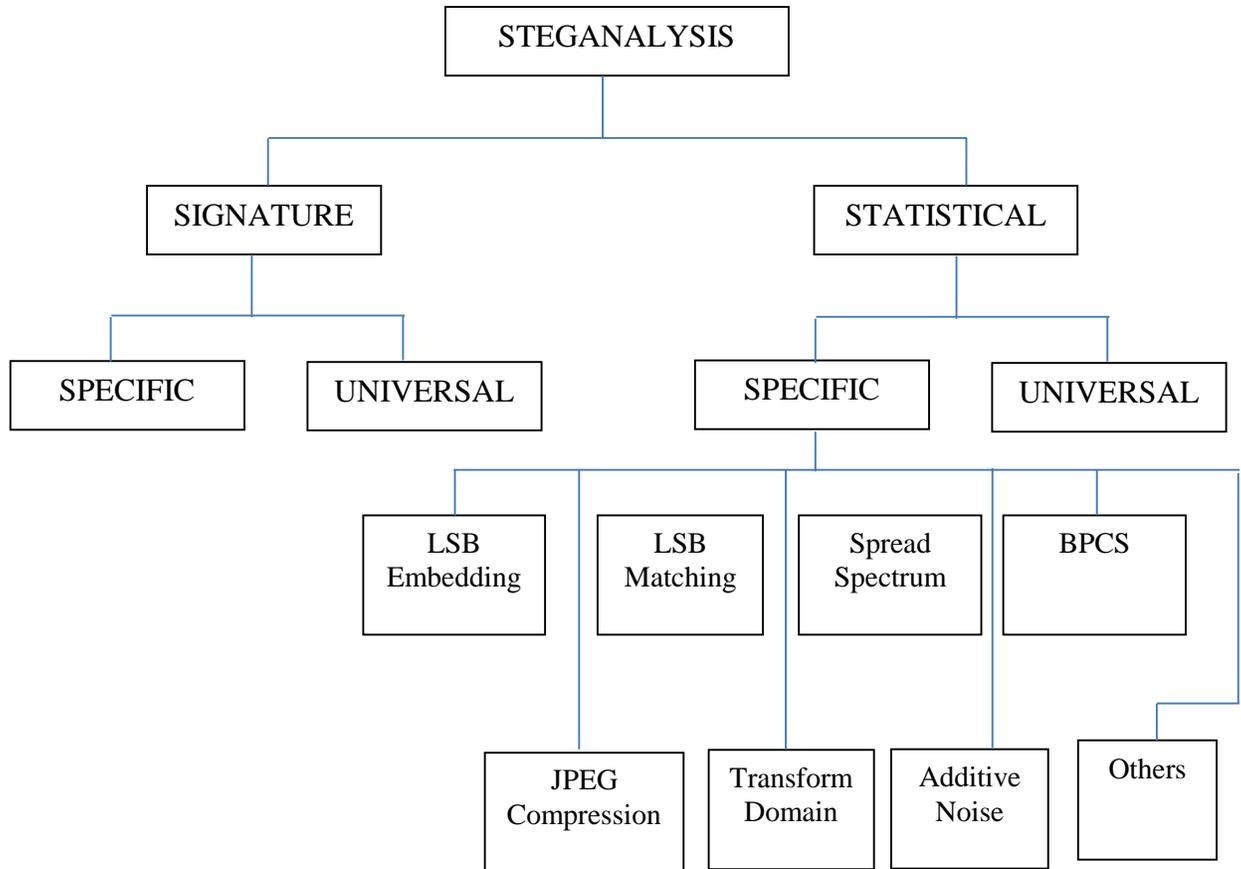
Figure 2.3 Different steganalysis techniques (Arooj & Mir, 2010)

In other words, steganalysis is the practice of attacking steganography methods for the detection, extraction, destruction and manipulation of the hidden data in a stego object. Attacks can be of several types for example, some attacks merely detect the presence of hidden data, some try to detect and extract the hidden data, some just try to destroy the hidden data by finding the existence without trying to extract hidden data and some try to

replace hidden data with other data by finding the exact location where the data is hidden. Steganographic attacks consist of detecting, extracting and destroying the hidden object of the stego media. Detection is generally carried out by identifying some characteristic feature of images that is altered by the hidden data. A good steganalyst must be aware of the methods and techniques of the steganography tools to efficiently attack.

2.3 Classification of Attacks

Johnson and Jajodia (1998) classified attacks in six main categories as in the following:

- Stego only attack: only stego object is available for analysis.

- Known cover attack: both cover and stego are known.

- Known message attack: in some cases message is known and analyzing the stego object pattern for this embedded message may help to attack similar systems.

- Chosen stego attack: steganographic algorithm and stego object are known.

- Chosen-message attack: here steganalyst creates some sample stego objects from many steganographic tools for a chosen message and analyses these stego objects with the suspected one and tries to find the algorithm used.

- Known stego attack: the steganography algorithm is known and both the original and stego-image are available (Johnson & Jajodia, 1998).

2.4 Different Approaches of Steganalysis

The different approaches of steganalysis are illustrated below:

Visual attacks: By analyzing the images visually, like considering the bit images and try to find the difference visually in these single bit images.

Structural attacks: The format of data file often changes as the data to be hidden is embedded, identifying these characteristic structural changes can detect the existence of image, for example in palette based steganography the palette of image is changed before embedding data to reduce the number of colors so that the adjacent pixel color difference should be very less. This shows that groups of pixels in a palette have the same color which is not the case in normal images.

Statistical attacks: In these types of attacks the statistical analyses of the images by some mathematical formulas is done and the detection of hidden data is done based on these statistical results. Generally, the hidden message is more random than the original data of the image thus finding the formulae to know the randomness reveals the existence of data.

More about machine learning approaches in steganalysis are discussed in detail in the book by Hans Georg Schaathun (2012)

2.5 Previous Studies

Detection of steganography, estimation of message length, and its extraction belong to the field of steganalysis. Some definitions and several methods of steganalysis were proposed in the literature Johnson and Jajodia (1998), Farid (2001), Aveibas, Memon, and

Sankur, (2003), Chandramouli and Memon (2001), Fridrich, Du and Meng (2000), Fridrich, et al., (2001), Jajodia (1998) and Voloshynovskiy (2002).

In Johnson and Jajodia (1998) study, the authors give an overview of some characteristics to detect the existence of hidden information. And the same authors in another paper they give a good description of popular free software's steganalysis.

Farid (2001) tells a steganalysis method based fisher linear classifier. Aveibas, et al (2003) take the regress analysis to analysis image based image metrics. And for LSB embedding methods, the most successful researchers that were published many papers is Fridrich, et al., (2000). And with the help of steganalysis, ones can find more robustness methods to resist attack and analysis Provos (2001).

Guillon, Furon, and Duhamel (2002) proposed a framework for steganalysis of SCS by modeling QIM steganography as an additive noise channel. Sullivan, Madhow, Chandrasekaran, and Manjunath, (2004) proposed a steganalysis scheme for QIM steganography using supervised learning. Zou, Shi, Su, & Xuan, (2006) proposed a supervised learning based steganalysis method which uses 2-dimensional Markov chain based framework to capture traces of message embedding. The proposed scheme in Zou et al. (2006) uses local neighborhood of the current pixel to predict pixel values. The prediction-error image is then generated by subtracting the predicted value from the actual pixel value and then comparing the difference of the two datasets. Feature vector consisting of three empirical transition matrices of Markov chain along the horizontal, vertical, and diagonal directions is used to train a binary support vector machines (SVM)-based classier.

Wang and Gong, (2012) proposed a steganalysis algorithm based on colors-gradient co-occurrence matrix (CGCM) for GIF images. CGCM structured with colors matrix and gradient matrix of the GIF image, and 27- dimensional statistical features of CGCM, which are sensitive to the color correlation between adjacent pixels and the break in texture, are extracted. This proposed steganalysis algorithm does not require a lot of computing time.

Arvis, Debain, Berducat and Benassi (2004) has proposed a multispectral method considering the correlations between the color bands. To study the efficiency of their method, they tested it in a classification problem on the image databases VisTex and Outex available on the internet. They also extended the co-occurrence method according to the two other approaches, which are: (fusion of texture and color descriptors and quantization of the color image) to have a comparison between the three approaches to the texture in color images.

Aljarf, Amin, Filippas, and Shuttelworth, (2013) proposed a steganalysis system for gray images using GLCM and color images based on four features which are contrast, energy, homogeneity, and correlation.

Anita, Ramesh, and Vaishali, (2016) introduced an unsupervised optimization technique before classification. They used individual classifiers of SVM and MLP and the fusion techniques that are used to combine these classifiers are Bayes, Dempster-Schafer, and Decision Template schemes. They used Euclidean distance measure and Mahalanobis distance measure to measure the performance of classifiers. They compared the accuracy of different classifiers based on the two measurements above.

Sahu and Chourasia, (2015) reviewed the main steganographic techniques for both lossy and lossless image formats, such as JPEG and BMP. The consequences are presented in terms of a taxonomy that focuses on three principal steganographic techniques for hiding information in image files. Those techniques include those modifying the image in the spatial domain, in the transform domain, and those modifying the image file formatting. Each of these techniques tries to satisfy the three most important factors of steganographic design (imperceptibility or undetectability, capacity, and robustness).One can deduce that while one technique may lack in payload capacity, another may lack in robustness.

Chapter 3
Methodology and the Proposed Technique

3.1 The Methodology Approach

The methodology approach in this research work is experimental. The proposed steganalysis model will be developed to detect the existence of hidden data inside cover gray-scale single-channel images and will be evaluated experimentally. The datasets of clean images for the experiments will be collected from various public sources, and the stego images will be created using two LSB steganography techniques. The proposed texture feature sets for the analysis of clean and stego images will include statistical functions, applied to parts of the image for the more effective detection of hidden messages.

3.2 Basic Premise of the Proposed Research

This research aims to enhance the steganalysis of images by focusing attention on parts of an image that are more often used in embedding in cover images. Most steganography techniques embed in the right half part of a pixels' bytes, the 1LSB, 2LSB, 3LSB and 4LSB, which is motivated by the fact that perceptible visual distortion can result from embedding beyond the fourth bit. Therefore, the feature set to be selected for this work will include measurements related to parts of bytes, as well as features that measure the relationship between bytes.

3.3 Introduction to the Proposed Model

In this research, the proposed steganalysis model will be based on selecting appropriate feature set elements that will help to detect embedding in a cover image, with

focusing on the right half of bytes of the image. The proposed model will be realized in several phases. The first phase is the feature selection based on statistical features. The second phase is the feature extraction from a set of clean and stego images, and the creation of a labeled training dataset of clean and stego images (50% of each). The last phase is the analysis and detection of an unknown test image, by extracting its feature set; and using a classifier model that has been trained using the training dataset, to classify the unknown test image as clean or stego.

Figure 3.1 shows a diagram of the general framework on which the proposed model is designed. (Xia et. al. 2014)
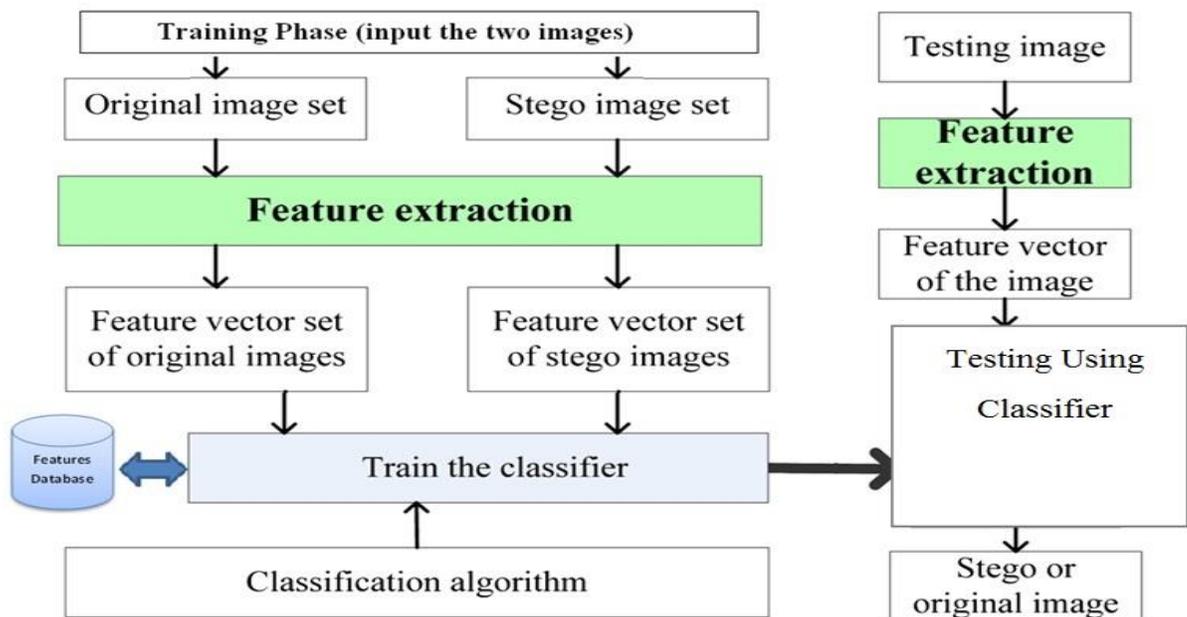
Figure 3.1 diagram of the steganalysis framework (Xia et. al. 2014)

3.4 Statistical Features Selection

The selected features for the proposed model will be statistical image texture measurements that are expected to be used as discriminators between clean and stego images. The feature set elements will be extracted from the images and stored in an Excel CSV files for subsequent batch classification of a set of images.

The proposed model will use the set of features as mentioned in the following sections.

3.4.1 Intra-Pixel Correlation Coefficient

The correlation coefficient is a mathematical technique that is used to study changes in measured values of two variables. In this study we will use two-dimensional (2-D) correlation coefficient to detect the association within a pixel's components, the left half byte (LHB) and the right half byte (RHB), in relation to their intensity (numeric value). In general, if high value of A is associated with high value of B in that case, a positive correlation exists and when high value of A associated with low value of B then a negative correlation exists. The function to measure 2-D correlation coefficient in Matlab is *corr2*:

$$r = corr2(A,B)$$

The result returns the correlation coefficient *r* between *A* and *B*, where *A* and *B* are matrices or vectors of the same size and it can be numeric or logical values while the result *r* is a scalar double.

In general, the correlation coefficient is a statistical measure of how well the trends in the predicted values follow the trends in the past actual values. It's another method to study image features and properties. In this thesis, we will compare two variables by

dividing the bytes into two parts right-half byte and left-half byte and study each part as separate entity to measure the degree of association between the two parts.

3.4.2 Image Comparison Using GLCM Image Analysis

The GLCM (Gray-level co-occurrence matrix) is a common technique in statistical image analysis that is used to estimate image properties related to second-order statistics. GLCM is a tool used to measure the relation between two neighboring pixels in one offset, as the second order texture, where the first pixel is called reference and the second one the neighbor pixel. GLCM is the two-dimensional matrix of joint probabilities $Pd,\theta(i, j)$ between pairs of pixels, separated by a distance d in a given direction θ. In this work, the homogeneity for features vector is used.

The GLCM functions characterize the texture of an image by calculating how often pairs of the pixel with specific values and in a specified spatial relationship occur in an image, creating a GLCM, and then extracting statistical measures from this matrix. GLCM have four properties which can provide useful information about the spatial distribution of the gray levels in the texture image but cannot provide information about the shape of an image. Those features can be described as in the below table:

Property	Description	Formula
Contrast	Returns a measure of the intensity contrast between a pixel and its neighbor over the whole image Range = [0 (size (GLCK,1) -1) *2] Contrast is 0 for a constant image	$\sum_{i,j}\lvert i-j\rvert^{2}p(i,j)$
Correlation	Returns a measure of how correlated a pixel is to its neighbor over the whole image Range = [-1 1] Correlation is 1 or -1 for a perfectly positively or negatively correlated image correlation is NaN for a constant image	$\sum_{i,j}\dfrac{(i-\mu i)(j-\mu j)\rho(i,j)}{\sigma_i\sigma_j}$
Energy	Returns the sum of squared elements in the GLCM Range = [0 1] Energy is 1 for a constant image	$\sum_{i,j}p(i,j)^{2}$
Homogeneity	Returns a value that measures the closeness of the distribution of elements in the GICM to the GLCM diagonal Range = [0 1] Homogeneity is 1 for diagonal GLCM	$\sum_{i,j}\dfrac{p(i,j)}{1+\lvert i-j\rvert}$

Figure 3.2 GLCM Features (Matlab)

GLCM function result can be either logical or numeric, and it must contain real, non-negative, finite integers.

Figure 3.3 shows how gray co-matrix calculates the first three values in a GLCM. In the output GLCM, element (1 , 1) contains the value 1 because there is only one instance in the input image where two horizontally adjacent pixels have the values 1 and 1, respectively. glcm (1 , 2) contains the value 2 because there are two instances where two horizontally adjacent pixels have the values 1 and 2. Element (1 , 3) in the GLCM has the value 0 because there are no instances of two horizontally adjacent pixels with the values 1 and 3. Gray co-matrix continues processing the input image, scanning the image for other pixel pairs (*i,j*) and recording the sums in the corresponding elements of the GLCM.

	1	2	3	4	5	6	7	8
	1	1	5	6	8			
	2	3	5	7	1			
	4	5	7	1	2			
	8	5	1	2	5			

	1	2	3	4	5	6	7	8
1	1	2	0	0	1	0	0	0
2	0	0	1	0	1	0	0	0
3	0	0	0	0	1	0	0	0
4	0	0	0	0	1	0	0	0
5	1	0	0	0	0	1	2	0
6	0	0	0	0	0	0	0	1
7	2	0	0	0	0	0	0	0
8	0	0	0	0	1	0	0	0
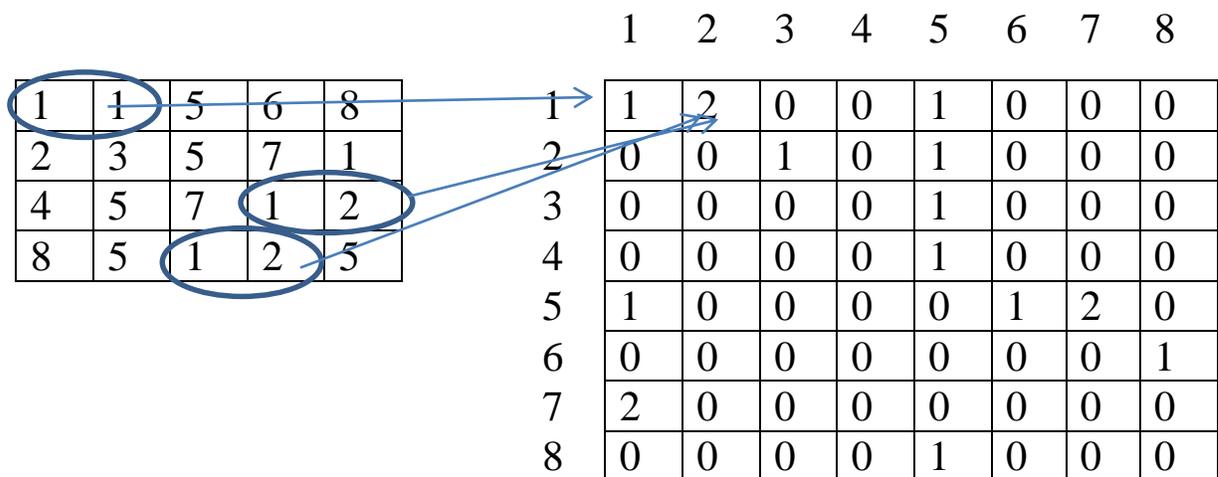
Figure 3.3 GLCM Matrix Process

Hence, a GLCM is a histogram of co-occurring grey-scale values at a given offset over an image.

The image GLCM features are important clues to determine the existence of hiding information or not from the detection process. In this study, we compare detection results of the selected features of the proposed model with results of the GLCM model.

3.4.3 Entropy

Entropy is a measurement of how *unpredictable* of the content of the image. Claude Shannon used entropy to measure the amount of information in the content (an image) in this study.

Entropy mathematical expression can be represented as in the following:

$$\text{Entropy} = -\sum_{i} Pi \; Log2 \; Pi$$

Where *Pi* is the probability of *i*. This means evaluating the summation of the probability of first pixels times log base two probability of first pixels and by doing this for all image pixels we get the entropy of an image as a result.

An example to demonstrate the concept of entropy is tossing a fair coin which contains a probability of 50% tail and 50% head, according to Shannon definition and by applying the above mathematical expression the result of this experiment is having one bit of information of each toss. However, when tossing an unfair coin which has the probability of 25% tail and 75% head, then the result will be equivalent to 0.811278 which means that there is some information that we can get from tossing an unfair coin but it is less than the information as for a fair coin.

In this study, we use entropy as a statistical measure of randomness that can be used to characterize the texture of the input image between left half byte and right half byte. Entropy can be logical, uint8, uint16, or double and must be real, nonempty, and no sparse. Entropy function in Matlab is defined as:

```
E = entropy(I)
```

3.4.4 Difference between Adjacent Bytes

This feature is presented as a measure of the degree of change in intensity between adjacent bytes. It is assumed that a tampered image will have higher differences in the values of adjacent bytes than in clean images, due to the change introduced by bit replacement in bytes of the stego image. The difference feature is calculated as the average of the absolute difference in value between every two adjacent bytes of an image. The absolute value of a difference is considered because a change due to bit replacement can increase or decrease a byte value, hence this feature measures the rate of change regardless if it is positive or negative.

3.4.5 Coefficient of Variation

The coefficient of variation (CV), which is the ratio of the standard deviation to the average of a set of values, provides an indicator of variability or dispersion of the observed values. When data is embedded in the bytes of an image, such embedding will influence the variability of the image's bytes, in comparison with the bytes of the clean image. In our work, we will measure the coefficient of variation of the right half bytes of an image.
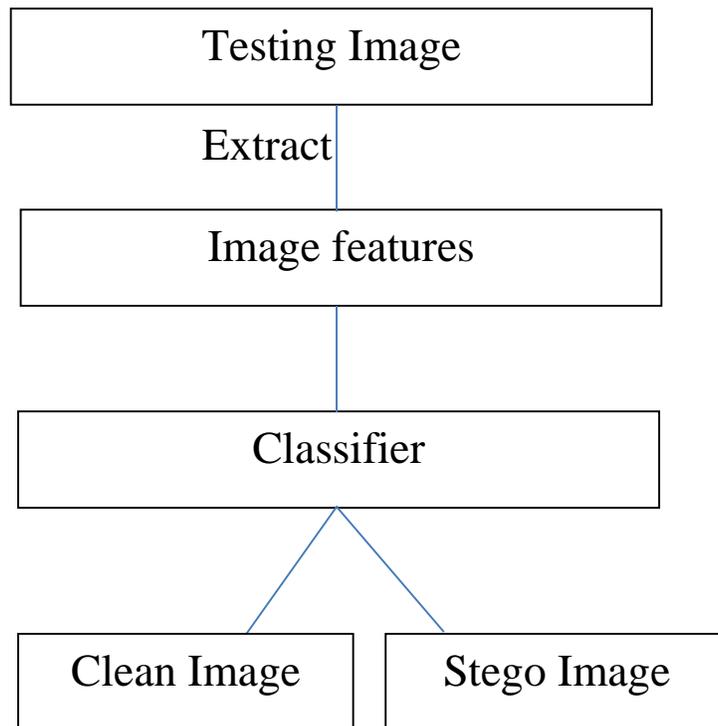
Figure 3.4 the process of the proposed detection system.

3.5 Classifier Selection

In the previous step, we used a set of features to train the system to capture clean and stego image features, and keep the result as datasets for each feature and use it in the detection phase. At this step, we must be able to classify any new grey-scale image if it is clean or stego image, we choose a discriminant analysis classifier to analyze the features datasets for the training phase and based on the training process to classify an unknown image. A new testing image will be tested as clean or stego based on the results of the training phase. The new testing images will include both clean and stego images.

3.5.1 Discriminant Analysis Classifier

A statistical analysis classifier (discriminant analysis) used to classify objects into groups based on a set of measurable object's features. In general, the process starts with observation of the object's features. The end result of the process is a system that is able to predict the category of an object. The result categories are stego or clean (it called dependent variable sometimes as well), while features that describe the object is called the independent variables.

The discriminant analysis classifier has been used in a variety of statistical applications including image processing and optical change recognition. In general, discriminant analysis is available as a function in many data analysis applications such as Matlab.

3.6 Training and Testing Steps

1. Dataset creation: A dataset of triplets of secret files, clean images, and stego images will be created using an existing steganography tool.
2. Feature extraction: Using the proposed feature set, the clean and stego images will be processed and the features extracted, which will result in two datasets: clean image features dataset and stego features dataset.
3. Training: The extracted features datasets of clean and stego images will be used in the training process.
4. Testing: A set of clean and stego images will be tested using the trained classifier, where each image will be classified as either clean or stego. The accuracy of the proposed model will be evaluated accordingly.

3.7 Feature-Extraction and Batch-Testing Modules

The two main modules of the proposed system are the feature-extraction module and the Batch-Testing module. Figures 3.5 and 3.6 present flowcharts of the two modules.
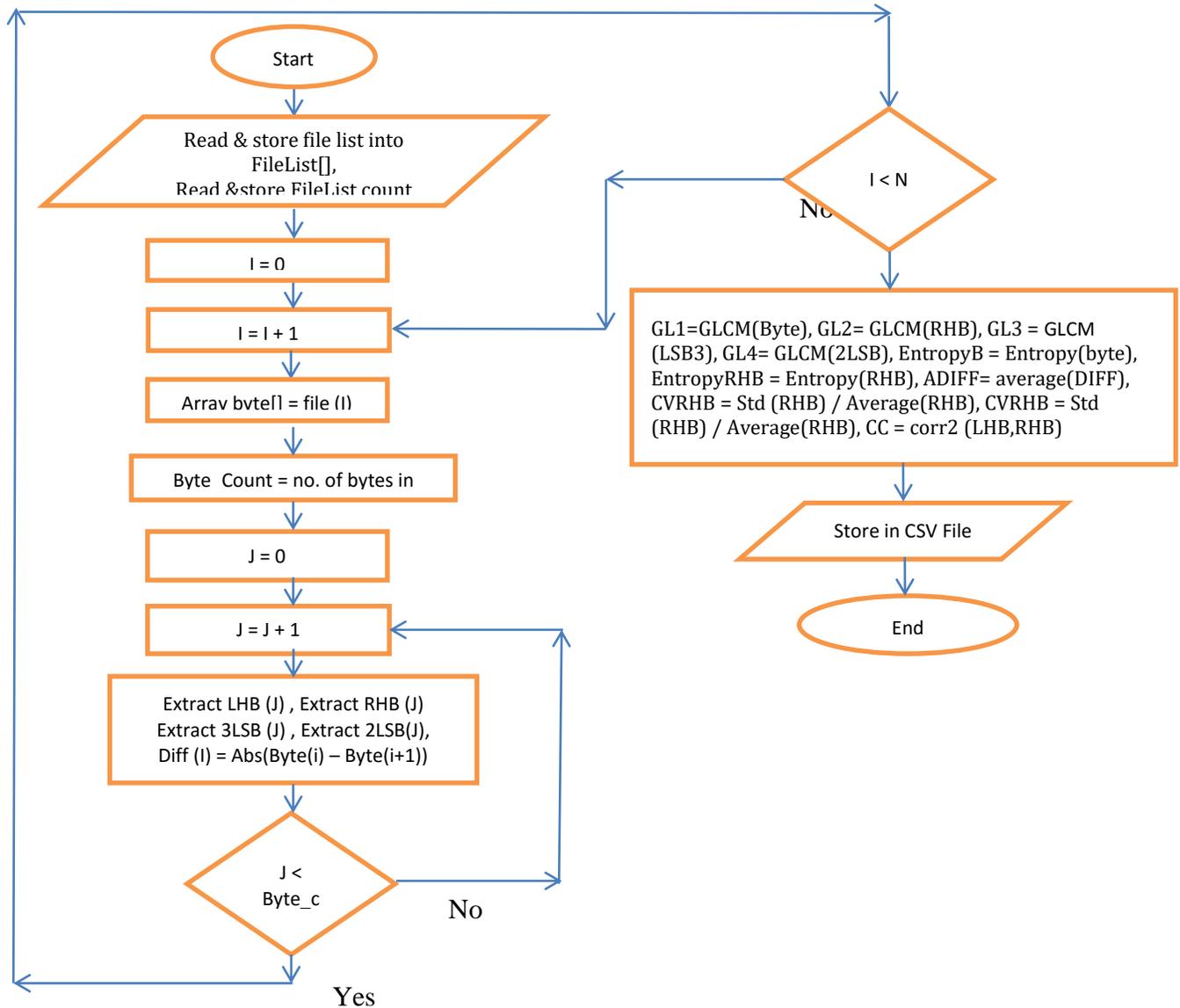
Figure 3.5 Feature extraction flowchart
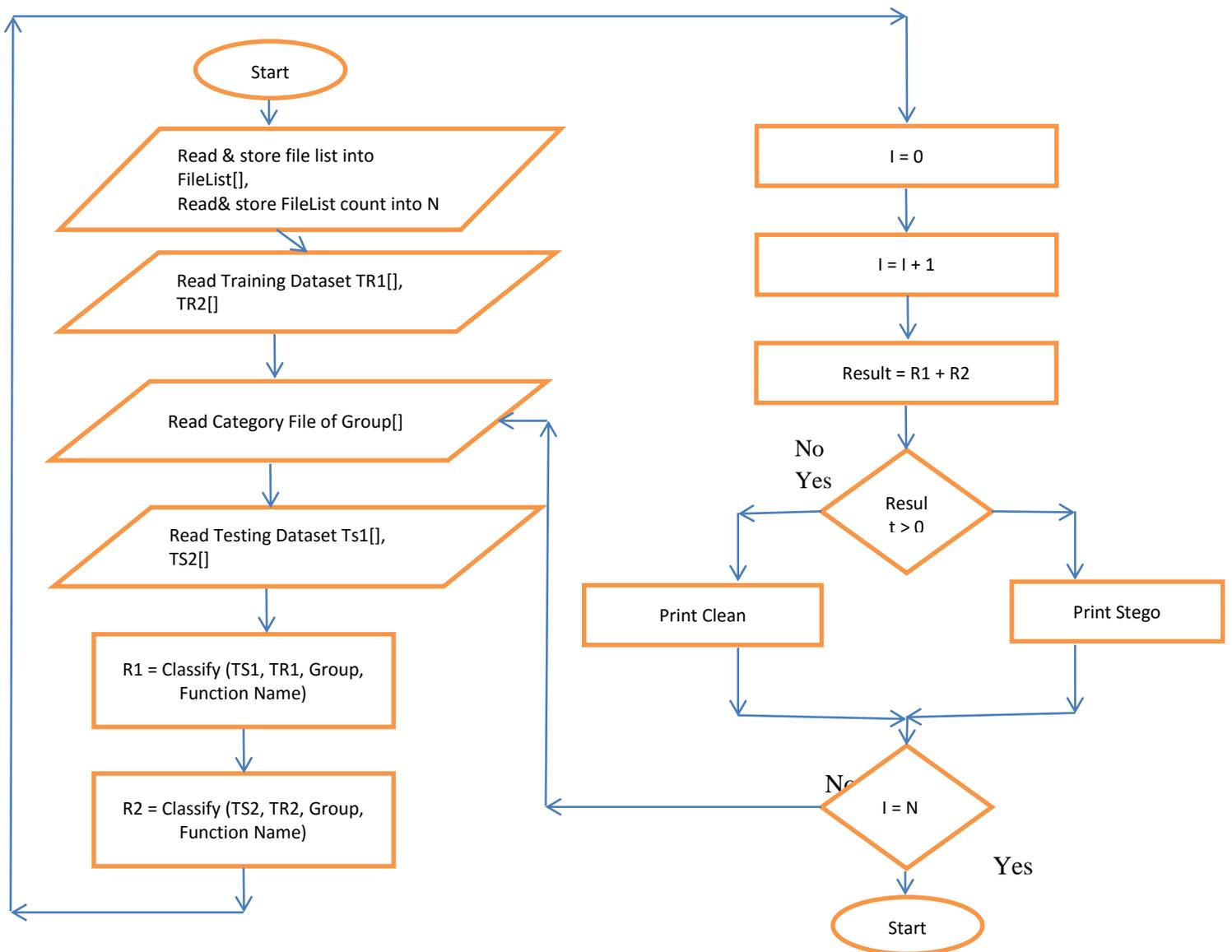
Figure 3.6 Batch test flowchart

Chapter 4

Experimental Results and Discussion

4.1 Introduction:

The proposed model was implemented in MATLAB as a working system, titled "Experimental Steganalysis System (ESS)". The ESS system embeds secret multimedia files in 8-bit depth images. For the present experimental work the cover images were chosen to be uncompressed 8-bit gray-scale BMP images, but the system can embed in 8-bit GIF images, color or gray-scale. The stego images that were analyzed in this work were created using two steganography models, the 2LSB model which embeds 2 bits per pixel (2 bpp) and 4LSB model, which embed 4 bits per pixel (4 bpp). The ESS system can process one image, or work in batch mode, to analyze a large number of images.

4.2 Evaluation Metrics

Detection performance of the proposed model has been evaluated using the following metrics:

TP: The number of true positive results, which represents the images that have embedded data and are classified as stego.

FP: The number of false positive results, which represents clean images that are classified as stego.

TN: The number of true negative results, which represents clean images that are classified as clean.

FN: The number of false negative results, which represents stego images that are classified as clean.

$$\text{Precision} = \frac{TP}{(TP + FP)} \qquad\qquad (1)$$

$$\text{Accuracy} = \frac{(TP + TN)}{(TP + FN + TN + FP)} \qquad\qquad (2)$$

4.3 Implementation

The implementation of the experimental steganalysis system (ESS) consists of the following modules:

1. **Steganography:** Batch embedding a secret message file inside a set of gray-scale BMP images, using 4LSB and 2LSB steganography methods.

2. **Features extraction:** Batch extraction of the selected features from the images of clean and stego image sets. The feature set values are stored in Excel CSV files. A training file is created by combining an equal number of rows of clean and stego features data.

3. **Classifier:** Classifying an image as clean or stego, based on the training data collected in the feature extraction phase. The classifier used in this process is the Discriminant Analyzer that is available in MTLAB.

4. **Checkit application:** an application which uses the extracted features of the training file to classify a single unlabeled image, to determine if it is a clean or a stego image.

5. **BatchTest application:** an application which uses the extracted features of the training file in analyzing a set of unlabeled clean and stego images, to determine whether they are clean or stego images.

The above modules were developed and implemented using MATLAB R2015a.

4.4 The Selected Features

The selected feature set is selected to discriminate between clean images and stego images of two types; 2LSB and 4LSB stego images. The refinement of the feature set was influenced by the experimental work, which resulted in separate feature sets for the two stego models.

Table 4.1 shows the feature set elements for the 2LSB stego model.

Table 4.1: Feature set for the 2LSB and 4LSB methods

Feature Name	Feature Description
CC1	Correlation coefficient between LHB and RHB
CV2	Coefficient of variation of RHB
GLCM1	Contrast, Correlation, Homogeneity Energy, of full bytes
GLCM2	Contrast, Correlation, Homogeneity, Energy, of RHB
GLCM3	Contrast, Correlation, Homogeneity, Energy, of 3LSB RHB
GLCM4	Contrast, Correlation, Homogeneity, Energy, of 2LSB of RHB
EntropyB	Entropy of full bytes
EntropyRHB	Entropy of RHB
DIFF	Average of absolute difference between successive bytes

For the 4LSB steganography model, the feature set includes the same feature set of the 2LSB method, excluding GLCM3 and GLCM2 features. These two features were removed from the 4LSB feature set because they did not make a difference in the detection accuracy of the 4LSB images.

4.5 Experimental Datasets

In this research we used two sets of images:

1. **Basic:** For model evaluation, in training and testing. The dataset represents a mixed collection of images from various public sources, such as Carnegie Mellon University (CMU) vision dataset (CMU, 2016), DECSAI gray-scale images dataset (DECSAI, 2016), Oxford University flower dataset (Flower Datasets, 2016) and Wikimedia Commons images (Wikimedia, 2016). Some of the images were gray-scale of the JPG format, and they were converted to 8-bit gray-scale BMP format using the FastStone Photo Resizer (FastStone, 2016). The other images were in color / JPG format, and they were converted to gray-scale 8-bit BMP format. The stego images for this dataset were generated using 2LSB and 4LSB embedding techniques. See Appendix A for basic image dataset.

2. **Extended:** An extended dataset from Caltech (CUB-200-2011, 2016), of 11,580 birds color JPG images, which were converted to gray-scale 8-bit BMP images.

The first 5000 images were chosen for batch testing in this experiment, while 1500 images were randomly selected from the images 5001-10000 for the training of the

batch testing module. The stego images were generated using the 2LSB and 4LSB
modules. A subset of 90 images this dataset was included in the basic dataset.

Both datasets were converted to a fixed resolution of 512x512 pixels, 8-bit gray-scale BMP
format, which resulted in image size of 257 KB. Figure 4.1 shows a sample of original
color JPG images and the converted gray-scale BMP images from the Caltech dataset.
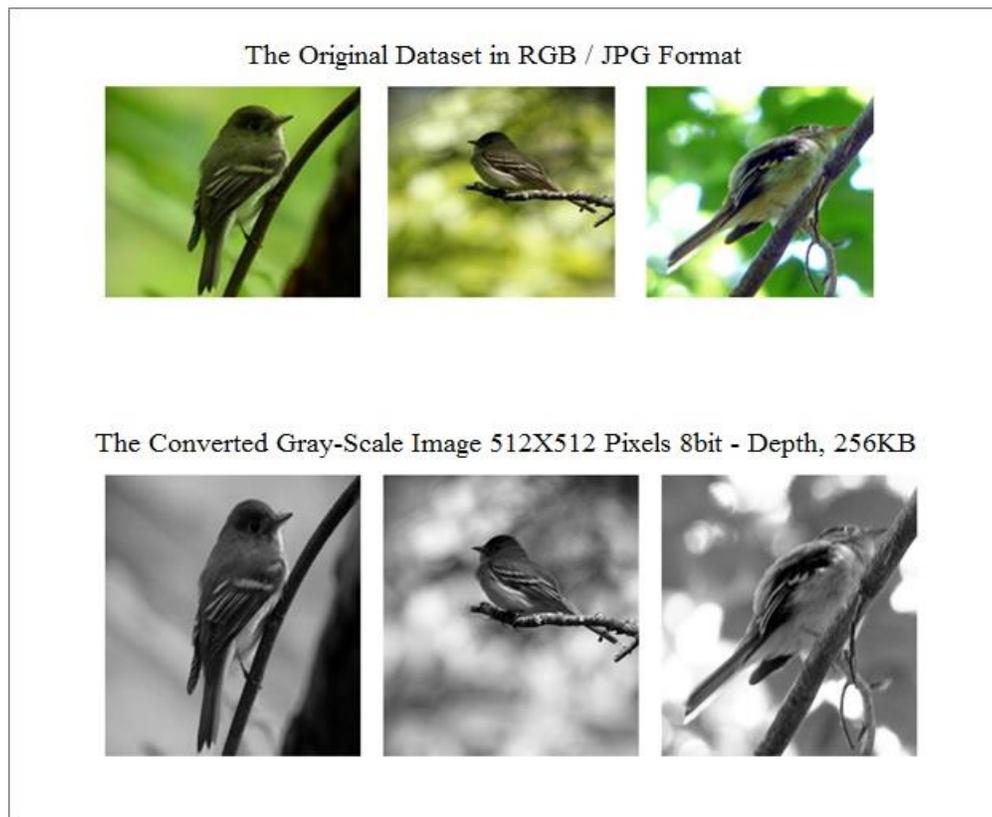
Figure 4.1: A sample of Caltech color JPG and gray-scale BMP images

The image in Fig. 4.2 shows a JPG color image of size 128Kb, which was used as the secret
message to be embedded using the 4LSB embedding technique to create the first stego
dataset.
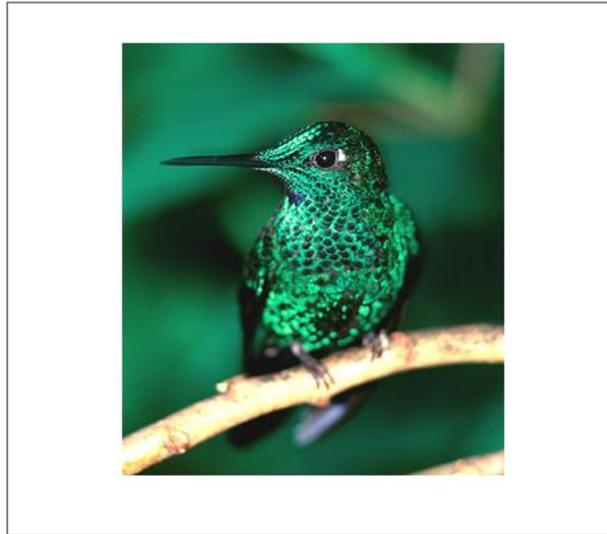
Figure 4.2: The Secret Image in JPG format 495 x 600 Pixels, Size 128KB, for 4LSB Embedding

The image in Fig. 4.3 of size 64KB, is used as the secret message to be embedded using the

2LSB embedding technique to create the second stego dataset.

Figure 4.3: The Secret Image in JPG format 637 x 669 Pixels, Size 64KB, for 2LSB Embedding

Table 4.2 shows a sample of the values of feature set elements for a group of images, which was generated by the feature extraction module.

Table 4.2 a sample of the extracted features

|		GLCM Byte				GLCM RHB				GLCM 3LSB				GLCM 2LSB				Entropy		
CCLR	CVRHB	Contrast	CC	Energy	Homogen	Contrast	CC	Energy	Homogen	Contrast	CC	Energy	Homogen	Contrast	CC	Energy	Homogen	Byte	RHB	Adiff
-0.02	0.61	0.20	0.98	0.11	0.92	5.12	0.09	0.79	0.91	9.99	0.06	0.62	0.82	17.55	0.03	0.40	0.69	7.67	0.33	3.85
0.06	0.65	0.40	0.95	0.12	0.88	6.15	0.15	0.74	0.89	10.72	0.08	0.59	0.81	18.04	0.04	0.38	0.68	7.72	0.40	4.52
-0.06	0.61	0.39	0.94	0.11	0.87	5.46	0.04	0.79	0.90	10.59	0.02	0.61	0.81	18.37	0.01	0.39	0.67	7.49	0.33	4.56
-0.05	0.61	0.12	0.98	0.16	0.94	5.02	0.20	0.78	0.91	9.56	0.14	0.62	0.83	17.17	0.07	0.39	0.69	7.55	0.36	3.24
-0.05	0.62	1.33	0.74	0.05	0.70	5.78	0.01	0.78	0.90	10.71	0.00	0.61	0.81	18.35	0.00	0.39	0.67	7.83	0.34	5.26
-0.02	0.59	0.32	0.94	0.20	0.90	3.88	0.12	0.84	0.93	8.03	0.12	0.68	0.86	15.28	0.20	0.40	0.73	7.11	0.27	3.33
0.05	0.60	0.50	0.94	0.08	0.85	4.86	0.07	0.80	0.91	9.44	0.06	0.64	0.83	16.76	0.06	0.41	0.70	7.83	0.31	4.13
-0.08	0.60	0.36	0.96	0.11	0.90	4.78	0.07	0.81	0.91	10.02	0.06	0.62	0.82	17.83	0.03	0.39	0.68	7.63	0.31	3.80
-0.09	0.62	0.37	0.92	0.16	0.89	4.65	0.19	0.80	0.92	8.64	0.18	0.64	0.85	15.27	0.16	0.41	0.73	7.54	0.34	3.71
-0.05	0.63	0.57	0.93	0.08	0.86	5.34	0.12	0.78	0.90	9.98	0.08	0.62	0.82	17.42	0.06	0.39	0.69	7.91	0.35	4.26
-0.11	0.64	0.06	0.98	0.21	0.97	5.37	0.17	0.77	0.90	9.72	0.11	0.62	0.83	17.61	0.04	0.39	0.69	7.06	0.37	3.14
-0.12	0.63	0.39	0.95	0.09	0.87	5.34	0.05	0.79	0.90	10.29	0.04	0.62	0.82	18.10	0.02	0.39	0.68	7.77	0.33	4.34
-0.01	0.63	0.35	0.96	0.12	0.89	5.80	0.11	0.76	0.90	10.81	0.05	0.60	0.81	18.50	0.02	0.38	0.67	7.66	0.37	4.59
-0.06	0.61	0.35	0.93	0.11	0.88	5.52	0.03	0.78	0.90	10.70	0.01	0.61	0.81	18.45	-0.01	0.39	0.67	7.73	0.34	4.48
-0.07	0.62	0.38	0.93	0.11	0.87	5.60	0.04	0.78	0.90	10.61	0.02	0.61	0.81	18.41	0.00	0.39	0.67	7.68	0.34	4.60
-0.03	0.61	0.67	0.92	0.10	0.84	5.22	0.10	0.79	0.91	9.72	0.09	0.62	0.83	17.12	0.07	0.40	0.69	7.86	0.34	4.27
-0.13	0.61	0.08	0.97	0.29	0.97	5.03	0.22	0.78	0.91	9.07	0.17	0.63	0.84	16.94	0.08	0.40	0.70	6.50	0.37	2.58
-0.07	0.59	0.33	0.95	0.20	0.90	3.85	0.26	0.82	0.93	7.26	0.35	0.65	0.87	12.96	0.32	0.42	0.77	7.34	0.31	2.94
0.52	0.49	0.15	0.99	0.22	0.94	3.03	0.05	0.88	0.95	6.07	0.07	0.76	0.89	11.28	0.15	0.55	0.80	5.93	0.21	2.78
-0.06	0.66	0.29	0.91	0.38	0.93	3.52	0.46	0.80	0.94	6.78	0.40	0.65	0.88	12.28	0.33	0.44	0.78	6.58	0.37	2.25

4.6 Experimental Result and Discussion

Two modes of tests were conducted to evaluate the accuracy of the proposed model in detecting the existence of a hidden message inside a gray-scale BMP image. The tests gave equal emphasis to testing setgo images and clean images. Therefore False-Negative-Rate (FNR) as well the False-Positive-Rate (FPR) metrics were evaluated, as well as the combined accuracy that is calculated from TNR, TPR, FNR, and FPR.

4.6.1 Model Evaluation Test

The aim of this test is to evaluate the proposed model's accuracy based on a limited dataset before it is applied in a larger scale field test. The basic dataset of 360 images (180 clean and 180 stego images) was used in this test. A 3-fold cross validation method was adopted were the dataset was divided into training and testing data as follows:

Fold-1: Stego and clean images 1-120 are combined into the training dataset of 240 images. Stego and clean images 121-180 are combined into the testing dataset of 120 images.

Fold-2: Stego and clean images 61-180 are combined into the training dataset of 240 images. Stego and clean images 1-60 are combined into the testing dataset of 120 images.

Fold-3: Stego and clean images 1-60 and 121-180 are combined into the training dataset of 240 images. Stego and clean images 61-120 are combined into the testing dataset of 120 images.

The 3-fold test was conducted separately for 4LSB stego images and 2LSB stego images.

Table 4.3 shows results of the 3-fold test for the 4LSB stego model, while table 4.4 shows

the results for the 2LSB stego model.

Table 4.3: 3-fold test results of the 4LSB stego model using the basic dataset

Metric	Average of 3 folds (%)
TNR	100.00
TPR	97.22
FNR	2.78
FPR	0.00
Detection Accuracy	98.61

The high detection accuracy of 98.61% and the low False-Negative-Rate (FNR) of

2.78% are strong indicators but need further testing on larger datasets to confirm the

results. For comparison with other research work, Al-Jarf, et. al. (2013) reported on the

results of steganalysis on gray-scale single channel images, however, the results did not

report on the detection accuracy and FNR, instead they published averages of the feature

set elements which cannot be used for comparison on detection accuracy.

Table 4.4 shows results of the 3-fold test for the 2LSB stego model. The detection

accuracy is very close to the results of the 4LSB model, but it is a little lower due to the fact

that the 2LSB stego image stores half the amount of embedded data compared to the 4LSB

stego image. However, both features sets are equally effective as they give over 95%

detection rate.

Table 4.4: 3-fold test results of the 2LSB stego model using the basic dataset

Metric	Average of 3 folds (%)
TNR	99.44
TPR	95.56
FNR	4.44
FPR	0.56
Detection Accuracy	97.50

4.6.2 Extended Model Evaluation Test

This test is an extended version of the model evaluation test in 4.2.1. A subset of the Caltech dataset, of 1500 images, randomly selected from the images 5001-10000, was used in a 3-fold Cross-validation test.

In each fold, 500 images were used for testing and the remaining 1000 images for training. Table 4.5 shows results of the 3-fold test for the 4LSB stego model, while table 4.6 shows the results for the 2LSB stego model.

Table 4.5: 3-fold test results of the 4LSB stego model using the extended Caltech dataset

Metric	Average of 3 folds (%)
TNR	100.00
TPR	97.47
FNR	2.53
FPR	0.00
Detection Accuracy	98.73

Table 4.6: 3-fold test results of the 2LSB stego model using the extended Caltech dataset

Metric	Average of 3 folds (%)
TNR	99.13
TPR	96.73
FNR	3.27
FPR	0.87
Detection Accuracy	97.93

The extended dataset results confirm the high accuracy results that were obtained using the basic dataset, despite the difference in image set sizes (180 vs. 1500), and the fact that images of the basic dataset are from mixed sources and formats, while the Caltech dataset is from one source and in one original format (color JPG). The higher detection rates in the extended dataset compared to the basic dataset can be attributed to using a larger dataset.

4.6.3 Field Test

The k-fold tests approach, as discussed in 4.4.2 and 4.4.3, are normally used when there is a limited size dataset. However, a real field test can give more dependable results when a large dataset is available, which allows for the training data to be independent from the field testing data.

To conduct a field test, a subset of the Caltech dataset, images 1-5000, were selected as the unlabeled testing dataset, while the training subset consisted of 1500 randomly selected images, from images 5001-10,000. The 1500 testing clean images were used to generate 1500 4LSB stego images and 1500 2LSB stego images, which results in two training sets:

TR1: 3000 labelled images (1500 clean + 1500 4LSB stego).

TR2: 3000 labelled images (1500 clean + 1500 2LSB stego).

Similarly, the 5000 testing images were used to generate 5000 4LSB stego images and 5000 2LSB images, which resulted in three testing sets:

TS1: 5000 un-labelled 4LSB stego images

TS2: 5000 un-labelled 2LSB stego images

TS3: 5000 un-labelled clean images.

The Batch-Test module was used in the analysis of the three testing sets, and in each batch-test, the two training sets (TR1 and TR2), were used, which means that each image was analyzed blindly against the two training tests without knowledge of which type it was (clean, 2LSB or 4LSB stego).

Tables 4.7 and 4.8 show the field test results of analyzing 5000 clean images, 5000 4LSB stego images, and 5000 2LSB stego images. The results show that even when applying blind steganography, i.e. using two different features sets and two training datasets, the detection accuracy is high and comparable in terms of the detection accuracy with the results of the basic and extended 3-fold tests. However, the new results have a better (lower) False-Negative-Rate (FNR), which is the most important of the performance metrics in steganalysis, where lower FNR means less stego image can escape detection.

Table. 4.7 Field test results of analyzing 5000 clean images and 5000 4LSB images

Image Type	Metric	Number	%
Clean Dataset (5000)	FP	48	0.96%
	TN	4952	99.04%
4LSB Stego Dataset (5000)	FN	124	2.48%
	TP	4876	97.52%
	Detection Accuracy of 10,000 images (Clean+4LSB)		98.28%

Table. 4.8 Field test results of analyzing 5000 clean images and 5000 2LSB stego images

Image Type	Metric	Number	%
Clean Dataset (5000)	FP	48	0.96%
	TN	4952	99.04%
2LSB Stego Dataset (5000)	FN	170	3.40%
	TP	4830	96.60%
	Detection Accuracy of 10,000 images (Clean+2LSB)		97.82%

Chapter 5

Conclusion and Future Work

5.1 Conclusion

The work in this thesis presented a new statistical steganalysis model, for the detection of the existence of hidden data inside 8-bit depth gray-scale BMP images. The proposed model enhanced the image texture feature set by analysing both full-bytes and parts of bytes of an image, taking into consideration that most steganography methods embed data in the right-half of a byte, to avoid visual and PSNR detections. The experimental work was carried out in three stages: 3-fold cross-validation of a basic dataset, 3-fold validation of an extended dataset and a field test using a large (5000 images) testing dataset and an independent training dataset of 1500 images.

Based on the obtained experimental results, the following conclusions are presented below:

1. The 3-fold cross-validation test, using a dataset of 180 mixed source images, yielded a detection accuracy of 98.61% for clean and 4LSB stego images, and 97.50% for clean and 2LSB stego images, which is higher than previously reported results for a similar case of using a single-channel cover image.

2. The 3-fold cross-validation test, using 1500 images from one source, yielded a similar detection accuracy to the basic dataset results, with better (lower) FNR results.

3. The large-scale field test, using 5000 clean images, 5000 2LSB stego images, and 5000 4LSB stego images, resulted in a detection accuracy of 98.28% for stego images that were embedded with 4 bpp, while the 2 bpp gave a close accuracy of 97.82%. No similar

experiment in terms of sample size and image format is available for comparison; however, such a high detection rate suggests that the proposed model offers a promising solution for the detection of hidden data, in single and possibly multi-channel images.

4. The implemented system provided a blind steganography approach, by detecting hidden data that were embedded using two steganography models, the 2LSB model with 2 bpp embedding, and the 4LSB model with 4 bpp embedding. It is possible to enhance the system by adding detection for other steganography models.

5. The selected binary classifier, the Discriminant Analyzer, has shown to be practically efficient in processing 15,000 clean and stego images.

6. The focus on the right half of each byte, which is the zone where embedding takes place in most steganography models, has contributed to the high detection accuracy.

5.2 Future Work

Based on the observed results and outcome of this research, the following suggestions for future work are presented:

1. Investigating the proposed model to detect hidden data in RGB images, and possibly in other media such as audio and video.

2. Enhancing the blind steganography capability of the proposed model by adding features that can detect embedding from other steganography models.

3. Investigating the use of combined multiple classifiers to enhance the detection process and improve the detection accuracy.

4. Enhancing the proposed model by adding and experimentally testing other statistical image texture features.

References

Aljarf, A., Amin, S., Filippas, J., & Shuttelworth, J. (2013). Develop a Detection System for Grey and ColourStego Images. *International Journal of Modeling and Optimization*, *3*(5), 458.

nderson, R. J., & Petitcolas, F. A. (1998). On the limits of steganography. *Selected Areas in Communications, IEEE Journal on*, *16*(4), 474-481.

Arvis V., Debain C., Berducat M. & Benassi A. (2004). Generalization of the concurrence matrix for colour images: Applicationto Colour Texture Classification. Image Anal Stereo.

Aveibas I., Memon N. & Sankur B. (FEB 2003). Steganalysis based on image quality metrics. IEEE Transactions on image processing, vol. 12, No. 2.

Chandramouli, R., & Memon, N. (2001). Analysis of LSB based image steganography techniques. In *Image Processing, 2001. Proceedings. 2001 International Conference on* (Vol. 3, pp. 1019-1022). IEEE.

Cheddad A., Condell J., Curran K. & Mc Kevitt P. (March 2010). Digital Image Steganography: Survey and Analysis of Current Methods. Signal Processing, Volume 90, Issue 3, , pages 727-752.

Christaline J. Anita, Ramesh R. & Vaishali D. (2016). Optimized JPEG Steganalysis. International Journal of Multimedia and Ubiquitous Engineering Vol.11, No.1, pp.385-396, 2016.

Codr, J. (2009). Unseen: An Overview of Steganography and Presentation of Associated Java Application C-Hide. *Retrieved January*, *8*, 2010.

Caltech-UCSD Birds-200-2011 Dataset (2016), http://www.vision.caltech.edu/visipedia/CUB-200-2011.html, viewed on 1/10/2016.

Duric, Z., Jacobs, M., & Jajodia, S. (2005). 6-Information Hiding: Steganography and Steganalysis. *Handbook of Statistics*, *24*, 171-187.

DECSAI Dataset (2016), http://decsai.ugr.es/cvg/CG/base.htm, viewed on 1/10/2016.

CMU Vision Group Datasets (2016), http://www.cs.cmu.edu/~cil/vision.html, viewed on 1/9/2016.

Farid, H. (2001). *Detecting steganographic messages in digital images* (Vol. 2, p. 12). Technical Report TR2001-412, Deparment of Computer Science, Dartmouth College.

Fridrich J., Goljan M. & Du R. (2001). Steganalysis based on PNG compatibility. SPIE Multimedia Systems and Applications IV.

Fridrich, J., & Goljan, M. (2002, April). Practical steganalysis of digital images: state of the art. In *Electronic Imaging 2002* (pp. 1-13). International Society for Optics and Photonics.

Fridrich, J., & Long, M. (2000). Steganalysis of LSB encoding in color images. In *Multimedia and Expo, 2000. ICME 2000. 2000 IEEE International Conference on* (Vol. 3, pp. 1279-1282). IEEE.

Fridrich, J., Goljan, M., & Du, R. (2001, October). Reliable detection of LSB steganography in color and grayscale images. In *Proceedings of the 2001 workshop on Multimedia and security: new challenges* (pp. 27-30). ACM.

FastStone Photo Resizer (2016), http://www.faststone.org/FSResizerDetail.htm, viewed on 1/10/2016.

Flower Datasets (2016), http://www.robots.ox.ac.uk/~vgg/data/flowers/index.html, viewed on 15/9/2016.

Gong, R., & Wang, H. (2012). Steganalysis for GIF images based on colors-gradient co-occurrence matrix. *Optics Communications*, *285*(24), 4961-4965.

Guillon, P., Furon, T., & Duhamel, P. (2002, April). Applied public-key steganography. In *Electronic Imaging 2002* (pp. 38-49). International Society for Optics and Photonics.

Johnson N.F. (1995) Steganography Technical Report. November 1995 http://www.jjtc.com/pub/tr_95_11_nfj/sec401.html

Johnson, N. F., & Jajodia, S. (1998). Exploring steganography: Seeing the unseen. *Computer*, *31*(2), 26-34.

Johnson, N. F., & Jajodia, S. (1998, April). Steganalysis of images created using current steganography software. In *Information Hiding* (pp. 273-289). Springer Berlin Heidelberg.

Johnson, N. F., & Jajodia, S. (1998, September). Steganalysis: The investigation of hidden information. In *Information Technology Conference, 1998. IEEE* (pp. 113-116). IEEE.

Morkel, T., Eloff, J. H., & Olivier, M. S. (2005, June). An overview of image steganography. In *ISSA* (pp. 1-11).

Nagaraj, V., Vijayalakshmi, V., & Zayaraz, G. (2013). Overview of Digital Steganography Metods and Its Applications. *International Journal of Advanced Science and Technology*, *60*, 45-58.

Nissar, A., & Mir, A. H. (2010). Classification of steganalysis techniques: A study. *Digital Signal Processing*, *20*(6), 1758-1770.

Provos, N. (2001, August). Defending Against Statistical Steganalysis. In *Usenix security symposium* (Vol. 10, pp. 323-336).

Provos, N., & Honeyman, P. (2001). *Detecting steganographic content on the internet*. Center for Information Technology Integration.

Provos, N., & Honeyman, P. (2003). Hide and seek: An introduction to steganography. *Security & Privacy, IEEE*, *1*(3), 32-44.

Sahu, P., & Chourasia, B. (2015). State of the Art in Modern Steganography-A review. *International Journal of Advanced Electronics and Communication Systems*, *4*(1).

Sullivan, K., Bi, Z., Madhow, U., Chandrasekaran, S., & Manjunath, B. S. (2004, October). Steganalysis of quantization index modulation data hiding. In *Image Processing, 2004. ICIP'04. 2004 International Conference on* (Vol. 2, pp. 1165-1168). IEEE.

WIKIMEDIA COMMONS Images (2016), https://commons.wikimedia.org/wiki/Category:Images, viewed on 1/11/2016.

Appendix A

Basic Image Dataset