REPUBLIQUE ALGERIENNE
DEMOCRATIQUE ET POPULAIRE
MINISTERE DE L'ENSEIGNEMENT
SUPERIEUR ET DE LA RECHERCHE
SCIENTIFIQUE

UNIVERSITE LARBI TEBESSI - TEBESSA
FACULTE DES SCIENCES EXACTES ET SCIENCES DE LA NATURE ET DE VIE
DEPARTEMENT DE MATHEMATIQUE ET INFORMATIQUE

MEMOIRE

DE FIN D'ETUDES POUR L'OBTENTION DU DIPLOMEDE MASTER

EN INFORMATIQUE

SPECIALITE: SYSTEME MULTY MEDIA

THEME

Ensemble Classifieurs pour la stéganalyse des images numériques

Présenté par : DJAFFALI Nacereddine

Devant le jury :

Zeggari Ahmed	**MCB**	**Président**
Zammar Ammar	**MAA**	**Examinateur**
Laimeche Lakhdar	**MCA**	**Encadreur**

Année Universitaire 2021/ 2022

Remerciement

Je tiens avant tout à remercier mon encadreur, Ms. Laimeche Lakhdar qui m'a aidé beaucoup. Je les remercie également pour leur temps et son investissement dans tous les aspects de mon travail.

Je tiens également à remercier les membres du Jury :

Dr. Zeggari Ahmed et Dr. Zemmar Amar.

Je remercie également les étudiants du Master 2 que j'ai eu le plaisir d'étudier avec et mes collègues au travail.

Je remercie toutes les personnes que j'ai pu rencontrer et avec lesquelles j'ai pu échanger.

Ces remerciements ne seraient pas complets sans remercier tous mes enseignants de l'année scolaire 2021/2022.

Merci à toute ma famille.

Djaffali Nacereddine

Dédicaces

Je dédie ce modeste travail :

A l'être le plus cher de ma vie, ma mère.

A celui qui m'a fait de moi un homme, mon père.

A mes chers frères et sœurs

-Rima Lilia Mohamed manel et Mouaad

A tous mes Nis et neveux

A tous mes amis de promotion Khairy, Mounime, Chams, Hamou, Farouk, Abed ali……,

A tous mes professeurs qui m'ont encadré toutes au long des années de mes études

A mon encadreur cher monsieur laimech lakhder qui a cru mon travail et m'a soutenu avec patience.

Je dédie ce travail à tous ceux qui ont participé à ma réussite.

Résumé

Le travail présenté dans ce manuscrit s'inscrit dans le domaine de la stéganalyse, autrement dit, la détection des informations cachées dans les images numériques. Nous proposons une méthode de stéganalyse basée sur la méthode d'extraction de caractéristiques HOG et l'ensemble classifieur : XGBoost. Nous avons testé la méthode proposée sur une base de données de 50 000 images numériques originales et 70000 images stéganographiées avec deux méthodes de stéganographie qui sont difficilement indétectables. La méthode proposée a également montré taux de détection élevé.

Mots clés : *Stéganographie, Stéganalyse, XGBoost, HOG, Tatouage.*

Abstract:

The presented work belongs in the field of steganalysis, in other words, the detection of hidden information in digital images. In this work, we propose a steganalysis method based on the feature extraction method HOG and the ensemble classifier: XGBoost. We tested the proposed method on a database of 50,000 cover images and 70000 digital images steganographed with two steganography methods that are difficult to detect. The proposed method also showed high detection rate.

Index term*: Steganography, Steganalysis, XGBoost, HOG, watermarking.*

ملخص

يقدم هذا المشروع طريقة للكشف على المعلومات المخفية داخل الصور الرقمية والتي تعرف بالستيغانااليزيا. في هذا العمل نقدم طريقة للستيغانااليزيا تعتمد على طريقة استخراج الميزات HOG و مجموعة المصنفات *XGBoost*. اختبرنا الطريقة المقترحة على قاعدة بيانات تضم 50000 صورة رقمية أصلية و70000 صورة رقمية تحتوي علي معلومات مخفية بطريقتين. أظهرت الطريقة المقترحة معدل عالي في الكشف على المعلومات المخفية.

الكلمات المفتاحية: الستيغانوغرافيا الستيغانااليزيا *XGBoost* ، *HOG*, الوشم الرقمي

Table des Matières

Liste des Figures

Liste des tableaux

Introduction

La dissimulation d'information cherche à cacher une information de n'importe qu'elle type dans un autre support qui peut être de type texte, image, audio ou vidéo. Les applications de la dissimulation se distinguent par leurs objectifs visés, en stéganographie, le but est de cacher un message dans un support numérique pour permettre à des partenaires de communiquer d'une façon secrète, le support n'a aucun lien avec le message à envoyer. Le tatouage numérique consiste à insérer une marque qui a un lien avec le support numérique. Il est utilisé pour la protection des droits d'auteurs, la protection des copies, l'indexation et la vérification de l'intégrité du document. Si la marque insérer dans un support numérique est différentes pour toutes les copies du support de base, on parle alors du fingerprinting, le but principal est de tracer la source des copies illégales. Malgré les objectifs distincts, ces trois approches partagent des points communs : un support pour la dissimulation (son importance est liée à l'application), des informations à cachées (que se soit un message, marque ou une empreinte) et une clé pour l'insertion et l'extraction ou détection. La différence entre la stéganographie et le tatouage est qu'en tatouage on cherche à marquer le support (on se limite souvent à la dissimulation d'un bit : marquer/pas marquer) pour protéger les droit d'auteurs ou encore de démontrer l'intégrité du document. Une autre différence importante se situe au niveau des attaques. En stéganographie le pirate cherche à lire le message dissimulé dans le support, tandis qu'en tatouage, va chercher à laver le support de toute marque possible.

La stéganographie et la cryptographie sont souvent très proche mais ne vise pas le même objectif, mais on peut dire que ces deux disciplines sont complémentaires. Dans le cas de la stéganographie, la communication n'est pas chiffrée. Elle ne peut pas être détecté par une tierce personne, ce dernier ne se doute pas que les parties communicantes échangent de message. La cryptographie permet d'établir une liaison sécurisée entre deux parties communicantes en chiffrant la communication ce qui la rend incompréhensible pour une tierce personne.

La sétganographie est la technique la plus utilisée dans les images numériques pour assurer une communication secrète, mais, malheureusement, elle peut être utilisée afin d'exécuter des actions illégales. Citons par exemple, de nombreux spécialistes relayés par les médias avancent l'hypothèse selon laquelle Ben Laden aurait coordonné les attentats du 11

septembre 2001 en utilisant des messages cachés dans des images de sites à caractères mauvais.

C'est dans ce contexte, rentre notre travail et qui consiste à la détection de la présence des informations cachées. Ce domaine est connu par la stéganalyse et qui est la contre partie de la stéganographie, elle consiste à attaquer les méthodes stéganographiques par détection, destruction, extraction ou modification des données encapsulées.

Nous allons essayer d'atteindre notre objectif à travers trois chapitres :

- Dans le premier chapitre, nous allons présenter des concepts généraux sur les techniques de la dissimulation de l'information à savoir la stéganographie, le tatouage numérique, le filigrane, l'architecture générale d'un système de dissimulation ainsi que les relations entre ces techniques de protection de l'information.
- Dans le deuxième chapitre, nous allons présenter le domaine de la stéganalyse et les différents types de stéganalyse en l'occurrence stéganalyse spécifique et universelle. Un état de l'art sur les différentes techniques de stéganalyse est ensuite présenté.
- Dans le dernier chapitre, nous présentons la méthode proposée ainsi que les résultats expérimentaux. Dans une première étape, des prérequis théoriques à savoir les ensembles classifieurs et la méthode HOG, sur lesquelles repose notre méthode proposée, sont détaillés. Dans une deuxième étape, les résultats expérimentaux sont détaillés et discutés. Finalement, Nous clôturons ce mémoire par une conclusion générale, ainsi que les perspectives visées.

Chapitre 1

STEGANOGRAPHIE ET LA SÉCURITÉ D'INFORMATION

1.1. Introduction

Aujourd'hui, qui contrôle l'information détient énormément de pouvoir. Dans ces situations où l'information représente un tel enjeu stratégique et économique, il est devenu nécessaire de mettre en œuvre des outils de sécurité adaptés aux nouvelles technologies de communications et de l'information. Actuellement, il est extrêmement simple de reproduire, modifier et transférer n'importe quel médium. C'est pour cette raison, les recherches se dirigent vers d'autres techniques de sécurité, la stéganographie, pour insérer une information (message, marque ou fingerprint) dans un médium afin de protéger ce dernier.

Le terme stéganographie désigne simplement le fait de cacher une information dans un support numérique. Il s'agit d'une libre adaptation de l'expression anglaise information hiding couramment utilisée dans la littérature. Ce chapitre a pour but de présenter les concepts de base de la stéganographie, le schéma général d'insertion et d'extraction en termine ce chapitre par une comparaison entre les techniques existantes de protection de l'information.

1.2. Définition de la stéganographie

La *stéganographie* est un domaine où l'on cherche à dissimuler discrètement de l'information dans un media de couverture (typiquement un signal de type texte, son, image, vidéo, etc...). Elle se distingue de la cryptographie qui cherche à rendre un contenu inintelligible à autre que qui-de-droit. Lorsqu'un acteur extérieur regarde un contenu cryptographié il peut deviner la nature sensible de l'information qui lui est cachée. L'intérêt de la stéganographie réside précisément dans la possibilité de communiquer en échangeant des contenus d'apparence anodine de telle sorte à ne pas éveiller de soupçons.

1.3. Objectifs de la stéganographie

De nombreux usages peuvent exister dans des domaines très variés mais souvent sensibles :

- Communiquer en toute liberté même dans des conditions de censure et de surveillance.

- Protéger ses communications privées là ou l'utilisation de la cryptographie n'est pas normalement permise car elle soulèverait des suspicions.
- Publier des informations ouvertement mais à l'insu de tous, des informations qui pourront ensuite être révélées.

1.4. Stéganographie dans l'histoire

La stéganographie à une très longue histoire qui remonte à la Grèce Antique. Herodotus, auteur grec, raconte les communications secrètes entre deux chefs de guerre qui utilisaient des esclaves pour transmettre un message afin d'organiser une révolte contre les Perses. Afin d'assurer le transfert des messages secrets ou plans de batailles sans aucune suspicion des adversaires, les cheveux d'un esclave de confiance ont étaient rasés pour tatouer le message sur son crâne. Une fois que ses cheveux avaient repoussés le message devenait invisible et l'esclave pouvait être envoyé avec l'ordre de se faire raser le crâne une fois arrivé à destination [1].

Une autre technique était utilisée afin de prévenir les Grecs d'une invasion du roi Xerxès de Perse en envoyant un message gravé dans le bois d'une tablette d'écriture recouverte de cire, d'apparence vierge (voir figure 1.1). Les Grecs vont mettre en place plusieurs mécanismes dédiés à la stéganographie. Des trous sur un disque représentant des lettres, des fils de couleurs différentes, permettaient de lire un message secret [2]

Fig. 1.1. Tablette contenant un message caché gravé sur le bois sous la cire [2].

Une méthode de stéganographie fut inventée par Gaspar Schott (1608-1666). Le principe est de coder un message secret selon des notes de musique. Autrement dit, associer une lettre à une note musicale (voir figure 1.2). L'avantage de ce procédé est que le message apparaissait comme une partition musicale, et donc passait totalement inaperçu [3].

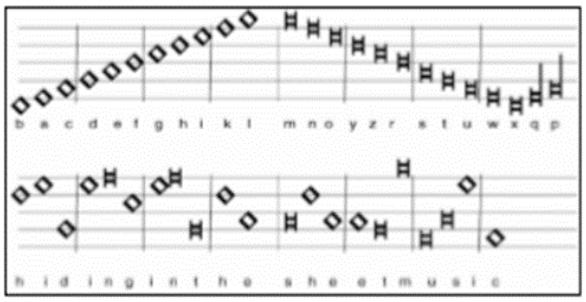


Fig.1.2 : Correspondance entre les lettres de l'alphabet et notes musicales [3].

Au 17ème siècle, Sir John Trevanion fut arrêté et emprisonné dans un château. Il reçut une lettre, que les gardiens avaient jugée sans danger. Lorsque John lut celle-ci, il détecta la présence suspecte de certaines virgules étrangement placées. Il repéra également qu'en prenant la troisième lettre de chaque mot suivant ces virgules, il pouvait former la phrase « Panel at East of Chapel slides » ce qui signifiait « Le panneau à l'extrémité Est de la chapelle peut glisser ». C'est ainsi qu'il demanda un instant de recueillement dans la chapelle et s'évada [3].

Une autre forme stéganographique était connue par le sémagramme qui consiste à transmettre les messages secrètes dans les initiales de chaque vers de poème, mot placé dans des vers ou utilisation de la ponctuation (points, hauteur de lettres et virgules). Alfred de Musset est l'utilisateur le plus connu de ce procédé puisqu'il a entretenu une relation secrète avec Georges Sand (entre 1833 et 1834) au travers de poème qu'il lui envoyait.

Dans les années 1940, la stéganographie a été très souvent employée et s'est ouverte à un grand nombre de formes. Durant la seconde guerre mondiale, plusieurs techniques de stéganographie ont été utilisées. L'encre invisible est la technique stéganographique la plus connue avec laquelle les messages secrets sont écrits sur un papier et qui apparaissent lorsque le papier est approché d'une source de chaleur (voir figure 1.3) [3].

Avec le développement des produits chimiques, par la suite des encres invisibles ont été développé pour communiquer des messages en toute sécurité comme le chlorate de soude ; l'écriture apparaît en passant sur l'encre sèche une petite éponge trempée dans une solution de vitriol de cuivre.

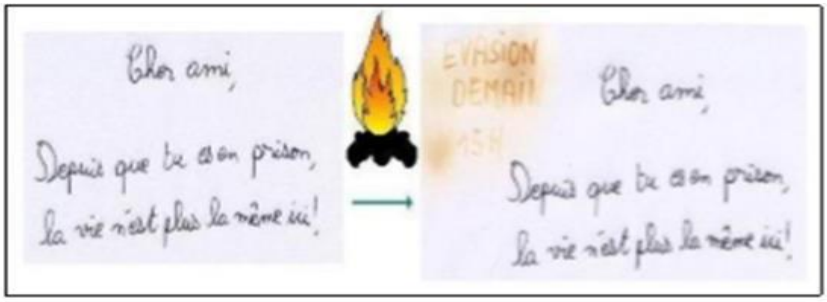


Fig.1.3 : Stéganographie par l'encre invisible [3].

Pendant la seconde guerre mondiale, un procédé de stéganographie a été réalisé par les agents allemands en Angleterre qui envoyaient des pulls en Allemagne contenant des nœuds dans la laine. Les tricots étaient démaillés à leur arrivée. Sur un mur sur lequel se trouvait l'alphabet sous la forme d'une règle, ils posaient l'extrémité du fil à un point et regardait la position du nœud. Grâce à l'emplacement du nœud sur la règle, ils retrouvaient petit à petit le message caché.

Dans les années 1980, Margaret Thatcher, premier ministre britannique a demandé de faire modifier le logiciel de traitement de texte utilisé par les membres du gouvernement afin de dissimuler dans les espaces séparant les mots l'identité de la personne utilisant le traitement de texte [2].

Dans les années 1990, les fabricants d'imprimantes HP et Xerox ont utilisé la stéganographie afin de tracer la falsification de dollar américain. Pour ce faire, de petits points jaunes sont ajoutés au cours de la phase d'impression dans chaque page (voir figure 1.4). Ces points qui sont visible sous la lumière bleue ou avec une loupe peuvent servir à identifier l'imprimante qui a été utilisée (le numéro de série, l'heure d'impression, la marque et le modèle de l'imprimante) (https://w2.eff.org/Privacy/printers/docucolor/).

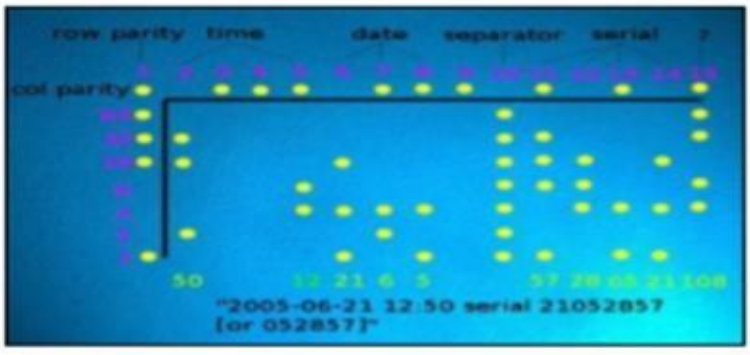


Fig.1.4 : Ajoute des points jaunes par les imprimantes HP.

En 1997, une technique de stéganographie connue par les micros points a été utilisée afin de tracer la falsification des billets suisse.

Cette technique consiste à réduire un texte ou une image en un point d'un millimètre ou moins. Celui-ci est ensuite disposé dans un texte ou une image normale (voir figure 1.5) [3].

Fig.1.5 : Stéganographie dans les billets suisses par la méthode des micros points [3].

De nombreux spécialistes relayés par les médias avancent l'hypothèse selon laquelle Ben Laden aurait coordonné les attentats du 11 septembre 2001 en utilisant des messages cachés dans des images de sites à caractères mauvais.

La stéganographie a été aussi utilisé dans les petites annonces des journaux pour transmettre des messages. On peut citer par exemple, en 2004 un mystérieux groupe terroriste, appelé AZF, qui menaçait de faire sauter des voies ferrées si une rançon ne leur était pas payée. Pour dialoguer avec les autorités, ce groupe exigeait l'utilisation de la rubrique "Messages personnels" de Libération, Cinq messages à « AZF » ont été passés par la police dans Libération [4].

1.5. Stéganographie aujourd'hui

Aujourd'hui, qui contrôle l'information détient énormément de pouvoir. Dans ces situations où l'information représente un tel enjeu stratégique et économique, il est devenu nécessaire de mettre en œuvre des outils adaptés aux nouvelles technologies: une protection renforcée de la vie privée et des droits d'auteur. Il est devenu extrêmement simple de reproduire parfaitement n'importe quel médium. Dans le cas des média numériques (son, image et vidéo), les recherches se dirigent vers une solution technique: insérer une marque dans le médium afin d'identifier l'ayant-droit légitime.

1.5.1. Types des supports utilisés dans la stéganographie

La stéganographie technique regroupe toutes les techniques qui ne jouent pas sur les mots. La stéganographie technique est intéressante car elle permet de dissimuler des données dans plusieurs types de médias.

1.5.1. Fichier image

Ce support de stéganographie est le plus populaire en ces dernières années par rapport à d'autres types de support de stéganographie, à cause de l'inondation des informations d'images électroniques disponibles avec l'avènement de l'appareil photo numérique et la distribution d'Internet en haute vitesse. Ça peut impliquer la dissimulation d'informations dans le bruit produit naturellement dans l'image. La plupart des types d'informations contiennent ce genre de bruit. Le bruit fait référence aux imperfections inhérentes au processus de rendu d'une image analogique en tant qu'image numérique.

Frédéric Raynal [5] propose six catégories de stéganographie dans les images:

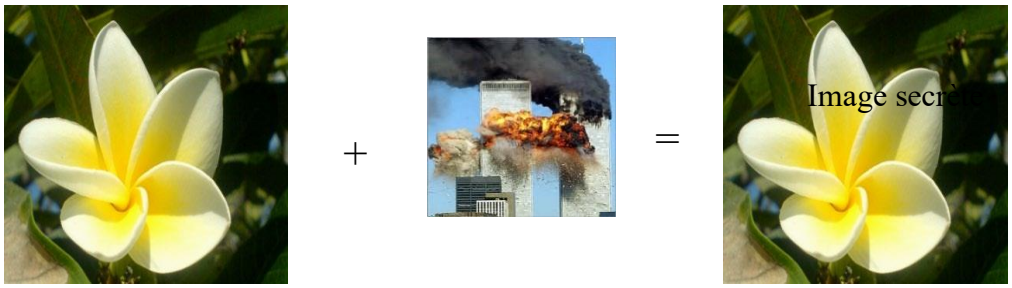
- Un système par substitutions remplace les parties redondantes du support par le message.
- Les techniques par transformations dissimulent l'information dans une transformée du support, comme par exemple, la transformée à cosinus discrète ou le domaine des ondelettes.
- Les techniques par étalement de spectre : les informations sont dissimulées dans tout l'image, et la perte de certaines informations doit pouvoir être compensée par les autres.
- Les méthodes statistiques modifient plusieurs statistiques du support (distribution des pixels, luminosité...) pour cacher le message et le récupèrent en testant ces hypothèses.
- Les techniques par distorsions altèrent le support, la différence avec le support initial constituant alors le message.
- Les méthodes par génération de support construisent un support autour du message pour le dissimuler.

Les méthodes de stéganographie se basent généralement sur la manipulation des bits de poids faibles des pixels, Il s'agit d'insérer sur les bits de chaque pixel d'une image et d'y ajouter le code binaire du fichier numérique secret.

Un pixel d'image (RVB) est souvent codé sur trois octets (soit 24 bits d'information), on comprend aisément qu'une telle dégradation de l'image soit indiscernable à l'œil nu.

Soit un octet de l'image à stéganographiée01101011 (figure 1.6), et un octet de l'image que l'on souhaite cacher 10011101. Le but est de remplacer les bits de poids faible de l'image qui cache par les bits de poids fort de l'image qu'on souhaite cacher.

Ainsi, on obtiendra l'octet 01101001.



Cover Image Image à cachée Image stéganographiée

Fig.1.6 : Exemple d'une image stéganographiée avec la méthode de remplacement LSB

1.5.2. Fichier audio

La stéganographie dans un fichier audio, consiste à dissimuler des messages dans le bruit (audio), ou dans les fréquences que les être humain ne peuvent pas entendre, c'est un autre domaine de dissimulation de données et d'informations qui repose sur l'utilisation d'une source existante comme un espace dans lequel cacher l'information [6].

1.5.3. Fichier vidéo

Les techniques sont équivalentes à celles utilisées dans les images. Cependant les vidéos sont souvent plus bruitées ce qui facilite l'imperceptibilité des données dissimulées mais les rend aussi moins robustes [7].

1.5.4. Fichier HTML

Certains logiciels de stéganographie proposent de cacher des messages dans des pages HTML : ils ne font que toucher la source pour camoufler le fichier secret en insérant des espaces entre balises, variant minuscules et majuscules dans les balises, ... Astucieux mais cela peut toutefois se détecter par analyse statistique et même par un coup d'œil à la source dont l'indentation exotique pourra attirer l'attention [5].

1.5.5. Systèmes de fichiers

Pour stocker un fichier, le système découpe ce dernier en un nombre de morceaux pour que chaque morceau puisse être logé dans un bloc. Comme la taille d'un fichier a rarement une taille multiple de la taille des blocs, généralement le dernier bloc ne sera pas rempli. Pour cacher des données, il suffit de les stocker dans ce dernier bloc ; si la taille de ces données dépasse l'espace du bloc non rempli, il faut les découper et les stockées sur autant de blocs nécessaires, garder la trace des blocs utilisés et l'ordre pour la récupération. Le problème de cette technique vient du fait que les fichiers peuvent être modifiés, supprimés, déplacés, etc. [8].

1.5.6. Fichier exécutable

Les fichiers exécutables peuvent être utilisés pour transmettre un message d'une façon secrète. Lors de la compilation d'un programme, le code source est transformé en un ensemble d'instructions qui sont facile à comprendre par la machine, pout l'exécution, le système d'exploitation lit les sections dont il a besoin. Donc il est possible de bénéficier les parties du code non exécuté [6].

1.6. Classification des schémas stéganographique

Le contexte dans lequel se situe un schéma de stéganographie permet de le classer dans une des catégories suivantes : *stéganographie pure*: aucune entente préalable, autre que le choix de l'algorithme, n'est nécessaire, Alice et Bob utilisent le canal pour échanger des informations ; *stéganographie à clé secrète*: Alice et Bob conviennent au préalable d'une clé qui leur sert à insérer puis extraire le message du stégo-médium ; stéganographie à clé publique: tout comme en cryptographie, Alice utilise la clé publique de Bob lorsqu'elle souhaite lui envoyer un message. Bob, pour sa part, l'extrait à l'aide de sa clé privée.

1.7. Fonctionnement d'un système stéganographique

La mise en œuvre d'un schéma de stéganographie s'effectue en deux phases distinctes : phase d'insertion et phase d'extraction

1.7.1. Phase d'insertion

C'est l'étape dans laquelle les messages sont cachés, elle requiert trois paramètres en entrée:

- Une image dans laquelle des données sont insérées,
- Une information à transmettre, une marque ou une empreinte digitale,

- Un algorithme d'insertion qui sélectionne dans l'image les sous-parties favorables à la dissimulation à l'aide d'une clé de stéganographie générer aléatoirement.

1.7.2. Phase d'extraction

Elle prend en entrée l'image stéganographiée et la clé secrète utilisée dans la phase d'insertion. Cette clé est alors utilisée par la suite pour déterminer les positions contenant l'information cachée.

1.8. Contraintes d'un système stéganographique

1.8.1. La capacité

La capacité d'insertion d'un système de stéganographie est définie par la taille en bits du message secret qui peut être intégré dans un média de taille donnée. La capacité d'insertion relative est le rapport entre la taille du message secret à dissimuler et la taille du médium utilisé. Dans le domaine spatial, pour une image numérique, la capacité d'insertion relative peut être exprimée en nombre de bits de message secret insérés par pixel (bpp). Dans le domaine fréquentiel, par exemple insertion dans les coefficients quantifiés d'une image JPEG, la capacité d'insertion relative peut être exprimée par le nombre des bits du message secret à insérer par chaque coefficient DCT quantifié non-nul (bpc) [4]. Notons que dans ce cas, comme le nombre de coefficients non-nuls dépend du contenu de l'image, la capacité d'insertion est variable d'une image à l'autre.

1.8.2. L'imperceptibilité

Toutes les exigences de sécurité pour les systèmes cryptographiques peuvent (doivent) également être considérées pour les systèmes de stéganographie. Cela signifie que la **sécurité** de l'algorithme de stéganographie ne doit pas s'appuyer seulement sur l'algorithme, qui devrait être publique, mais sur le caractère secret de la clé. Dans la stéganographie, il ne devrait pas être possible de distinguer une image d'origine d'une image stego si la clé est inconnue. Par ailleurs, les modifications apportées sur l'image originale afin de pouvoir incorporer le message secret ne devrait pas modifier les propriétés statistiques de l'image. La technique qui étudie la sécurité des systèmes de stéganographie est la stéganalyse.

1.8.3. La robustesse

Elle quantifie la résistance du message dissimulé aux diverses attaques (transformations) apportées au médium stégo.

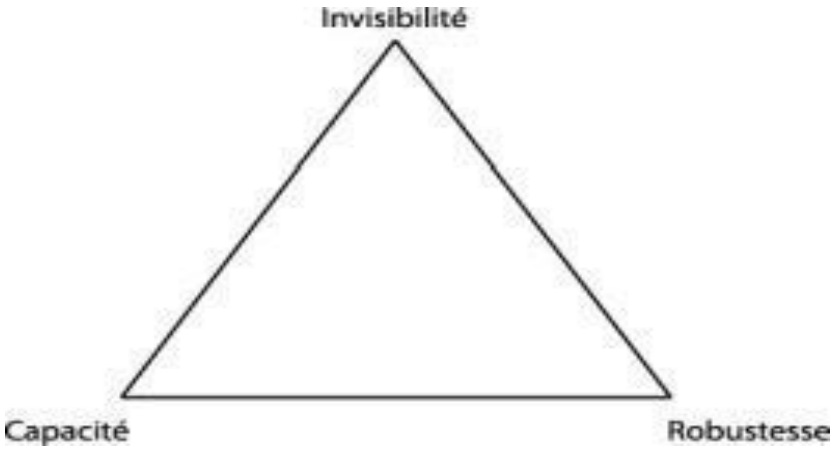


Fig.1.7. : Compromis entre capacité, invisibilité et robustesse

1.9. Applications de la stéganographie

Après la définition de la stéganographie, des questions apparaissent de façon automatique : à quoi peut bien servir la stéganographie? Pourquoi dissimuler un message si l'on n'a rien à se reprocher? Dans le passé, cette science a toujours été utilisée à des fins d'espionnage, pourtant, de nos jours, la stéganographie n'est pas toujours synonyme d'insécurité et peut, au contraire, servir à protéger le droit. Ci-dessous, nous donnons quelques exemples d'utilisation de la stéganographie [9].

1.9.1. Utilisations malveillantes

Internet est une source inépuisable de ressources, la quantité innombrable d'images qui circulent sur le web ainsi que les nombreux fichiers audio qui s'échangent via les communications point à point (P2P) rendent difficile la détection de messages cachés dans ses médias. Un attaquant peut alors utiliser la stéganographie pour cacher un code malveillant fragmenté sur plusieurs images stégo, et procéder ensuite au réassemblage du code malveillant directement sur l'ordinateur de la victime. La stéganographie peut être aussi utilisée pour dissimuler des messages interdits sur des images ou photos anodines et les échanger via l'internet sans provoquer le moindre signe de soupçon. Par exemple, des réseaux terroristes ou des pédophiles qui s'échangent des messages secrets à travers le web. De plus, la stéganographie est aussi utilisée pour l'espionnage industriel. En effet, cette technique semble bien adaptée au transfert de l'information confidentielle volée ou obtenue par corruption.

1.9.2. Utilisations légitimes

Dans certains pays non démocratiques où la liberté d'expression est totalement interdite, la stéganographie apparaît comme un moyen de communiquer plus librement. Dans ces pays, l'utilisation de la stéganographie est illicite mais son usage dans ce cas est légitime. Aussi, lors d'une guerre entre deux nations, la stéganographie est utilisée, de part et d'autre, pour transmettre des messages secrets qui ne doivent pas tomber dans des mains ennemies et si la communication est interceptée, le message caché ne doit être dévoilé (message chiffré ou caché de façon aléatoire ou chaotique).

1.10. Liens avec d'autres techniques de dissimulation d'information

La figure 1.8. résume les différentes techniques de sécurité de l'information.

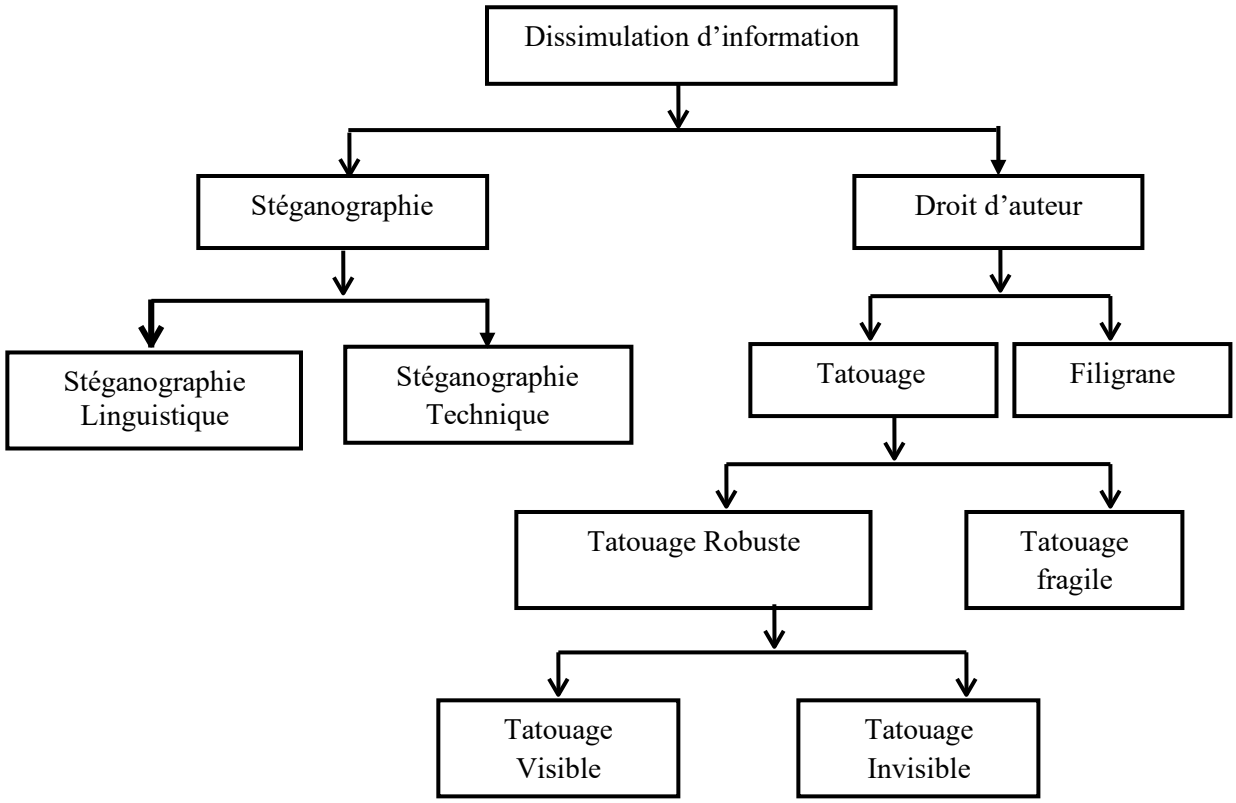


Fig.1.8: Techniques de la sécurité de l'information

1.10.1. Tatouage numérique

Le tatouage comprend deux types : le tatouage fragile et le tatouage robuste.

Le tatouage numérique fragile n'est utilisé que pour prouver l'authenticité des documents et l'intégrité des données. La protection de la marque "tatoue" étant très faible, le message qu'il transporte n'est pas vraiment important [10].

La marque fait partie du document et ainsi, lorsque celui-ci est modifié, le marquage l'est également. Ce type de tatouage pose quand même un problème, car même s'il permet de prouver qu'un document a subi une transformation, il ne prouve pas pour autant qui est l'auteur du document.

Le tatouage robuste est plus dur à contourner et doit résister à diverses attaques. Il doit posséder les deux propriétés suivantes :

- La marque doit être très résistante vis à vis des différentes attaques connues telles que : rééchantillonnage, impression puis scanne, à la compression, à la coupure, aux bruits et aux changements de format.
- La marque doit être facilement reconnaissable après extraction et ceci malgré le dommage subi par les différentes attaques. Dans le cas contraire, la marque pourrait être incompréhensible ou avoir changée de sorte qu'elle n'ait plus rien à voir avec celle insérée à l'origine. Ce type de tatouage est utilisé surtout dans les applications de protection de copyright et de contrôle de copies

1.10.2. Filigrane

Le but de cette application est de pouvoir contrôler et faire le suivi des copies de document. Cela implique de créer une marque originale pour chaque document distribué. Les marques doivent être très robustes, afin de résister aux attaques ayant pour but de les détruire [11].

1.10.3. Cryptographie

C'est la science d'écriture d'un message en code secret afin de préserver sa sécurité et sa confidentialité. Le but est donc de brouiller un message afin de le rendre incompréhensible pour les personnes non autorisées. Le message initial est appelé message en clair et, après chiffrement, message chiffré ou cryptogramme. Le chiffrement et le déchiffrement sont réalisés principalement à partir d'algorithmes, en utilisant des clés secrètes ou publiques [12].

1.11. Comparaison entre les techniques de la dissimulation de l'information

Malgré les objectifs distincts, il est facile de remarquer, que ces trois approches requièrent des paramètres communs [4] :

- Chaque approche nécessite des informations que se soit un message, une marque ou une empreinte.

- Un support pour dissimuler ces informations, son importance dépend de l'application, aucune pour la stéganographie, capital pour le tatouage et le filigrane.
- Utilisation d'une clé pour insérer ou extraire/détecter l'information, la fonction extraction/détection dépend de l'objectif de l'application, extraction en stéganographie et détection pour les deux autres approches.

1.11.1. Stéganographie vs cryptographie

La stéganographie est un art proche de la cryptographie [13]. C'est pourquoi il est essentiel de les distinguer. La dissimulation d'information cherche à dissimuler la présence d'informations pertinentes au sein de plusieurs autres; le message secret est caché dans un support de manière à passer inaperçu lors de la communication. Quant à la cryptographie l'objectif est de rendre l'information incompréhensible à une personne ne possédant pas les connaissances adéquates : une personne surveillant le canal de communication par lequel transite le message sait qu'un échange a lieu, mais est ainsi incapable d'en interpréter le contenue.

1.12. Conclusion

La technique qui aurait été présentée, la stéganographie, consiste à cacher un message dans un support anodin. Elle peut, de surcroît, se combiner à la cryptographie, qui se charge de dissimuler le sens de la missive et non plus son existence. Le résultat est alors particulièrement efficace. Le message secret s'abrite d'abord derrière son invisibilité. En cas de découverte, il restera à le décoder. Un défi pour les services secrets qui, dans le cas du terrorisme, doivent réaliser ces deux opérations au plus vite pour que l'information recueillie ne soit pas obsolète.

Chapitre 2

LA STEGANALYSE DES IMAGES NUMERIQUE

2.1. Introduction

La stéganographie laisse en général des traces qui peuvent être détectables dans l'image stéganographiée. Cela peut permettre à un adversaire, en utilisant des techniques de stéganalyse, de révéler qu'une communication secrète se déroule. Parfois, un attaquant est aussi appelé un « gardien ». Il existe deux types d'attaquant : passifs et actifs. Un attaquant passif examine uniquement la communication pour savoir si la communication contient des messages cachés. Cet adversaire ne modifie pas le contenu de la communication et l'autorise si aucune preuve de message secret n'est trouvée, sinon, il l'a bloque. Un attaquant actif peut provoquer volontairement l'interruption, la distorsion, ou la destruction de la communication, bien qu'il n'y ait aucune preuve de communication secrète.

2.2. Stéganalyse (Steganalysis)

La stéganalyse est l'art d'attaquer les méthodes de stéganographie pour la détection, l'extraction, destruction et manipulation des données cachées dans un médium stéganograhié. Les attaques peuvent être de plusieurs types par exemple, certaines attaques détectent simplement la présence de l'information cachée, certains essaient de détecter et d'extraire les informations cachées, certains essaient simplement de détruire les données cachées en trouvant l'existence sans essayer d'extraire des données cachées et certains essaient de remplacer les données cachées avec d'autres données en trouvant l'emplacement exact où les données sont cachées.

La détection suffit à déjouer le but même de la stéganographie même si le message secret n'est pas extrait car détecter l'existence de données cachées suffit si elles doivent être détruites.

La détection est généralement effectuée en identifiant certaines caractéristiques des images qui sont modifié par les données cachées. Un bon attaquant doit connaître les méthodes et les techniques de stéganographie pour les attaquer efficacement.

2.3. Classification des attaques en fonction des informations dont dispose l'attaquant

Les attaques des systèmes stéganographiques sont classées dans cinq classes [4] :

- **Attaque avec stégo-medium seul (Stego-only attack)**: Seul le stégo-medium est connu. L'insertion d'un message change certaines caractéristiques statistiques du cover-médium (*e.g.:* histogramme, égalité des cardinaux,…). L'attaque est basée sur cette altération.
- **Attaque avec cover et stégo medium (Known cover attack)**: Le medium de couverture et le stégo-medium sont disponibles. Ce type d'attaque est basé sur la comparaison entre le cover-médium et le stégo-médium (*e.g.* attaque visuelle).
- **Attaque sur message connu (Known message attack)**: Certaines parties du message caché sont connues de l'utilisateur. L'attaquant va essayer de retrouver dans le stégo-medium les parties du message qu'il connaît afin de faciliter l'analyse des documents futur. Même avec le message connu cette attaque est très difficile et généralement considérée comme équivalent à l'attaque stégo-only.
- **Attaque avec un algorithme choisis (Chosen stego attack)**: L'algorithme et le stégo- medium sont connus.
- **Attaque avec un message choisis (Chosen message attack)**: Le stéganalyste génère un stégo-medium à l'aide du message choisis. Le but est d'observer le résultat pour cracker l'algorithme.

2.4. Différentes approches de la stéganalyse

Plusieurs méthodes peuvent être utilisées pour la détection des images stéganographiées dont les principales :

- ***Attaques visuelles*** : L'insertion d'un message dans le dernier plan de bit peut se faire de façon aléatoire sur l'ensemble des pixels de l'image ou de façon séquentielle à partir de début de l'image. L'idée de cette attaque est basée sur le fait que une image peu texturée, le plan LSB est corrélé avec l'image d'origine. L'insertion du message perturbe le plan LSB en proportion de la taille de message. Les attaques visuelles appliquent des filtres sur l'image originale (figure 3.1) et l'image stéganographiée, supprimant les composantes les plus visibles (bits de poids forts) et renforçant les autres (bits de poids faible) [9].

Figure 2.1. : Image originale

La figure 3.2 représente le dernier plan de bit de l'image suivante, dégradé de 1280x960 pixels, avant et après insertion d'un message dans le plan LSB.

Le dernier plan de bit de l'image initiale montre une régularité qui correspond à la régularité de l'image initiale (figure 3.1). On remarque ainsi que l'image stéganographiée est très bruitée et laisse apparaître de façon claire la présence d'un message dans l'image.

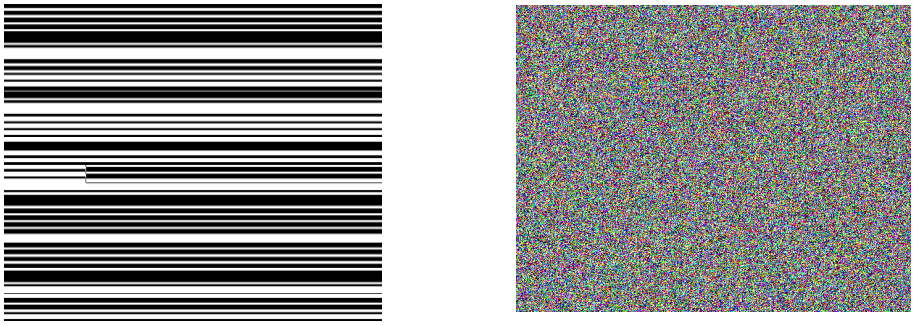


Figure 2.2.: Dernier plan de bit avant et après insertion de l'image dégradé

L'image Lina de la figure 3.3, a été choisie car c'est une image naturelle qui possède des zones homogènes assez grandes et en assez grand nombre. En comparant les images de la figure 3.4, on peut conclure que le même test ne montre aucun artéfact à l'œil.

Figure 2.3 : Image Lena originale

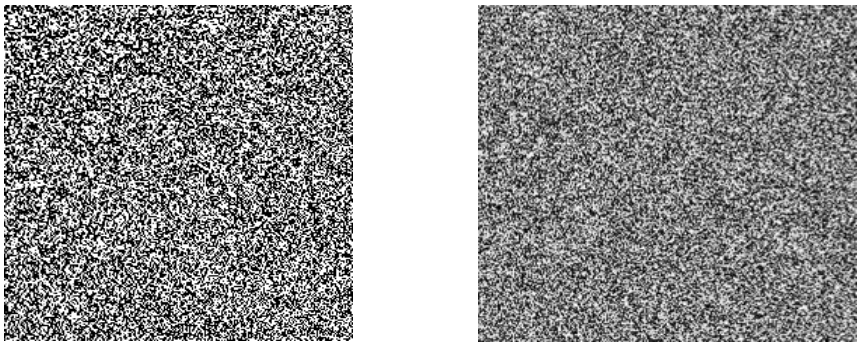


Figure 2.4.: Dernier plan de bit avant et après insertion de l'image Lena

- ***Attaques structurelles*** : le format du fichier de données change souvent car les données à masquer sont intégrées, l'identification de ces changements structurels caractéristiques peut détecter l'existence d'une image, par exemple dans la stéganographie basée sur la palette, la palette de l'image est modifiée avant d'intégrer les données pour réduire le nombre de couleurs de sorte que la différence de couleur des pixels adjacents soit très inférieure. Ceci montre que des groupes de pixels dans une palette ont la même couleur, ce qui n'est pas le cas dans les images normales [9].
- **Attaques statistiques**: Dans ce type d'attaques, la détection d'irrégularité statistique est une autre catégorie d'analyse. Elle se base sur la quantification de distorsion du média analysé, comparativement à des distributions statistiques théoriques représentant un média de base.

  Selon le type des mesures effectuées pour la distinction entre les images de couverture et les images stéganogrphiées, nous distinguons deux types de stéganalyse. La stéganalyse universelle et la stéganalyse spécifique.

2.5. Stéganalyse spécifique vs. Stéganalyse universelle

2.5.1. Stéganalyse spécifique

La stéganalyse ciblée, également appelée *spécifique*, a pour principe d'essayer de déterminer les faiblesses de sécurité d'un algorithme particulier, en étudiant son "*implémentation*" et/ou ses "*failles statistiques*", pour pouvoir identifier la présence d'un message caché, par cet algorithme, dans un médium donné. Pour ce faire, le stéganalyste, qui a connaissance au préalable de l'algorithme de dissimulation (il cible un algorithme stéganographique particulier), génère un ensemble de supports stéganographiés avec le même algorithme pour 1) comprendre et analyser les différentes étapes de l'algorithme, et 2) comparer la statistique

des images de couverture qu'il a à sa disposition avec celles qui ont été générées. À travers cette opération, le stéganalyste tente de déterminer les points caractérisant ainsi que les faiblesses de l'algorithme ciblé, pour pouvoir discriminer les images *stego* des images *cover*. On peut donc dire que la stéganalyse ciblée se base sur l'identification des caractéristiques spécifiques, qui distinguent un algorithme stéganographique donné des autres algorithmes [14].

2.5.2. Stéganalyse universelle

La Stéganalyse statistique universelle ne nécessite en principe aucune information a priori sur les méthodes de stéganographie utilisées pour la détection du message caché. A cet effet, elle s'appuie sur un processus d'apprentissage utilisant des images originales et des images stéganographiées dont le résultat sera exploité par un processus de classification. Les réseaux de neurones ou des algorithmes de classification standards peuvent être utilisés pour construire le modèle de détection à partir des données expérimentales [14].

2.6. Travaux connexes

L'histoire de la stéganalyse d'images commence avec les premières études proposées par Johnson et Jajodia [15] et Chandramouli et al. [16]. La stéganalyse d'images a parcouru un long chemin en commençant par la stéganalyse visuelle et l'extraction manuelle des caractéristiques jusqu'à l'utilisation de l'apprentissage en profondeur et de l'extraction automatique des caractéristiques.

Au début, les chercheurs ont essayé de trouver une signature ou un motif pour détecter des techniques stéganographiques spécifiques bien connues [16]; cependant, ce type n'a que des applications limitées. Avec l'évolution et la variété des techniques de stéganographie, des techniques de stéganalyse robustes sont devenues plus nécessaires. De nombreuses techniques de stéganalyse ont commencé à extraire des caractéristiques statistiques qui peuvent refléter des changements invisibles dans le support numérique. Par exemple, Chaeikar et al. [17] ont proposé une technique de stéganalyse statistique aveugle pour détecter la stéganographie d'image par remplacement du bit le moins significatif (LSB). Les auteurs ont constaté que l'harmonie naturelle des couleurs du pixel est affectée lors de l'intégration des données. Par conséquent, une caractéristique statistique qui analyse la corrélation des couleurs est extraite des pixels de l'image pour détecter l'existence du message secret. Dans un premier temps, les pixels ont été classés en trois classes en fonction de la similarité des couleurs avec les pixels voisins, et le niveau de suspicion des pixels a été identifié en fonction de la moyenne et de

l'écart type. Cela conduit à un ensemble de données utilisé pour former SVM pour détecter et estimer la longueur du message intégré.

Dans [18], ont proposés une méthode de stéganalyse pour détecter la stéganographie à spectre étalé. Le cœur de leur méthode consiste à découvrir les matrices de message dans le cover et stego médium en utilisant une méthode bien connue des moindres carrés. La matrice porteuse est initialisée de manière aléatoire, puis les matrices porteuse et de message sont mises à jour selon une méthode de descente de gradient univariée. Leur technique est basée sur les travaux de Li et al. [19], où le but est de réduire la complexité des calculs et de ne s'appuyer sur aucune connaissance préalable du nombre porteuses à spectre étalé. Par conséquent, la technique proposée extrait consécutivement les bits de données de chaque porteuse en extrayant la variance pour réduire le coût de calcul. Pour détecter et estimer le nombre de messages intégrés sans connaissance préalable, la technique proposée vise à atteindre la perturbation de l'image stéganographiée résiduelle au minimum en réduisant la variance de la stégo-image résiduelle.

Dans[20], les auteurs ont proposé une technique de stéganalyse universelle, où la loi de Zipf [21] est exploitée pour extraire les caractéristiques de la transformée en ondelettes. L'idée de base de la loi de Zipf dans la représentation des images comprend trois phases. La première phase est la sélection taille de masque pour compter la fréquence d'apparition des motifs. La deuxième phase consiste à minimiser le nombre de motifs en identifiant des coefficients d'ondelettes significatifs, ce qui conduit à une distribution plus significative pour la fréquence des motifs. Dans la troisième phase, la courbe Zipf est produite, qui représente la fréquence du motif et le nombre d'axes du motif. Enfin, les caractéristiques Aire sous la courbe de Zipf, Point d'inflexion et Métrique de similarité automatique de sous-bande) sont extraites de la courbe Zipf produite. Pour détecter les images stego, le classificateur de forêt aléatoire est formé à l'aide de l'ensemble de données UCID

Une nouvelle technique de stéganalyse qui vise à réduire la compléxité de calcul et de temps tout en haute performance est proposée par Guttikonda et Sridevi [22]. Chaque transformation de Walsh Hadamard basée sur un coefficient et chaque matrice de cooccurrence de niveaux de gris est utilisée pour extraire les caractéristiques des domaines de transformation et spatiaux, respectivement. Pour réduire la dimensionnalité des caractéristiques et sélectionner les caractéristiques les plus pertinentes, l'algorithme Pine Growth Optimization a été appliqué. Enfin, les caractéristiques sélectionnées sont utilisées pour former le classificateur Cross Integrated Machine Learning afin de distinguer les images de couverture et stego. Les résultats de l'expérience ont montré l'efficacité de la technique proposée en termes de

précision de détection et de temps d'exécution, où elle a réduit le temps d'environ par rapport à la technique Multi-SVM existante.

L'apprentissage très profond et l'extraction automatique de caractéristiques sont appliqués dans les travaux de Wu et al. [23]. Plus précisément, un nouveau modèle CNN appelé Deep Residual learning Network (DRN) est proposé pour la stéganalyse d'images. Les auteurs ont prouvé que le réseau de neurones très profond qui contient de nombreuses couches peut refléter des propriétés statistiques complexes, ce qui conduit à une distinction plus efficace des images stéganographiées. L'idée principale de leur technique est d'alimenter le réseau avec des composantes de bruit de l'image, au lieu de l'image originale pour forcer le réseau à considérer le signal faible produit par l'intégration des données. Par la suite, DRN est formé pour apprendre les caractéristiques efficaces des images de couverture et stego. Pour la classification binaire, une couche entièrement connectée avec un classificateur softmax a été réalisée. Les résultats expérimentaux menés à l'aide de l'ensemble de données BOSS base ont montré la supériorité de la technique proposée par rapport à d'autres techniques basées sur les réseaux de neurones profonds.

Une autre technique basée sur les réseaux de neurones profonds qui extrait les caractéristiques de plusieurs domaines est proposée par Wang et al. [24]. Tout d'abord, deux méthodes de stéganalyse célèbres sont simulées, à savoir le modèle riche en espace SRM et le résidu DCT pour détecter les caractéristiques de stéganographie dans les domaines spatial et transformé. À l'étape suivante, les caractéristiques linéaires précédentes avec des caractéristiques SRM non linéaires sont transmises à la couche CNN pour extraire les caractéristiques générales. Enfin, la couche entièrement connectée est appliquée pour la classification des images stego et de couverture. Grâce aux expériences, les auteurs ont prouvé l'efficacité de la prise en compte de l'extraction des caractéristiques non linéaires ainsi que de l'extraction des caractéristiques de plusieurs domaines, où la précision de détection est augmentée de 0,3 à 6 % et de 2 à 3 %, respectivement.

Tan et.al ont utilisé un réseau CNN avec quatre couches de convolution pour la stéganalyse d'images [25]. Leurs expériences ont montré qu'un CNN avec des poids initialisés aléatoires ne peut généralement pas converger et initialiser les poids de la première couche avec le noyau KV peut améliorer la précision. Qian et al.[26] propose une modèle stéganalyse utilisant l'architecture CNN standard avec la fonction d'activation Gaussien, et a en outre prouvé que l'apprentissage par transfert est bénéfique pour un modèle CNN pour détecter un algorithme de stéganographie avec de faibles taux d'insertion. Les performances de ces schémas sont comparables ou meilleur que la technique SPAM, mais sont faible que le

technique SRM. Xu et al.[27] proposé une structure CNN avec certaines techniques utilisées pour la classification des images, comme la normalisation par lots, 11 convolution, et mise en commun avec le pooling moyenne. Ils ont également effectué un prétraitement avec un filtre passe-haut avec une couche d'activation absolue (ABS). Leurs expériences ont montré de meilleures performances. En améliorant le Xu-CNN, ils ont atteint une performance plus stable. Dans le domaine JPEG, Xu et al.[28] proposé un réseau basé sur image décompressée et obtenu une meilleure précision de détection que les méthodes traditionnelles dans le domaine JPEG. En simulant le schéma de stéganalyse traditionnel des éléments fabriqués à la main, Fridrich et al.[30] a proposé une structure CNN avec histogramme couches, qui est formé par un ensemble de fonctions d'activation gaussienne. Ye et al.[31] a proposé une structure CNN avec un groupe de filtres passe-haut pour le prétraitement et adopté un ensemble de fonctions d'activation hybrides pour mieux capturer l'enrobage signaux. Avec l'aide de la connaissance et des données du canal de sélection augmentation, leur modèle a obtenu des performances significatives améliorations que le SRM classique. Fridrich [32] a proposé une architecture réseau différente pour faire face à la stéganalyse images de taille arbitraire par extraction manuelle de caractéristiques. Leur schéma entre les éléments statistiques des cartes d'entités dans le classificateur de réseau profond.

2.7 Conclusion

Nous avons présenté dans ce chapitre le domaine de la stéganalyse et les différents types de stéganalyse en l'occurrence la stéganalyse spécifique et universelle. Un état de l'art sur les différentes techniques de stéganalyse est ensuite présenté. Nous avons constaté que les méthodes de stéganalyses proposées récemment sont basées sur l'apprentissage en profondeur. Alors peu sont les méthodes utilisant l'apprentissage basé sur l'ensemble classifieur. C'est la raison pour laquelle, nous allons présenté dans le chapitre suivant une méthode de stéganalyse basée sur le classifieur XGBoost.

Chapitre 3

CONCEPTION ET RÉSULTATS EXPÉRIMENTAUX

Introduction

Les performances des méthodes de stéganalyses dépendent généralement du choix de la méthode d'extraction de caractéristiques, les hyperparamètres du classifieur et les méthodes de stéganographie. Dans ce chapitre, nous proposerons une méthode de détection d'informations cachées dans les images JPEG basé la méthode d'extraction de caractéristique HOG et l'ensemble classifieur XGBoost.

3.2 Système proposé

La Figure .III.1 montre la l'architecture de la méthode proposée pour la détection des informations cachées dans les images JPEG. Il est clair que la création du modèle de détection est basée sur l'étape d'extraction des caractéristiques, et la classification.

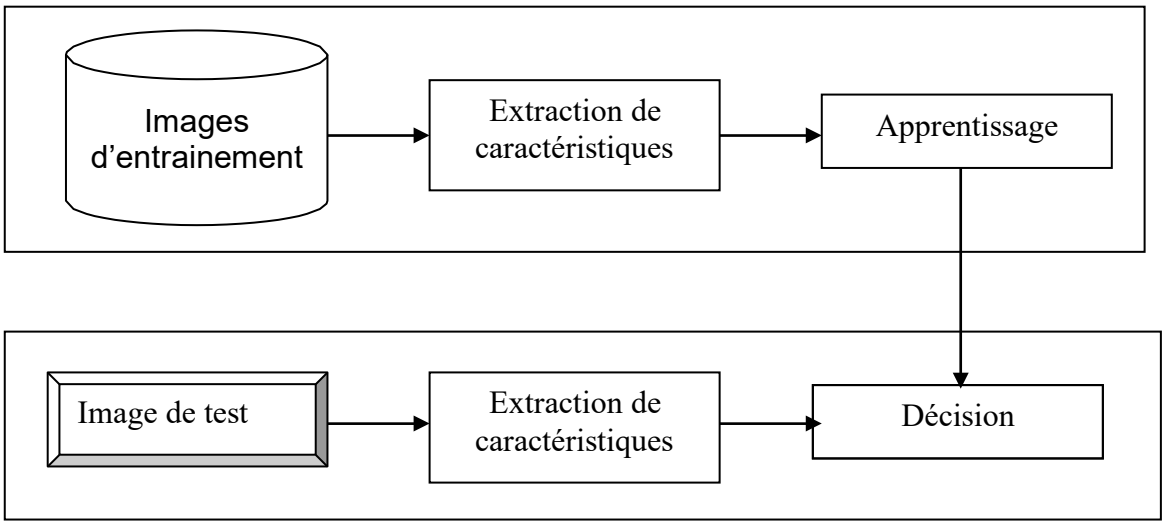


Fig. 3.1: Etape de la méthode proposée

3.2.1 La méthode d'extraction de caractéristiques HOG

La méthode d'extraction de caractéristiques HOG est une nouvelle méthode utilisée en vision par ordinateur pour la détection d'objet et la détection des régions d'intérêts. Le principe de base de cette méthode est de calculer les histogrammes locaux de l'orientation du gradient sur une grille dense, c'est-à-dire sur des zones régulièrement réparties sur l'image.

L'objectif du de la méthode HOG est que l'apparence et la forme locale d'un objet dans une image peuvent être décrites par la distribution de l'intensité du gradient ou la direction des contours. La figure III.2, illustre les différentes étapes de la méthode HOG. Le principe de cette méthode consiste à diviser une image à des régions adjacentes de petite taille, appelées cellules, et en calculant pour chaque cellule l'histogramme des directions du gradient ou des orientations des contours pour les pixels à l'intérieur de cette cellule. La combinaison des histogrammes forme alors le descripteur HOG. Pour de meilleurs résultats, les histogrammes locaux sont normalisés en contraste, en calculant une mesure de l'intensité sur des zones plus larges que les cellules, appelées des blocs, et en utilisant cette valeur pour normaliser toutes les cellules du bloc. Cette normalisation permet une meilleure résistance aux changements d'illuminations et aux ombres.

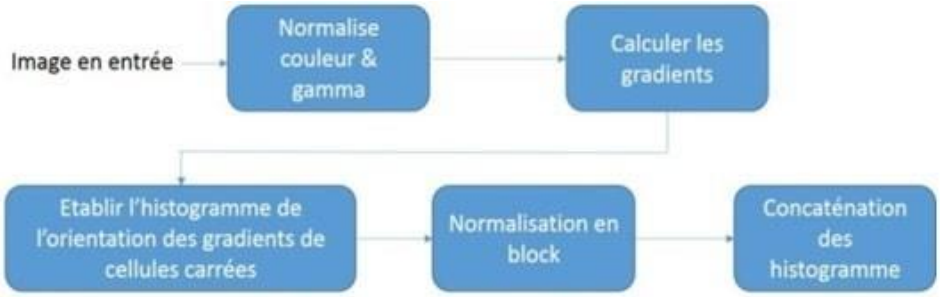


Fig. 3.2 : Schéma des blocs de méthode HOG

3.2.2. Apprentissage d'ensemble classifieur: XGBoost

XGBoost est en fait une version particulière de l'algorithme de Gradient Boost. En effet, il s'agit d'un assemblage de "weak learners" qui prédisent les résidus, et corrigent les erreurs des "weak learners" précédents.

La particularité d'XGBoost réside dans le type de "weak learner" utilisé. Les "weak learners" sont des arbres décisionnels. Les arbres qui ne sont pas assez bons sont "élagués", c'est à dire qu'on leur coupe des branches, jusqu'à ce qu'ils soient suffisamment performant. Sinon ils sont complètement supprimés. Cette méthode est appelé le "pruning" (élagage).

Ainsi, XGBoost s'assure de ne conserver que de bons weak learners. De plus, XGBoost est informatiquement optimisé pour rendre les différents calculs nécessaires à l'application d'un Gradient Boosting rapides.

Enfin, XGBoost propose un panel d'hyperparamètres très important. Il est ainsi possible grâce à cette diversité de paramètres, d'avoir un contrôle total sur l'implémentation du Gradient Boosting. Pour toutes ces raisons, XGBoost est souvent l'algorithme gagnant des

compétitions Kaggle, il est rapide, précis et efficace, permettant une souplesse de manœuvre inédite sur le Gradient Boosting.

3.3 Résultats expérimentaux

L'évaluation d'une méthode de stéganalyse revient à déterminer si une image est stéganographiée ou non. La présente section porte sur différentes expérimentations réalisées dans le but d'évaluer la performance de la méthode développée.

3.3.1 Base d'images

La base de données ALASKA est l'ensemble des images de l'Alaska Kaggle compétition [35]. Cette base de données offre un ensemble de données beaucoup plus important de 80 000 images, provenant de plus de 40 caméras (Smartphones, tablettes, appareils photo de bas gamme et DLSR plein format haut de gamme) et traitées de manière réaliste et très hétérogène. Dans nos expériences, nous avons sélectionné 50 000 images au hasard et nous avons les redimensionnées avec une taille de 512 × 512 pixels, puis convertis au niveau de gris.

3.3.2 Construction de la base des images stéganographiées

Afin de construire les images stéganographiées, nous avons utilisé dans ce travail deux méthodes de stéganographies les plus connues en termes de la difficulté de détection: J-UNIWARD et UERD.

A. Méthode de stéganographie J-UNIWARD

J-Uniward [33], pour UNIversal Wavelet Relative Distortion est une méthode de modélisation de la distorsion stéganographique causée par les données insérées. Elle vise à fournir une fonction qui détermine quelles régions de l'image de couverture sont moins prévisibles et plus difficiles à modéliser. Les changements introduites lors l'insertion des données cachées dans ces zones sont plus difficile à détecter que s'elles sont introduites uniformément l'image de couverture. En détectant ces zones prévisibles et imprévisibles, cette méthode fournit un moyen de déterminer où les changements d'insertion seraient les moins perceptibles. Cette méthode est couplée avec un schéma de codage, tel que le codage en treillis de syndrome (STC), pour créer un algorithme de stéganographie données de contenu adaptatif.

B. Méthode de stéganographie UERD

UERD [34] est un schéma d'insertion stéganographique qui vise à minimiser la probabilité détection de la présence d'informations cachées, en minimisant l'impact de l'insertion sur les paramètres statistiques de l'image de couverture. Il atteint ceci par l'analyse des paramètres des coefficients DCT, ainsi que des blocs DCT et leurs voisins. Grâce à cela, la méthode peut déterminer si la région peut être considérée comme bruitée et si l'insertion aura un impact sur les caractéristiques statistiques tels que les histogrammes de l'image de couverture. Les régions non texturées sont celles où les paramètres statistiques sont prévisibles et où l'insertion entraînerait des changements notables. Le schéma ne fait pas exclure l'utilisation de valeurs telles que les coefficients DC ou les coefficients DCT nuls lors de l'insertion, car leurs profils statistiques peuvent les rendrais appropriés à l'insertion.

3.3.3 Protocole de tests

Pour la création des images stéganographiées, nous avons construit deux bases d'images de 50000 images avec les méthodes de stéganographie décrites ci-dessus (J-UNIWARD et UERD) et avec différents taux d'insertion. Dans le cadre de l'expérimentation menée dans la phase d'apprentissage, nous avons utilisé une base de données contenant 35000×3 images (35000 images de couverture et 70000 images stéganographiées).

3.3.4 Résultats

A. Sélection des paramètres de l'algorithme HOG

La méthode HOG [36] dépend de deux paramètres importants qui sont la taille de bloc et le pourcentage de chevauchement entre les blocs adjacents. Ce que nous intéressent c'est bien la taille de bloc (plusieurs travaux montrent que un pourcentage de chevauchement égale à 50% est suffisant pour que l'algorithme fonctionne efficacement). Pour avoir des meilleurs paramètres (paramètres optimaux), nous avons testé l'algorithme HOG avec des blocs de différents tailles, par exemple 3×3, 5×5,… et avec le même pourcentage de chevauchement et le classifieurs XGBoost avec ses paramètres par défaut. L'objectif de la variation dans la taille des blocs est d'avoir des meilleurs résultats en fonction du taux d'erreur de détection et de la taille de vecteur de caractéristiques.

Le tableau III.1 montres que notre méthode de proposée d'une façon générale est efficace, avec un taux de détection **93.20 %** pour un bloc de taille 9×9. Il est clair que ces paramètres sont très influencés par la taille de bloc. L'augmentation de ce dernier provoque des mauvaise performances vis-à-vis la taille de vecteur de caractéristiques.

Tableau 3.1:Résultats d'exécution du HOG avec XGBoost pour la détection de J-UNIWARD

W*W	Accuracy	Taille Vecteur de caractéristiques
3*3	0.887	81
5*5	0.906	225
7*7	0.915	441
9*9	**0.932**	729
11*11	0.927	1089
13*13	0.921	1521
15*15	0.907	2025
17*17	0.864	2601
19*19	0.831	3249

Nous avons utilisé la précision (Accuracy) comme une métrique de performance pour tester les performances de notre méthode de stéganalyse proposée. En effet, nous l'avons définie selon les valeurs de TP (True Positive), FN (False Negative), TN (True Negative), et FP (False Positive) ; où est le nombre de vrais positifs, ce qui signifie que certains images stéganographiées sont correctement classés comme images stéganographiées avec notre méthode ; FN est le nombre de faux négatifs, ce qui signifie que certains images stéganographiées sont classés comme des images non stéganographiées avec notre méthode; TN est le nombre de vrais négatifs, ce qui signifie que certains images non stéganographiées sont correctement classés comme images non stéganographiées avec notre méthode; et FP est le nombre de faux positifs, ce qui signifie que certains images non stéganographiées sont classés comme images stéganographiées avec notre méthode.

Les figures III.3 et III.4, présente les courbes ROC pour la détection des méthodes de stéganographie J-UNIWAED et UERD qui nous permet de calculer la précision à partir des valeurs des Faux positif et vrai positif

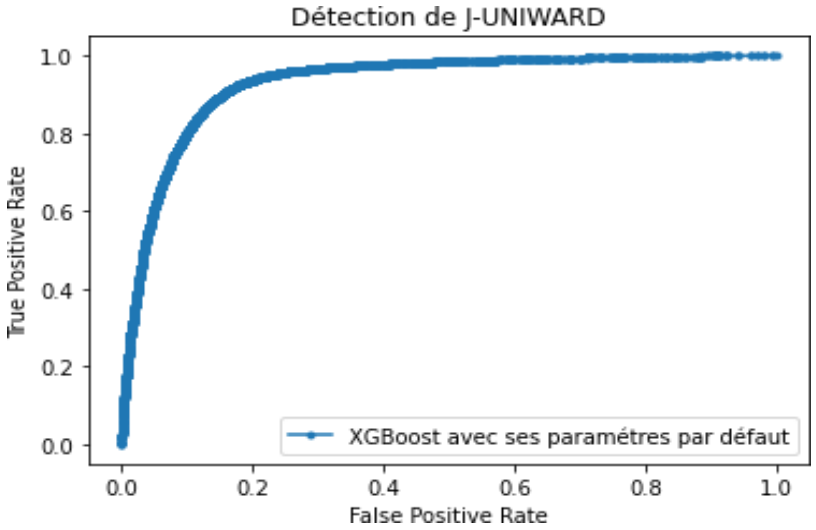


Fig. 3.3 : Courbe ROC de la méthode proposée: Détection de la méthode de stéganographie J-UNIWARD

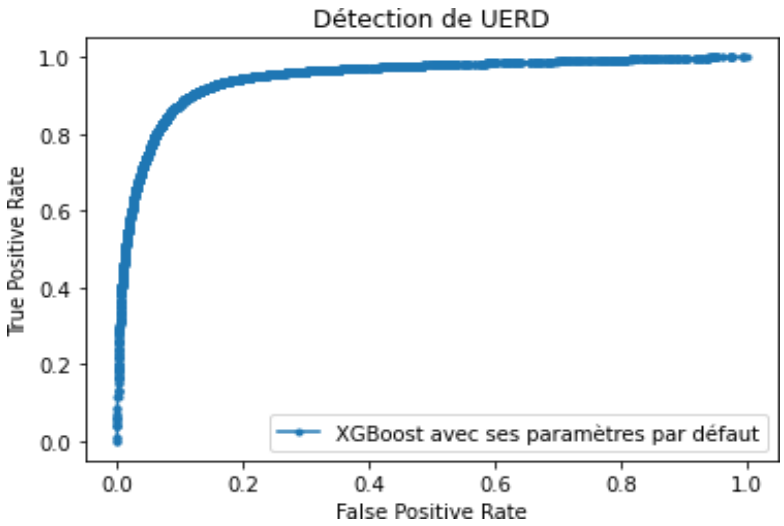


Fig. 3.4 : Courbe ROC de la méthode proposée: Détection de la méthode de stéganographie UERD

Nous pouvons dire que la méthode proposée est généralement est efficace, atteignant une précision de classification de **93.2** %, et de **94.70** % pour la détection des méthodes de stéganographie J-UNIWAED et UERD, respectivement.

B. Sélection des Hyperparamètres de XGBoost

Dans le but d'améliorer les performances de la méthode proposée, nous allons, dans cette section, jouer sur les hyperparamètres du classifieur XGBoost.

Le réglage ou l'optimisation des hyperparamètres consiste à choisir un ensemble d'hyperparamètres approprié pour un algorithme d'apprentissage automatique. C'est une tâche très importante dans tout cas d'utilisation de Machine Learning. Ces paramètres doivent être spécifiés manuellement à l'algorithme et fixés via une passe de formation. Dans les modèles basés sur des arbres, les hyper-paramètres incluent des éléments tels que *la profondeur maximale de l'arbre,* le *nombre d'arbres*, *le nombre de variables* à prendre en compte lors de la construction de chaque arbre*, le nombre minimum d'échantillons sur une feuille*, la fraction d'observations utilisé pour construire un arbre, et quelques autres.

Dans nos expérimentation, dans un premier temps, nous avons crées un dictionnaire de certains paramètres sur lesquels nous avons entrainé le classifieur XGBoost. Les clés de ce dictionnaire sont essentiellement les paramètres et les valeurs sont une liste de valeurs des paramètres sur lesquels s'entraîner (Tableau III.2).

Tableau 3.2: Hyperparamètres de XGBoost

Clés	Valeurs			
Nombre d'arbre	500	1000	1500	
Taux d'apprentissage	0.05	0.1	0.15	0.2
Profondeur	2	4	6	8
Boosting	gblinear		gbtree	

Ensuite, nous avons utilisés *Randomized SearchCV* qui teste chaque valeur de ces Hyperparamètres (Tableau III.3) afin de trouver les paramètres qui donnent le meilleur taux de détection (conduisant à **480** combinaisons d'apprentissage).

Tableau 3.3: Meilleur Hyperparamètres de XGBoost

Nombre d'arbre	1500
Taux d'apprentissage	0.15
Profondeur	2
Boosting	Gbtree

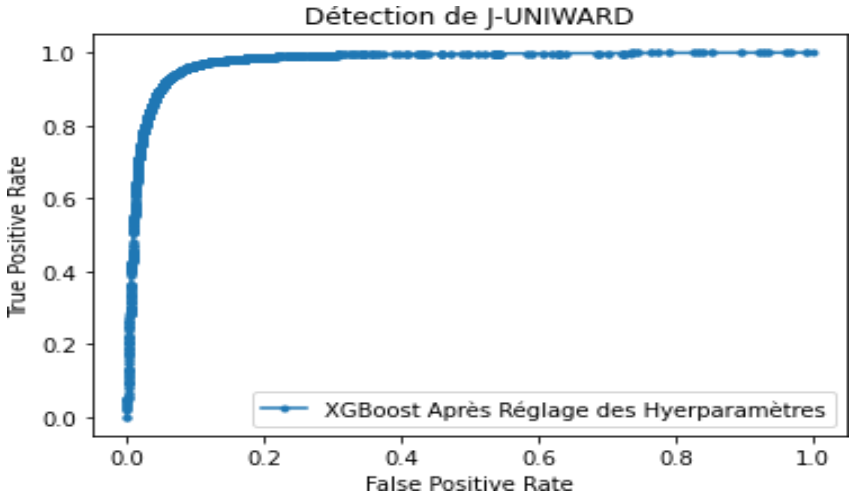


Fig. 3.5 : Courbe ROC de la méthode proposée: Détection de la méthode de stéganographie J-UNIWARD

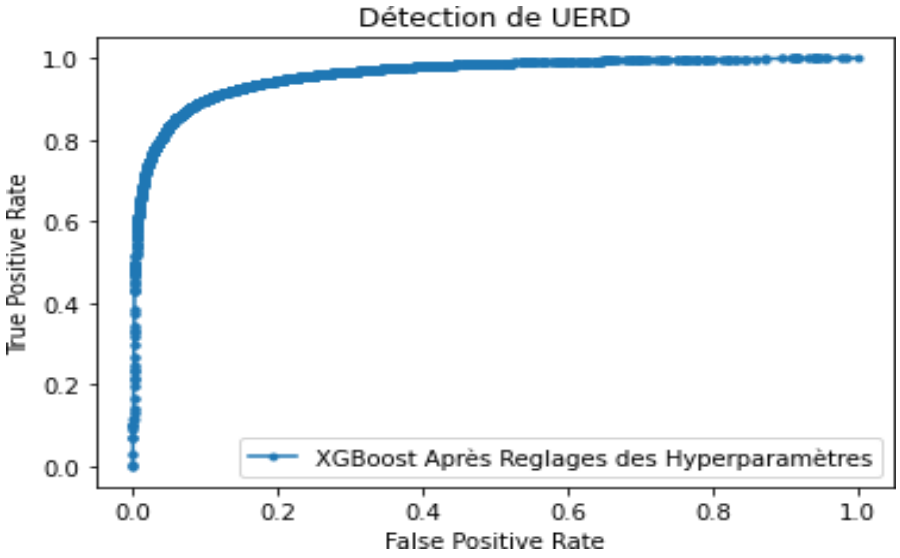


Fig. 3.6: Courbe ROC de la méthode proposée: Détection de la méthode de stéganographie UERD

Les figures III.5 et III.6 montrent que le réglage des hyperparamètres du classifieur XGBoost conduit à une amélioration importante de la précision en particulier pour la détection de la méthode de stéganographie J-UNIWARD atteignant une précision de classification de **97.2** %, et de **95.30** % pour la détection de la méthode de stéganographie UERD, respectivement.

3.4 Conclusion

Nous avons proposé une méthode de stéganalyse pour la détection des informations cachées dans les images Jpeg. Notre méthode repose sur la méthode d'extraction de caractéristique HOG et une méthode d'ensemble classifieur : XGBoost en raison de son extrême sensibilité aux ses hyper paramètres.

Les expérimentations réalisées montrent que les performances d'un système basé sur l'apprentissage automatique peuvent être améliorées soit par la configuration de la méthode d'extraction de caractéristiques soit par le réglage des hyperparamètres du classifieur.

Conclusion

Afin d'assurer les quatre fonctions de la sécurité de l'information (confidentialité, intégrité, non-répudiation, et traçabilité), trois techniques peuvent être utilisées: la dissimulation de l'information, la cryptographie et la biométrie.

De nos jours, la stéganographie a connu un grand succès afin d'assurer le la protection du transfert de l'information. En fait, plusieurs systèmes stéganographiques ont été développés et jugés efficaces vue de leurs haute performances dans l'étape d'insertion. Malheureusement, ces systèmes peuvent être utilisés d'une façon illégale : vol de l'information et le transfert des informations jugées illégales.

L'utilisation malveillante de la stéganographie conduit à l'émergence de sa contre partie la stéganalyse et qui consiste à la détection de la présence de l'information cachée, mésuser la taille de l'information cachée ou sa destruction.

Dans ce travail, nous avons proposé une méthode de stéganalyse pour la détection des informations cachées dans les images Jpeg. Notre méthode repose sur la méthode d'extraction de caractéristique HOG et une méthode d'ensemble classifieur : XGBoost en raison de son extrême sensibilité aux ses hyper paramètres.

Les expériences ont été réalisées sur une base de données contenant 50000 images de couverture et deux bases d'images stéganographiées. Les résultats expérimentaux ont montré un taux de détection élevé, qui peut également être amélioré en changeant la méthode d'extraction de caractéristique. Les travaux futurs de cette étude se concentreront sur l'utilisation d'autres techniques d'apprentissage en profondeur comme DCTNET et ICANET.

Bibliographies

[1] Barbier J., Filiol E., and Mayoura K., (2006), *New Features for Specific JPEG Steganalysis.* In Proceedings 3rdInternational Conference on Computer, Information, and Systems Science, and Engineering, CISE.

[2] Barbier J., (2007), *Analyse de Canaux de Communication dans un Contexte non Coopératif*, Thèse pour obtenir le grade de docteur, ESAT - Laboratoire de Virologie et Cryptologie, B.P. 18, 35 998 Rennes Cedex..

[3] Johnson N.; (2002), *Survey of Steganography Software*, Technical Report.

[4] Raynal F., (2002), *Etude d'outils pour la Dissimulation d'Information*, Thèse pour obtenir le grade de docteur, Université paris XI.

[5] Raynal F., Fabien A.P. Petitcolas and Caroline F. (2002), *L'art de dissimuler les informations*. Pour la Science, 36 : 26–31.

[6] Cheikh LO, La Stéganographie Appliquée aux Textes & Images, Travail de diplôme, IUP de Rouen- Institut Universitaire Professionnalisé, 2008.

[7] Fabian Galand, Stéganaographie, Traité de Sécurité des systèmes d'information, Techniques de l'Ingénieur, Ch H 5870, 2004.

[8] www.noxistes.org/manipuler_le_slack_space_4.php

[9] Laimeche L.., (2018), *Stéganalyse universelle des images JPEG*, Thèse pour obtenir le grade de docteur, Annaba, 2018.

[10] Christian Rey (2003), Tatouage d'Image: Gain en Robustesse et Intégrité des Images, Thèse pour obtenir le grade de docteur, Université d'Avignon, 2003.

[11] Christian Rey (3003), Tatouage d'Image: Gain en Robustesse et Intégrité des Images, Thèse pour obtenir le grade de docteur, Université d'Avignon, 2003.

[12] Jean-Pierre CLUTIER (2002), Cryptographie et certification, Travail de diplôme, Conservatoire National Des Arts et Métiers (CNAM) ,2002.

[13] Jenny Dentand, (2005) Stéganographie, travail de diplôme, Haute Ecole de Gestion de Genève (HEG-GE), 2005.

[14] Ge, Y.; Zhang, T.; Liang, H.; Jiang, Q.; Wang, D. A (2021), Novel Technique for Image Steganalysis Based on Separable Convolution and Adversarial Mechanism. Electronics 2021, 10, 2742. https://doi.org/ 10.3390/electronics10222742

[15] Johnson, N.F.; Jajodia, S. (1998) Steganalysis of images created using current steganography software. In International Workshop on Information Hiding; Springer: Berlin/Heidelberg, Germany, 1998; pp. 273–289.

[16] Chandramouli, R.; Li, G.; Memon, N.D. (2002) Adaptive steganography. In Security and Watermarking of Multimedia Contents IV; International Society for Optics and Photonics: Bellingham, WA, USA, 2002; Volume 4675, pp. 69–78.

[17] Chaeikar, S.S.; Zamani, M.; Manaf, A.B.A.; Zeki, A.M. (2018) PSW statistical LSB image steganalysis. Multimed. Tools Appl. 2018, 77, 805–835.

[18] Soltanian, M.; Ghaemmaghami, S. (2017) Blind consecutive extraction of multi-carrier spread spectrum data from digital images. In Proceedings of the 2017 Iranian Conference on Electrical Engineering (ICEE), Tehran, Iran, 2–4 May 2017; pp. 1835–1839.

[19] Li, M.; Kulhandjian, M.K.; Pados, D.A.; Batalama, S.N (2013).; Medley, M.J. Extracting spread-spectrum hidden data from digital media. IEEE Trans. Inf. Forensics Secur. 2013, 8, 1201–1210.

[20] Laimeche, L.; Merouani, H.F.; Mazouzi, S (2018). A new feature extraction scheme in wavelet transform for stego image classification. Evol. Syst. 2018, 9, 181–194.

[21] Wilson, L. Zipf, George K (2017): Human Behavior and the Principle of Least Effort; Addison Wesley: New York, NY, USA, 1949.

[22] Guttikonda, J.B.; Sridevi, R. (2019). A new steganalysis approach with an efficient feature selection and classification algorithms for identifying the stego images. Multimed. Tools Appl. 2019, 78, 21113–21131.

[23] Wu, S.; Zhong, S.; Liu, Y (2018). Deep residual learning for image steganalysis. Multimed. Tools Appl. 2018, 77, 10437–10453.

[24] Wang, Z.; Chen, M.; Yang, Y.; Lei, M (2020).; Dong, Z. Joint multi-domain feature learning for image steganalysis based on CNN. EURASIP J. Image Video Process. 2020, 2020, 1–12.

[25] S. Tan and B. Li (2014) , "Stacked convolutional auto-encoders for steganalysis of digital images," in Proceedings of Signal and Information Processing Association Annual Summit and Conference, APSIPA 2014, Siem Reap, Cambodia, Dec. 2014, pp. 14.

[26] Y. Qian, J. Dong, W. Wang and T. Tan (2017), "Feature learning for steganalysis using convolutional neural networks." Multimedia Tools and Applications, 1-25, 2017(2).

[27] G. Xu, H.-Z. Wu, and Y.-Q. Shi, (2016) "Structural design of convolutional neural networks for steganalysis," IEEE Signal Process. Lett., vol. 23, no. 5, pp. 708712, May 2016.

[28] G. Xu, (2017), "Deep Convolutional Neural Network to Detect J-UNIWARD," in Proceedings of the 5th ACM Workshop on Information Hiding and Multimedia Security, Drexel University in Philadelphia, PA, June 2017, IH&MMSec17, pp. 6773.

[29] K. He, Xu. Zhang, S. Ren and J. Sun, (2015) "Spatial Pyramid Pooling in Deep Convolutional Networks for Visual Recognition," in IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 37, no. 9, pp. 1904- 1916, Sept. 1 2015.

[30] V. Sedighi and J. Fridrich, (2017) "Histogram layer, moving convolutional neural networks towards feature-based steganalysis," in Proc. Media Watermarking, Security, and Forensics, Part of IS&T International Symposium on Electronic Imaging (EI2017), 2017, pp. 5055.

[31] J. Ye, J. Ni and Y. Yi, (2017) "Deep Learning Hierarchical Representations for Image Steganalysis," in IEEE Transactions on Information Forensics and Security, vol. 12, no. 11, pp. 2545-2557, Nov. 2017.

[32] J. Fridrich and J. Kodovsky, (2012) "Rich models for steganalysis of digital images," IEEE Trans. Inf. Forensics Security, vol. 7, no. 3, pp. 868882, June 2012.

[33] Holub, V.; Fridrich, J.; Denemark, T. (2014) Universal distortion function for steganography in an arbitrary domain. EURASIP J. Multimed. Inf. Secur. 2014.

[34] L. Guo, J. Ni, W. Su, C. Tang, and Y.-Q. Shi, (2015) “Using statistical image model for JPEG steganography: uniform embedding revisited,” IEEE Transactions on Information Forensics and Security, vol. 10, no. 12, pp. 2669–2680, 2015.

[35] R. Cogranne and P. Bas. (2019) ”Alaska”. https://alaska.u.fr., March 2019

[36] D. Maltoni, D. Maio, A. K. Jain, and S. Prabhakar, \Handbook of Fingerprint Recognition", Second Edition, Springer, 2009.