An Investigation for Steganalysis in Color Images

Samah F. Aziz
samah.fakhri@uohamdaniya.edu.iq
AlHamdaniya University, Mosul, Iraq

Ahmed S. Nori
ahmed.s.nori@uomosul.edu.iq
College of Computer Science and Mathematics, University of Mosul, Iraq

Received on: 22/06/2011 **Accepted on: 16/08/2011**

ABSTRACT

With science developing and techniques used in Information hiding, there are another techniques wall together for Steganalysis.

Steganography is considered as the new and the complementary system of Cryptography that took a long time in transferring secret and important messages through the networks Internet. Then there was the emergence of what complements Steganography as a science that analysis and discover the content of the secret messages and this science is Steganalysis.

This study tackled and manifested the ideas of analysis processes that can be followed to interpret the secret messages and discovering them either by means of knowing about their existence only or the capability of extracting them in full.

The work relied on two important technologies; the first is called the Support Vector Machine (SVM) and the second is called Fisher Linear Discriminator (FLD). The SVM technology has been used with the blind application idea while FLD has been used with the blind and non-blind application ideas using colored images which are PNG and BMP.

Results proved the high efficiency of the two technologies in detecting the image that includes the secret messages and comparisons were varied between the two technologies in terms of detection rate, fault and the execution time.

Keywords: Steganography, Steganalysis, SVM, FLD, PNG, BMP.

التقصي حول الكشف عن الإخفاء في الصور الملونة

أحمد سامي نوري **سماح فخري عزيز**

كلية علوم الحاسوب والرياضيات / جامعة الموصل

تاريخ استلام البحث: 2011/06/22 **تاريخ قبول البحث: 2011/08/16**

الملخص

مع تطور العلم والتقنيات المستخدمة في تقنيات إخفاء المعلومات (Information Hiding)، هنالك تقنيات تتطور بموازاتها في تحليل الإخفاء (Steganalysis).

يعد الإخفاء (Steganography) النظام الجديد والمتمم لعلم التشفير الذي استغرق وقتا طويلا في تناقل الرسائل السرية والمهمة عبر الشبكات والانترنيت. ومن ثم ظهر ما يقابل الإخفاء من علم يعمل على تحليل وكشف محتوى الرسائل السرية فكان علم كشف الإخفاء (Steganalysis).

اعتمد العمل على تقنيتين مهمتين في عمليات التصنيف تسمى الأولى آلة المتجه الداعم SVM (Support Vector Machine) والثانية مميز فيشر الخطي FLD (Fisher Linear Discriminator). حيث تم استخدام تقنية SVM مع فكرة التطبيق الأعمى، في حين تم استخدام FLD مع أفكار التطبيق الأعمى وغير الأعمى (Blind and Non-Blind). والاعتماد على أكثر من نوع من الصور الملونة ذات الامتداد BMP و PNG. وأثبتت النتائج الكفاءة العالية للتقنيتين في الكشف عن الصورة التي تحوي الرسائل السرية وتفاوتت المقارنات مابين التقنيتين من ناحية نسبة الكشف ومقدار الخطأ وزمن التنفيذ.

الكلمات المفتاحية: الاخفاء، تحليل، آلة المتجه الداعم، مميز فشر الخطي، PNG، BMP.

1- مقدمة

خلال السنوات الماضية، أصبح علم أمن المعلومات محل اهتمام الكثير من الباحثين الذين تحاول جهودهم أن تتوصل إلى حلول وتقنيات جديدة لضمان حماية المعلومات التي ترسل وتستقبل عبر الانترنيت دون حدوث أي اختراق أو كشف، لذلك كان لابد من تطوير أمنية المعلومات وإنشاء تقنيات ووسائل جديدة، من هنا ظهر علم إخفاء المعلومات (Information Hiding) الذي تضمّنَ تقنية الإخفاء (Steganography). إذ تعد تقنية الإخفاء من طرائق الحماية التي تجعل الاتصال غير مرئي عن طريق إخفاء رسائل معينة داخل غطاء معين.

يهتم علم الإخفاء بسرية محتويات الرسالة إضافة إلى تحقيق سرية الاتصال. عندما يشك المتطفل بوجود معلومات مخفية فانه يحاول أن يحل أو يدمر أو يغير الرسالة ، ثم إرسالها إلى المستلم الذي يعلم كيف يحلها. [4]

تهدف تقنية الإخفاء (Steganography) إلى إخفاء البيانات داخل بيانات أخرى، بطريقة لا تؤدي إلى التأثير في هذه الأخيرة، بحيث لا تثير أي شبهة أو شك قد يؤديان إلى كشف المعلومات المخفية.

إن الذي شجع على إحياء وتطوير تقنية الإخفاء هو هذا الانفجار الهائل في تقنية الحاسوب والاتصالات، والشيء المميز فيها أنها تواكب التقنيات الحديثة واستخداماتها في جميع الوسائط الحاسوبية من صور ونصوص وصوت وفيديو ... الخ. إذ أصبحت أمنية المعلومات من الموضوعات الحساسة والمهمة جدا في حياة البشر خاصة بعد انتشار الحكومات الالكترونية في معظم دول العالم. [2]

القصد من هذه المقدمة، أنه مع تطور العلم والأساليب المستخدمة في الإخفاء فهنالك أساليب تتطور بموازاتها في فن تحليل وكسر هذا العلم. فهدف القائم بالإخفاء هو عدم إثارة أي نقطة للشك بوجود بيانات مخفية، أما هدف محلل الإخفاء هو الشك في كل الرسائل المُرسَلة، و فحصها للتأكد من وجود بيانات مخفية مرسلة. تسمى العملية التي تتم فيها محاولة طرف ما اكتشاف وجود المعلومات المخفية، أو قراءتها، أو تغييرها أو حذفها بـ Steganalysis. [1]

2- الدراسات السابقة

في عام 2002 قدم الباحث Farid بحثا لاكتشاف الرسائل المخفية التي تستعمل نماذج Higher Order Statistics حيث يصف طريقة جديدة لاكتشاف الرسائل المخفية عن طريق استخدام Wavelet لبناء هذه النماذج في الصور الطبيعية مع اعتماد تقنية Fisher Linear Discriminant Analysis للتمييز بين الصور الأصلية والمخفية. [9]

أما في عام 2003 فقدم الباحثان Hany Farid و Siwei Lyu بحثا يصف طريقة Multi-scale Wavelet Decomposition يعتمد على بناء أنموذج إحصائي Higher Order لكشف البيانات المخفية في الصور الرمادية. ويستعمل تقنية آلة المتجه الداعم Support Vector Machine (SVM) لكشف هذه الاختلافات الإحصائية في صورة الاختبار. [14]

وفي العام نفسه قدم الباحثان Hany Farid و Siwei Lyu بحثا يصف طريقة استخدام الـ Multi-scale Wavelet Decomposition في بناء نماذج First and Higher Order بالاعتماد على Quadrature Mirror Filters (QMF) لكشف البيانات المخفية في الصور الطبيعية. بالاعتماد على تقنية Fisher Linear Discriminant Analysis للتمييز بين الصور الأصلية والمخفية. [10]

قدم الباحثون Jiang و Wong وMemon وWu عام 2005 تقنية Steganalysis جديدة في صور Halftone بدون معرفة صورة الغطاء الأصلي. نحول أولاً صور Halftone إلى صور Grayscale-Like باستخدام Low-Pass Filtering, ومن ثم تحلل باستخدام Quadrature Mirror Filters للحصول على مجموعة من الميزات الإحصائية. بالاعتماد على تقنية Fisher Linear Discriminant (FLD) Analysis سيتم التمييز بين الصور الأصلية والمخفية. [13]

في عام 2006 قدم الباحثان Hany Farid و Siwei Lyu بحثا إقترح Steganalysis عالمياً يستخدم Higher Order Statistics عن طريق Wavelet Decomposition يعتمد على QMF للحصول على إحصائيات الصورة. مع اعتماد One Class and Multi Class SVM للتصنيف. من مساوئ الخوارزمية عدد الميزات المطلوب لتدريب SVM يكون عالياً, وذلك يتطلّب وقت أكثر لاكتشاف صورة stego. [17]

أما في عام 2007 قدم الباحثون Ge وGao وWang بحثا اقترحوا فيه أنموذجاً عاماً لتطبيق تقنيات تعليم الآلة لكشف المعلومات المخفية باستخدام تقنية إخفاء في البت الأقل أهمية. والاعتماد على مصنف إحصائي لتصنيف الصور. [11]

وفي عام 2009 قدم الباحث Zhang بحثا يقترح فيه طريقة كشف جديدة وهي كشفLSB matching . يتم حساب كل من المدرج الإحصائي لصورة الاختبار وLocal Maximums و Local Minimums ومن ثم حساب المنطقةَ بين Upper Envelope وLower Envelope للمدرج الإحصائي, ويتم تحويّلْ المدرج الإحصائي باستخدام DWT لحساب الفرق بين كل من Local Maximums و Local Minimums و Neighbors للحصول ميزات الصورة, وإدخال هذه الميزات إلى Fisher linear discriminator لتصنيف الصور. [19]

3- الإخفاء وتحليل الإخفاء

إخفاء المعلومات ضمن وسط إلكتروني يتطلّب تعديلات على خصائص ذلك الوسط التي قد تُقدم شكلاً من الاضمحلال أَو الخصائص غير العادية، هذه الخصائص قد تمثل كالتواقيع التي أذاعت وجود الرسالة المُضَمَّنة، وهذا يضعف الغرض من الإخفاء.

تحليل الإخفاء Steganalysis هو علم اكتشاف الرسائل المخفية باستخدام الإخفاء. من الناحية الأخرى, Steganalysis يمكن أن يعمل بوصفه طريقة فعالة لتحكيم أداء أمن تقنيات الإخفاء.[16] الهدف من Steganalysis أن يكتشف وجود الرسالة السرية في الأجسام (objects) ويُميز الأجسام بالرسالة السرية مِن الأجسام بدون أي رسالة سرية. [4]

الهجمات والتحليل على المعلومات المخفية قد تأخذ عدة أشكال: اكتشاف أو انتزاع، أو تعطيل أَو تحطيم المعلومات المخفية. المهاجم قد يُضمّن معلومات مضادة أيضاً على المعلومات المخفية الحالية. إن Steganalysis بشكل عام يُحاول هزيمة هدف الإخفاء باكتشاف المعلومات المخفية واستخلاصها أو تحطيمها .[2]

عملية كشف الإخفاء تتم من قبل جهة أخرى غير المرسل والمستقبل وهو محلل الإخفاء Steganalyst وهو الشخص الذي يطبق تحليل الإخفاء في محاولة لكشف وجود المعلومات المخفية. [20]

4- الهجمات على الإخفاء

جَذب الإخفاء وتحليل الإخفاء الكثير من الانتباه حول العالم في المستقبل القريب, وذلك لاهتمام البعض بضمان اتصالاتهم خلال إخفاء حقيقة خاصة بأنهم يتبادلون المعلومات وآخرون يهتمون باكتشاف وجود مثل هذه الاتصالات. لهذا السبب كانت الحاجة لتصميم وتقييم تقنيات كشف قوية قادرة على تفادي أو تقليل مثل هذه الأعمال. [15]

يُمكن أَن تكون الهجمات على الإخفاء بأنواع مختلفة تعتمد على الأسباب أَو الغرض من الهجوم ضد بيانات_stego. والهجوم الناجح هو الذي يشمل كشف المحتوى المخفي.

يمكن تصنيف أنواع الهجمات على الإخفاء إلى صنفين عامّين هما:

1. الهجوم السلبي Passive Attacks

هذا الهجوم يكشف حضور أو غياب الرسالة السرية المتضمنة في بيانات_stego أو تحديد نوع تضمين الخوارزمية المستخدمة. والمهاجم قادر على اعتراض البيانات فقط. وتتضمن:

- ❖ فحص الوسط إذا كان يحتوي على رسالة سرية أم لا.
- ❖ استخلاص الرسالة السرية إذا كان بالإمكان استخلاصها.
- ❖ تحطيم الرسالة السرية.

2. الهجوم الفعال Active Attacks

تخمين أو انتزاع خصائص الرسالة أو خوارزمية التضمين لتحوير بيانات_stego من اجل تدمير البيانات المضمنة. يحاول محلل الإخفاء تخمين وانتزاع الرسالة السرية بدون تحطيمها, أي انه قادر على معالجة البيانات. [2]

وبشكل عام، تُصنف تقنيات تحليل الإخفاء إلى صنفين وكما يأتي:

1. تحليل الإخفاء المستهدف Targeted Steganalysis

وتعني عملية الكشف عن خوارزمية إخفاء معروفة. إذ يُمكن أن يكشف الرسالةَ السرية أو حتى يُخمّن نسبة التضمين مع معرفة خوارزمية الإخفاء.

2. تحليل الإخفاء الأعمى Blind Steganalysis

يتضمن اكتشاف مدى من خوارزمية الإخفاء. هذا النوع أولاً ينتزع بعض الميزات من المحتويات (الصور)، أَو بشكل محدد أكثر، الصور، ثم يختار أَو يصمم مصنف Classifier ويُدربه باستخدام الميزات التي انتزعت من مجموعات الصور بعد تدريبها, وأخيراً، يُصنف المجموعة الجديدة اعتمادا على الميزات السابقة. [19]

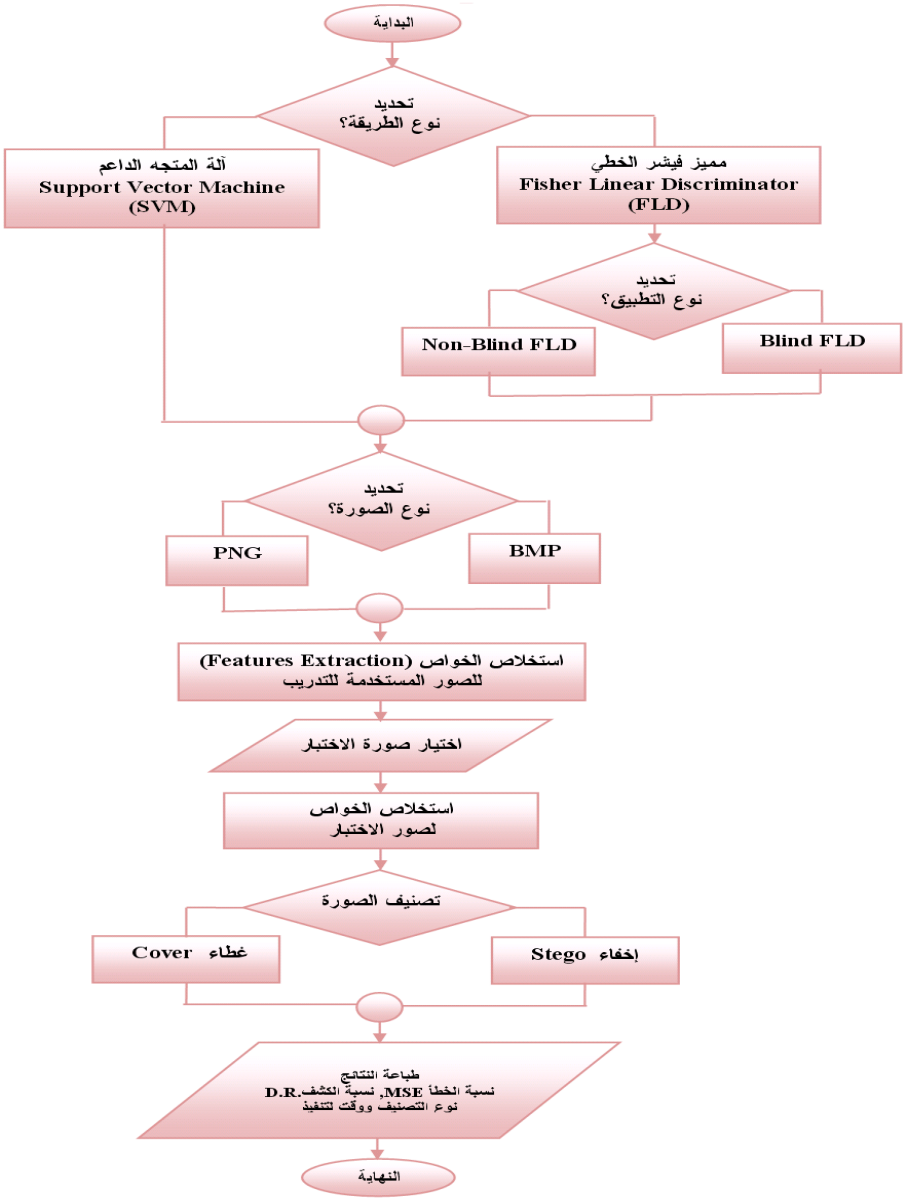
5- فكرة البحث الأساسية

نتيجة الإخفاء المستمر وظهور تقنيات إخفاء جديدة تم اقتراح نظام للتقصي عن تحليل الإخفاء (Steganalysis) باستخدام أنواع مختلفة من الصور، وذلك باعتماده على بناء نماذج إحصائية يتم تكوينها باستخدام عمليات معينة. قبل الدخول إلى النظام الجديد, تم استخدام نوعين من صور الغطاء الملونة للكشف عن المعلومات السرية، بالنسبة للصور الملونة:

- ❖ الامتداد BMP تم الإخفاء باستخدام برنامج (S-Tools V 4.0).
- ❖ الامتداد PNG تم الإخفاء عن طريق برنامج كُتب باستخدام بيئة Matlab (R2008a) البرمجية.

يتم تطبيق النظام المقترح على مرحلتين رئيستين هما كالآتي:

1. عملية استخلاص الخواص Features Extraction.
2. عملية تطبيق الخوارزمية الجديدة، عن طريق تقنية آلة المتجه الداعم (SVM)، وتقنية مميز فيشر الخطي FLD. ويمكن توضيح ذلك من خلال الشكل (1).



الشكل (1). المخطط الانسيابي العام للتنفيذ

6- استخلاص الخواص Features Extraction

تعتمد عملية استخلاص الخواص من الصور الملونة بالامتدادين (PNG , BMP) على بناء أنموذج إحصائي يستخدم فيه المدرج الإحصائي (Histogram) وتحويل فوريير (Discrete Fourier Transform) وذلك للحصول على متجهات الخواص (Features Vectors).

وتوضيح الخوارزمية يكون كالآتي:

الإدخال: صورة الاختبار (BMP أو PNG).

الإخراج: متجهات الخواص (Features Vectors)، وعددها (24) متجه.

أما الخطوات فتشمل:

1. تحليل الصورة إلى مستوياتها اللونية الثلاثة (RGB).

2. بناء مدرج إحصائي لكل مستوى لوني (R , G , B) (1Dimension)

3. حساب العزوم الأولى والثانية (First and Second Moments), بمعنى آخر, المتوسط (Mean) والتباين (Variance) لمعاملات المدرج الإحصائي. وتكون نتيجتهُ (6) إحصائيات.

4. تطبيق تحويل فوريير (Discrete Fourier Transform) على المدرج الإحصائي ولكل مستوى لوني (RGB).

5. حساب العزوم (الأولى، الثانية، الثالثة والرابعة) بمعنى آخر, المتوسط, التباين، الالتواء (Skewness) و(Kurtosis) لمعاملات تحويل فوريير (DFT). وهنا ينتج (12) إحصائية.

6. حساب الطاقة الكلية (Total Energy) لمعاملات تحويل فوريير (DFT) ولكل مستوى لوني. وهذا ينتج (3) إحصائيات.

7. حساب متوسط الفرق بين معاملات المدرج الإحصائي و تحويل فوريير المنفصل لكل مستوى لوني، هذا ينتج أيضا (3) إحصائيات.

8. تجميع الإحصائيات في متجهات الخواص لتصبح (24) إحصائية.

7- آلة المتجه الداعم Support Vector Machine

قُدّمت هذه التقنية في عام (1992) من قبل الباحث (Vapnik) [6]، وهي عبارة عن خوارزمية تعلم عن طريق مشرف أو موجه Supervised تستعمل للتصنيف مستندة إلى نظرية التعلم الإحصائية Statistical Learning Theory.[11]

تقنية آلة المتجه الداعم (SVM) كسبت مؤخراً أهمية في حقل تعلم الآلة وتصنيف الأنماط, إذ أنها أُبتكرت أصلا لحل مسائل تمييز الأنماط Pattern Recognition عن طريق تحديد المستوى الفاصل Hyperplaneللبيانات المراد فصلها. [12] الهدف الأساس من تقنية آلة المتجه الداعم (SVM) هو إيجاد أفضل مستوى فاصل Hyperplane للبيانات المراد فصلها وتصنيفها على صنفين.

تعد SVM تقنية مفيدة لتصنيف البيانات, وذلك لأنها ليست مستخدمة لحل مسائل التصنيف الخطية فقط ولكنها تعدّ أيضا علماً منهجياً قوياً لحل المسائل في التصنيف اللاخطي. [6] يعتمد بناء الأنموذج على عدد من المعلمات مثل المستوى الفاصل (Hyperplane) ومضاريب لاكرانج (Lagrange Multipliers). [5] تعد عملية استخلاص الخواص للحصول على المتجهات خطوة أساسية وأولية للبدء في تقنية آلة المتجه الداعم SVM.

بعدها يتم تدريب التقنية على مجموعة البيانات الناتجة من استخلاص الخواص للحصول على الأوزان المثالية، لاعتمادها بعملية التصنيف وإعطاء التصنيف النهائي. إن ناتج العملية السابقة هو قاعدة بيانات تحوي على مجموعة من متجهات الخواص التي تم الحصول عليها لمجموعة من الصور المستخدمة لتدريب التقنية.

أما في عملية التصنيف فيتم تطبيق الخطوات السابقة من عملية التدريب وباستخدام صورة جديدة (صورة الاختبار) لينتج لنا مجموعة من النماذج الإحصائية والتي ستعتمد أساساً للتصنيف بين الصورة الحاوية على معلومات سرية والصورة غير الحاوية تلك المعلومات.

8- تقنية مميز فيشر الخطي Fisher Linear Discriminator (FLD)

عبارة عن طريقة تصنيف إحصائية، إذ قُدم في عام (1936) من قبل الباحث فيشر (Fisher)، [3] وهو تقنية تصنيف قياسية مستخدمة على نطاق واسع ومثبتة في الكثير من تطبيقات العالم الحقيقي، ومنها (تمييز الأنماط). [18]

تهدف الطريقة إلى إيجاد أفضل إسقاط خطي مثالي (Optimal Linear Projection) بين مجموعتين أو أكثر في عملية التصنيف, والحصول على أعلى تمييز بين المجتمعات وذلك يكون بجعل نسبة التباين بين المجموعات إلى التباين داخل المجموعات كبيرا. [9]

تقوم فكرة هذه الخوارزمية على تدريب مجموعة متجهات الخواص الناتجة من عملية استخلاص الخواص لمجموعة الصور التي تم التدريب عليها, ومن ثم التصنيف لصورة الاختبار.

9- النتائج

أُخذت مجموعة صور ملونة ذات أبعاد 400×600 وبالامتدادين BMP و PNG. تم اعتماد برنامج S-Tools 4.0 بتقنية LSB حيث بلغ حجم ملف الصورة المخفي (87 kilobytes) والذي يمثل اكبر حجم يمكن إخفاءه في صور BMP (400×600) باستخدام برنامج S-Tools 4.0. أما إخفاء نص بحجم 62786 حرفا في صور PNG كان باستخدام LSB.

وكان عدد الصور المستخدمة للتدريب (100 غطاء/ 100 إخفاء).
أما عدد الصور المستخدمة في الاختبار (100 غطاء/ 100 إخفاء).
لكل نوع من أنواع الصور (BMP و PNG).

تم التطبيق لتقنية آلة المتجه الداعم على مرحلتين وهي:

- **مرحلة التدريب Training Phase**

شَملت هذه المرحلة تدريب تقنية آلة المتجه الداعم SVM على مجموعتين من متجهات الخواص: المجموعة الأولى تمثل متجهات خواص صور الغطاء، والمجموعة الثانية هي متجهات خواص صور الإخفاء. وتم اعتماد ثلاثة من هذه الميزات الناتجة من عملية استخلاص الخواص. بعد ذلك تقوم التقنية بعملية التصنيف وإعطاء التصنيف النهائي.

- **مرحلة الاختبار Testing Phase**

مرحلة الاختبار شملت تطبيق تقنية آلة المتجه الداعم SVM على مجموعة جديدة من

متجهات الخواص لمجموعة جديدة من صور الغطاء والإخفاء. في عملية التصنيف تؤخذ النماذج الإحصائية لصورة الاختبار وعلى أساسها يتم تصنيف الصورة أما صورة تحوي معلومات سرية تسمى صورة الإخفاء (Stego-Image) أو صورة لا تحوي معلومات سرية تسمى صورة غطاء (Cover-Image).

وكانت نتائج اختبار صور ملونة ذات امتداد BMP باستخدام تطبيق تقنية آلة المتجه الداعم SVM كما هو مبين في الجدول (1).

نسبة الكشف (DR) = عدد الصور المضمنة بصورة صحيحة / عدد الصور في الاختبار

مقدار الخطأ (MSE) $= \sum_{M,N} [\ stego\text{-}im(m,n) - cover\text{-}im(m,n)]^2 / (M*N)$

M: عدد الصفوف لصور الإدخال.

N: عدد الأعمدة الصور الإدخال.

الجدول (1). نتائج تطبيق تقنية SVM على صور BMP الملونة (400 × 600)

ت	اسم الصورة	نسبة الكشف Detection Rate %	مقدار الخطأ MSE	زمن التنفيذ (ثانية)	نوع التصنيف
1	yellow bird	91	0.03	0.25	cover
2	Coffee	95	0.08	0.17	cover
3	blue car	91	0.04	0.22	cover
4	sunflower	93	0.03	0.24	cover
5	horse race	97	0.09	0.59	cover
6	Flowers	92	0.05	0.27	cover
7	Child	96	0.07	0.21	cover
8	Natural	92	0.05	0.25	cover
9	Bird	96	0.05	0.35	cover
10	Girl	95	0.06	0.20	cover
1	s-yellow bird	92	0.02	0.26	stego
2	s-coffee	93	0.05	0.23	stego
3	s-blue car	94	0.07	0.26	stego
4	s-sunflower	93	0.04	0.22	stego
5	s-horse race	95	0.06	0.22	stego
6	s-flowers	96	0.06	0.23	stego
7	s-child	92	0.03	0.22	stego
8	s-natural	93	0.04	0.22	stego
9	s-bird	96	0.08	0.40	stego
10	s-girl	93	0.07	0.22	stego

ومن خلال ملاحظة الجدول (1) يتبين ما يأتي:

- إن هناك فرقا في المقاييس المستخدمة (نسبة الكشف, مقدار نسبة الخطأ، زمن التنفيذ) بحيث تراوحت نسب الكشف مابين (91-97)% بالنسبة للصور ذات التصنيف (Cover)، في حين كانت النسب هي (92-96)% بالنسبة لصور الإخفاء وهذا يعني الأداء العالي للتقنية المستخدمة في الكشف عن (Cover) و(Stego).
- وفيما يخص مقدار الخطأ نلاحظ أن النسبة تراوحت مابين (0.03 - 0.09) للصور المصنفة (Cover)، في حين كانت النسبة هي (0.02 -ss 0.08) بالنسبة للصور المصنفة (Stego). مما يدل على أن التقنية المستخدمة لا تملك مقدار خطأ ملحوظ.

❖ وإذا كان للزمن أهمية فنلاحظ أن النسبة تراوحت بين (0.17 - 0.59) ثانية لصور Cover و(- 0.40 0.22) ثانية لصور Stego دليلاً على سرعة التقنية في الكشف عن الصور المطلوبة.

علما أن النتائج المثبتة في الجدول (1) كانت كلها صحيحة وبالنسبة للكشف عن الـ (Cover) و(Stego).

ولإثبات كفاءة التقنية SVM تم تطبيقها على أنموذج شائع الاستخدام وهو (PNG). أما النتائج فيمكن مشاهدتها في الجدول (2).

الجدول (2). نتائج تطبيق تقنية SVM على صور PNG الملونة (400 × 600)

ت	اسم الصورة	نسبة الكشف Detection Rate %	مقدار الخطأ MSE	زمن التنفيذ (ثانية)	نوع التصنيف
1	Balloon	73	0.28	0.06	cover
2	Sea	75	0.28	0.10	cover
3	Book	82	0.33	0.12	cover
4	Girls	68	0.21	0.12	cover
5	Tree	79	0.31	0.09	cover
6	Snake and bird	73	0.20	0.09	cover
7	flowers	72	0.24	0.09	cover
8	pen	72	0.26	0.08	cover
9	digital	72	0.23	0.08	cover
10	cat	71	0.18	0.07	cover
1	balloon	74	0.28	0.12	stego
2	sea	74	0.25	0.05	stego
3	book	71	0.22	0.04	stego
4	girls	77	0.31	0.06	stego
5	tree	74	0.28	0.05	stego
6	Snake and bird	71	0.26	0.12	stego
7	flowers	71	0.22	0.05	stego
8	pen	73	0.25	0.05	stego
9	digital	77	0.27	0.07	stego
10	cat	72	0.23	0.09	stego

فيما يخص نتائج تطبيق تقنية FLD بنوعيها الأعمى (Blind FLD) و الغير الأعمى (Non-Blind FLD) تم اعتماد نفس المعلومات السابقة المستخدمة في تطبيق تقنية SVM.

تم استخدام النوع الأول (Blind) من مميز فيشر الخطي على صور BMP، والنوع الثاني (Non-Blind) على BMP وPNG.

وفيما يخص نتائج اختبار الصور الملونة بامتداد BMP باستخدام تقنية مميز فيشر الخطي الأعمى كما هو مبين في الجدول (3).

الجدول (3). نتائج تطبيق تقنية Blind FLD على صور BMP الملونة (400 × 600)

ت	اسم الصورة	زمن التنفيذ (ثانية)	نوع التصنيف	نسبة الكشف Detection Rate %
صور غطاء				
1	yellow bird	0.12	Cover	100 %
2	coffee	0.13	Cover	
3	blue car	0.13	Cover	
4	sunflower	0.14	Cover	
5	horse race	0.13	Cover	
6	flowers	0.14	Cover	
7	child	0.13	Cover	
8	natural	0.13	Cover	
9	bird	0.14	Cover	
10	girl	0.11	Cover	
صور إخفاء				
1	s-yellow bird	0.11	Stego	100 %
2	s-coffee	0.11	Stego	
3	s-blue car	0.10	Stego	
4	s-sunflower	0.13	Stego	
5	s-horse race	0.11	Stego	
6	s-flowers	0.13	Stego	
7	s-child	0.14	Stego	
8	s-natural	0.11	Stego	
9	s-bird	0.14	Stego	
10	s-girl	0.10	Stego	

وتم تطبيق التقنية FLD نوع (Non-Blind) على أنموذج صور (BMP). وكانت نتائج الاختبار باستخدام تقنية مميز فيشر الخطي غير الأعمى وكما هو مبين في الجدول (4) و(5).

الجدول (4). نتائج اختبار صور غطاء ملونة ذات امتداد BMP باستخدام Non-Blind FLD

ت	صورة غطاء	زمن التنفيذ (ثانية)	تصنيف صورة غطاء	نسبة الكشف Detection Rate %
1	yellow bird	0.14	cover	100 %
2	coffee	0.12	cover	
3	blue car	0.12	cover	
4	sunflower	0.13	cover	
5	horse race	0.15	cover	
6	flowers	0.16	cover	
7	Child	0.12	cover	
8	natural	0.14	cover	
9	Bird	0.14	cover	
10	Girl	0.14	cover	

الجدول (5). نتائج اختبار صور إخفاء ملونة ذات امتداد BMP باستخدام Non-Blind FLD

ت	صورة إخفاء	زمن التنفيذ (ثانية)	تصنيف صور إخفاء	نسبة الكشف Detection Rate %
1	s-yellow bird	0.14	stego	100 %
2	s-coffee	0.12	stego	
3	s-blue car	0.12	stego	
4	s-sunflower	0.13	stego	
5	s-horse race	0.15	stego	
6	s-flowers	0.16	stego	
7	s-child	0.12	stego	
8	s-natural	0.14	stego	
9	s-bird	0.14	stego	
10	s-girl	0.14	stego	

وبالنسبة لنتائج اختبار صورة PNG باستخدام هذه التقنية بنوعها غير الأعمى كما هو مبين في الجدول (6) (7).

الجدول (6). نتائج اختبار صور غطاء ملونة ذات امتداد PNG باستخدام Non-Blind FLD

ت	صورة غطاء	زمن التنفيذ (ثانية)	تصنيف صورة غطاء	نسبة الكشف Detection Rate %
1	Color	0.13	stego	80%
2	Sun	0.09	cover	
3	Green	0.10	cover	
4	gray car	0.09	stego	
5	digital	0.09	cover	
6	apple	0.09	cover	
7	yellow car	0.09	cover	
8	player	0.10	cover	
9	motor	0.09	cover	
10	Pen	0.09	cover	

الجدول (7). نتائج اختبار صور إخفاء ملونة ذات امتداد PNG باستخدام Non-Blind FLD

ت	صورة إخفاء	زمن التنفيذ (ثانية)	تصنيف صور إخفاء	نسبة الكشف Detection Rate %
1	s-color	0.13	cover	80 %
2	s-sun	0.09	stego	
3	s-green	0.10	stego	
4	s-gray car	0.09	cover	
5	s-digital	0.09	stego	
6	s-apple	0.09	stego	
7	s-yellow car	0.09	stego	
8	s-player	0.10	stego	
9	s-motor	0.09	stego	
10	s-pen	0.09	stego	

ومن خلال ملاحظة الجداول (6) و (7) يتبين ما يلي:

- ❖ نلاحظ أن الزمن تراوح مابين (0.09 - 0.13) ثانية لصور الغطاء والإخفاء وهذا بسبب اختيار الصورتين في الوقت نفسه للتصنيف وهو دليل على سرعة تنفيذ التقنية في الكشف عن الصور.
- ❖ استطاعت التقنية أن تثبت مدى سرعتها من خلال النتائج التي حصلنا عليها, وهذا يدل على كفاءة التقنية المستخدمة.

ويعرض الشكل (2). نموذجا من الصور التي أجريت عليها اختبارات الكشف بامتداد BMP وPNG.

أ- صورة أصلية بامتداد BMP

ب- صورة أصلية بامتداد PNG

ج- صورة غطاء بامتداد BMP

د- صورة غطاء بامتداد PNG

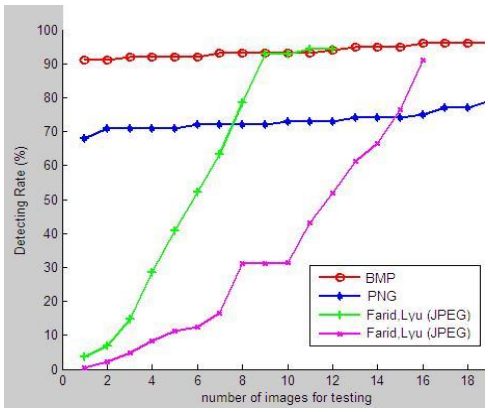
هـ- صورة إخفاء بامتداد BMP باستخدام برنامج إخفاء S-Tools

و- صورة إخفاء بامتداد PNG باستخدام LSB

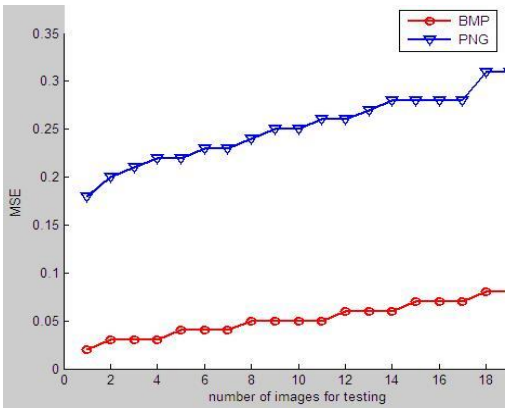
الشكل (2). نموذج من الصور التي أجريت عليها اختبارات كشف صورة

وبمقارنة النتائج الخاصة بالجداول (1) و (2) مع فقرة الدراسات السابقة (2 ,5) من حيث نسبة الكشف لتقنية SVM على صور BMP وPNG لاحظ الشكل (3).

أما ما يخص المقارنة النتائج الخاصة بالجداول (1) و (2) من حيث مقدار نسبة الخطأ لتقنية SVM على صور BMP و PNG فلاحظ الشكل (4).



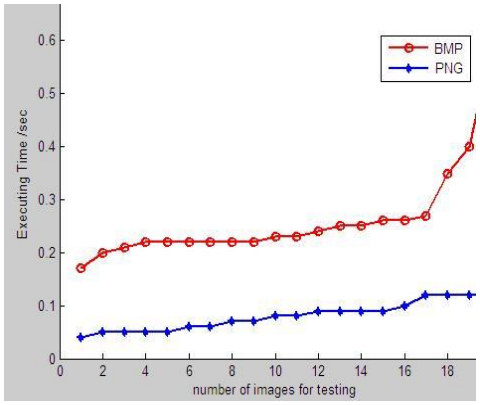
الشكل (3). يوضح نتائج مقارنة الدراسات السابقة (5,2) لنسبة الكشف باستخدام تقنية SVM



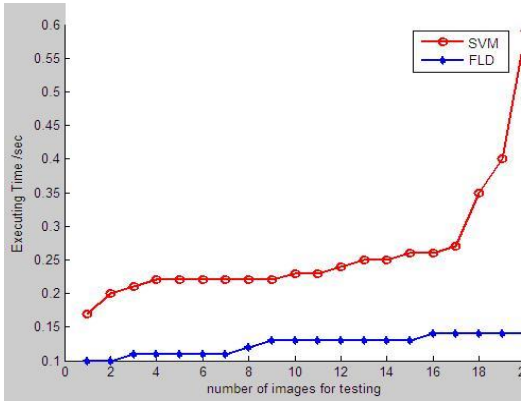
الشكل (4). يوضح مقارنة قيمة مقدار الخطأ لتقنية SVM على صور BMP وPNG

وفيما يخص مقارنة النتائج الخاصة بالجداول (1) و (2) من حيث زمن التنفيذ لتقنية SVM على صور BMP و PNG فلاحظ الشكل (5).

وبالنسبة لمقارنة النتائج الخاصة بالجداول (1) و (3) من حيث زمن التنفيذ لتقنية SVM blind و FLD blind على صور BMP لاحظ الشكل (6).

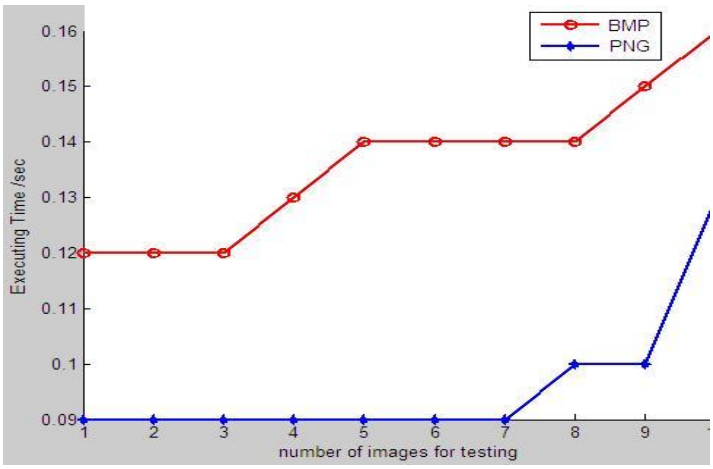


الشكل (5). يوضح مقارنة مقدار وقت التنفيذ لتقنية SVM على صور BMP وPNG



الشكل (6). يوضح مقارنة مقدار وقت التنفيذ لتقنية SVM blind وFLD blind على صور BMP

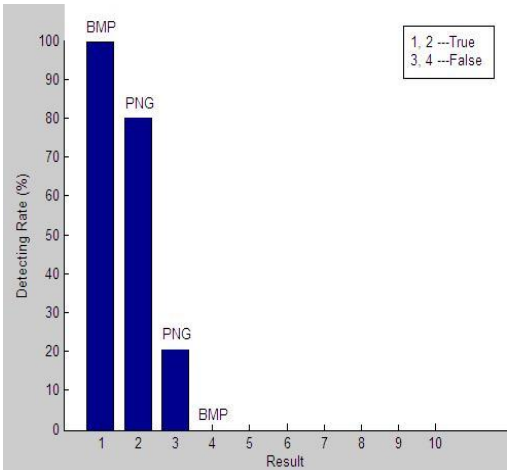
أما فيما يخص مقارنة النتائج الخاصة بالجداول (4) و (5) و (6) و (7) من حيث زمن التنفيذ لتقنية FLD non-blind على صور BMP و PNG لاحظ الشكل (7).



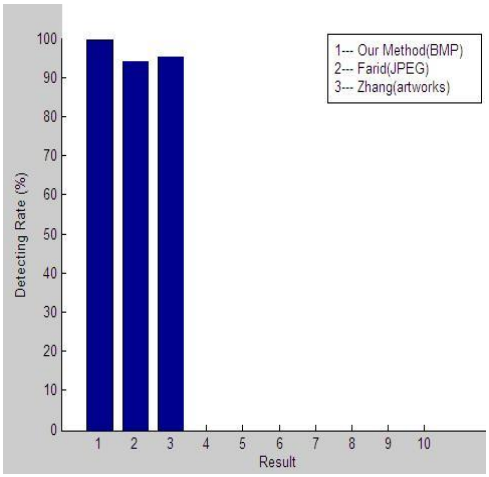
الشكل (7). يوضح مقارنة مقدار وقت التنفيذ لتقنية FLD non-blind على صور BMP وPNG

وفيما يخص مقارنة النتائج من حيث مقدار نسبة الكشف لتقنية FLD non-blind على صور BMP و PNG لاحظ الشكل (8).

وبمقارنة النتائج الخاصة بتقنية FLD مع فقرة الدراسات السابقة (7,1) من حيث نسبة الكشف لاحظ الشكل (9).



الشكل (8). يوضح مقارنة مقدار نسبة الكشف لتقنية FLD non-blind على صور BMP و PNG



الشكل (9). يوضح نتائج مقارنة الدراسات السابقة (7,1) لنسبة الكشف باستخدام تقنية FLD

10- الاستنتاجات Conclusions

من الأمور الأساسية لكل محلل هو:

- الشك بكل ما يرد أمامه من ملفات على الشبكة.
- فكرة عمل محلل الإخفاء هو بطبيعة الحال اعقد وله درجة من الصعوبة أكثر من فكرة عمل الإخفاء.
- لا بد للمحلل أن يكون مطلعاً وملماً بأكثر الطرائق المستخدمة في الإخفاء ولاسيما الحديثة منها.

وبشكل عام يمكن استنتاج:

1. للتوصل إلى عملية كشف صحيحة ودقيقة لا بد من توفر إحدى المعلومات التالية (Cover, Secret ,Message ,Algorithm).
2. هناك علاقة طردية مابين كمية البيانات المخفية وتقنية الكشف.
3. أثبتت النتائج جودة تقنية آلة المتجه الداعم SVM في الكشف عن الإخفاء في الصور بالاعتماد على مقياس نسبة الكشف التي تجاوزت 90%.
4. فكرة التدريب العشوائي لمجموعة التدريب تكون أفضل لان نسبة الكشف تتغير حسب التدريب بدلا من تثبيت مجموعة التدريب, وكلما زادت أحجام مجموعة التدريب كان الكشف أفضل.
5. امتازت تقنية مميز فيشر الخطي (FLD) بأداء عالٍ وسرعة كبيرة في الكشف نظرا لقلة الوقت المستغرق في التدريب.
6. أعطت تقنية FLD مع أفكار (Blind و Non-Blind) نتائج جيدة جدا وكانت الأفضلية لـ (Non-Blind).
7. التحكم في اختيار الخواص (الميزات الإحصائية) المستخدمة من الصورة أعطى قابلية اكبر للكشف عن الإخفاء.
8. وكمقارنة بين التقنيتين فيما بينهما (وبشكل عام) كانت الأفضلية لتقنية SVM على FLD من ناحية (نسبة الكشف ومقدار الخطأ)، أما من ناحية سرعة التقنية فكانت تقنية FLD هي الأسرع.
9. وكمقارنة بين التقنيتين فيما بينهما (وبشكل خاص للصور المستخدمة) كانت الأفضلية لصور BMP على PNG من ناحية (نسبة الكشف ومقدار الخطأ)، أما من ناحية زمن التنفيذ فكانت PNG هي الأسرع.

المصادر

[1] برزنجي، فوزي، 2008، "إخفاء البيانات داخل الصورة"، جامعة السليمانية، العراق. fowzibarazenji@yahoo.com www.boosla.com

[2] الحمامي، علاء حسين والحمامي، محمد علاء، 2008، "إخفاء المعلومات: الكتابة المخفية والعلامة المائية"، إثراء للنشر والتوزيع، الشارقة.

[3] الراوي، عمر فوزي صالح، 2004، "استخدام الدالة التمييزية في السيطرة النوعية مع تطبيق على ولادات الأطفال الخدج"، رسالة ماجستير، قسم الإحصاء، كلية علوم الحاسبات والرياضيات، جامعة الموصل، العراق

[4] الغريري، شهد عبد الرحمن حسّو، 2003، "تصميم نظام حماية هجين و تطبيقه على النصوص"، رسالة ماجستير، قسم علوم الحاسبات، كلية علوم الحاسبات والرياضيات، جامعة الموصل، العراق.

[5] قاسم، عمر صابر، 2009، "تطبيق التقنيات الذكائية في المعلوماتية الحياتية "، أطروحة دكتوراه، قسم الرياضيات، كلية علوم الحاسبات والرياضيات، جامعة الموصل، العراق.

[6] Ahmad A., Khalid M. and Yusof R., "Machine Learning Using Support Vector Machines", Universiti Teknologi Malaysia. {marzuki,rubiyah}@utmkl.utm.my

[7] Cheng J.**,** Kot A., and et.al., ICASSP 2005, "STEGANALYSIS OF BINARY TEXT IMAGES", Nanyang Technological University, Singapore.

[8] Du Q., 2007 IEEE,"Modified Fisher's Linear Discriminant Analysis for Hyperspectral Imagery", *Senior Member,* IEEE GEOSCIENCE AND REMOTE SENSING LETTERS, VOL. 4, NO. 4, OCTOBER 2007.

[9] *Farid* H., 2002, " DETECTING HIDDEN MESSAGES USING HIGHER-ORDER STATISTICAL MODELS", Department of Computer Science, Dartmouth College, Hanover. farid@cs.dartmouth.edu

[10] Farid H. and Lyu S., 2003, " Higher-order Wavelet Statistics and their Application to Digital Forensics", IEEE Workshop on Statistical Analysis in Computer Vision (in conjunction with CVPR).

[11] Ge S., Gao Y. and Wang R., 2007 ACM, " Least Significant Bit Steganography Detection with Machine Learning Techniques ",National Laboratory for Novel Software Technology Nanjing University 210093 Nanjing, Jiangsu, China, geshengeorge@gmail.com gaoy@nju.edu.cn R.wang@massey.ac.nz

[12] Ivanciue, O., (2007), "Applications of support Vector Machines in Cjemistry", Wiley-VCH, John Wiley & Sons, Volume 23 pp. 291400

[13] Jiang M., Wong E., Memon N. and Wu X., ICASSP 2005, "Steganalysis of Halftone Images", *IEEE Int'l Conf on Acoustics, Speech, and Signal Processing,* Philadelphia, PA.2005.

[14] Lyu S. and Farid H., Springer-Verlag Berlin Heidelberg 2003, " Detecting Hidden Messages Using Higher-Order Statistics and Support Vector Machines", Dartmouth College, Hanover, USA

[15] Rocha A., and Goldenstein S., (2008), " Steganography and Steganalysis in Digital Multimedia: Hype or Hallelujah?", RITA, Vo. 15, No. 1, pp.83-110.

[16] Shi Y., Xuan G., and et.al., "Steganalysis Based on Moments of Characteristic Functions Using Wavelet Decomposition, Prediction-Error Image, and Neural Network", New Jersey Institute of Technology, Newark, NJ, USA, Tongji University, Shanghai, China.

[17] Siwei Lyu and Hany Farid, 2006, "Steganalysis using Higher-Order Image Statistics", *IEEE Transaction on Information Forensics and Security*, vol. 1, pp. 111-119, 2006.

[18] Yang J., Jina Z., Yang Y. and Frang A., 2003, "Essence of kernel Fisher discriminant: KPCA plus LDA", Department of Computer Science, Nanjing University of Science and Technology, Nanjing.

[19] Zhang J., Hu Y. and Yuan Z., ACADEMY PUBLISHER 2009, "Detection of LSB Matching Steganography using the Envelope of Histogram", Guangdong University of Business Studies, Guangzhou P.R. China, JOURNAL OF COMPUTERS, VOL. 4, NO. 7, JULY 2009.
Zhangjundan123@yahoo.com.cn okhyp@yahoo.com.cn

[20] http://coeia.edu.sa/ar/asuurance-awarness/articles/57-cryptography-and-steganography-and-pki/557-steganography.html