International Journal of Multimedia and Ubiquitous Engineering
Vol.11, No.1 (2016), pp.385-396
http://dx.doi.org/10.14257/ijmue.2016.11.1.37

Optimized JPEG Steganalysis

J. Anita Christaline[1], R. Ramesh[2] and D. Vaishali[3]

Department of ECE[1], SRM University, Chennai, India, Saveetha Engineering College[2] Chennai, India, Department of ECE[3], SRM University, Chennai, India
[1]anita.j@vdp.srmuniv.ac.in, [2] ramesh@saveetha.ac.in,
[3]vaishali.b@vdp.srmuniv.ac.in

Abstract

Feature based image Steganalysis demands the best feature model for accurate steganalysis. The extracted feature model includes the components in DCT features of JPEG image. Existing research in this field show extraction of different types of image features that show slightly improved classification accuracies. Though few recent methods of image steganalysis involve extracting all possible features of the image, they suffer dimensionality problem. The dataset used in our research include raw images from the BOSS database. The original dimension of the feature set extracted has 8726 features from 2000 images. While a larger feature set is expected to have all important information about the steganographic changes, it affects the classifier accuracy due to redundancy. To overcome the curse of dimensionality, we intend to introduce an unsupervised optimization technique before classification. The individual classifiers implemented are SVM and MLP and the fusion techniques implemented to combine these classifiers are Bayes, Dempster Schafer and Decision Template schemes. The performances of classifiers are analyzed for optimization based on Euclidean distance measure and Mahalanobis distance measure. Comparing individual classifiers, it has been found that SVM classifier outperforms MLP classifier for both Euclidean distance measure and Mahalanobis distance measure. Among the fusion schemes, the accuracy of Bayes fusion scheme proves to be best compared to Decision template and Dempster Schafer schemes. Also, the best possible classification accuracy has been obtained for Euclidean distance based optimization followed by Bayes fusion classifier scheme. The classification accuracies obtained in our research are better compared to existing methods.

***Keywords**: JPEG Steganalysis, DCT features, Optimization, Fusion Classifiers, SVM, Bayes fusion scheme*

1. Introduction

Labeled as Steganography, the science of covert communication is the art of hiding secret information into some digital cover media like audio, image or video. As human perceptibility is less for changes in image data, image steganography has gained prominence in the recent past. Apart from commercially available steganographic tools like Ez Stego [1], JP Hide and Seek [1], J-Steg [1], Outgess [1] and F5 [2], improved versions are being developed for specific applications. As steganography is being used illegally to outwit the law enforcement authorities [3], the need to crack the secret communication (steganalysis) has gained momentum. While steganalysis can be universal (blind) steganalysis or embedding specific steganalysis, universal method of steganalysis finds wide application. According to [4] universal steganalysis is a two class pattern classification problem, where the clean image and the corrupted image have to be separated. This classification is based on the features that undergo changes during steganography (embedding changes). The two categories of universal steganalysis are statistical blind steganalysis and computational intelligence (CI) based steganalysis [4].

ISSN: 1975-0080 IJMUE

While statistical blind steganalysis uses wavelet features, Markov features and DCT features, CI based steganalysis uses Bayesian theory, genetic algorithms and neural network based classifiers. In either case, the classification accuracy depends on the quality of the chosen feature set. This demands better understanding about the properties and design of feature space. Obtaining a better feature model is a never ending battle between steganography and steganalysis. As many of the features in a chosen model are redundant, best features have to be chosen by proper optimization techniques. Large dimensionality of the feature set may lead to overtraining of the classifier and hence poor classification accuracy [5]. Together with more computational time, the large dimensionality problem has demanded the feature reduction techniques as the future in steganalysis.

To take up this challenge, we intend to analyze the steganographic images in JPEG domain. Hence the concepts related to JPEG image steganalysis, methods of feature set optimization and machine learning classifiers for steganalysis have been implemented and tested in this research.

2. JPEG Steganalysis

JPEG has become the most common format for storing and transmitting images over the internet. This popularity of JPEG has made it an easy cover for image steganography. Most steganographic algorithms hide the information in the DCT coefficients of JPEG images [6-9]. With emerging JPEG steganographic algorithms, JPEG steganalysis is an emerging area of research. As most of the JPEG steganographic techniques modify the DCT coefficients, detecting the changes in the DCT coefficients as image model would help in their steganalysis. This research work intends to identify all the possible changes in DCT coefficients of a JPEG image. As known the DCT modes of a JPEG image occurs in 64 parallel channels and show two types of dependencies. These include the intra block (frequency) and inter block (spatial) dependencies. While the frequency dependencies show the relationship among the coefficients within a block (8x8), the spatial dependencies show the relationship among the coefficients of different blocks. Shi *et al.* [10] have modeled the dependencies of the JPEG DCT coefficients as Markov transition matrices and have derived Markov features from them based on a threshold. Owing to the need for reduced dimensionality and efficient training of the classifiers, most of the past researchers [11-14] have used limited DCT coefficients in the form of conditional probability distributions, co-occurrence matrices, joint probability distributions and their calibrated versions [15-16]. This limited usage treats all the DCT modes equally even though they are statistically different, which leads to inefficient steganalysis. Based on this intuitive, our research extracts all possible changes among the inter block and intra block JPEG DCT coefficients as features. While other authors have used limited features to create the image model, we intend to use the large dimensional feature set of the DCT coefficients by calculating the differences in DCT coefficients in all the directions followed by appropriate feature set optimization technique.

3. Investigative Setup

3.1 Image Database

The raw images used in this research are taken from the BOSS BOSS (Break Our Steganographic Systems) database [17]. This database is available for researchers in steganalysis. The database has 9074 full resolution, uncompressed images in .pgm format. These original images taken from Canon EOS, Leica, Panasonic, NIKON and Pentax cameras are resized and cropped to get 512x512 images. 1000 images are chosen for our research and converted into JPEG format with quality factor 75. These JPG images are the cover images in which the secret data is embedded to create another 1000 stego images.

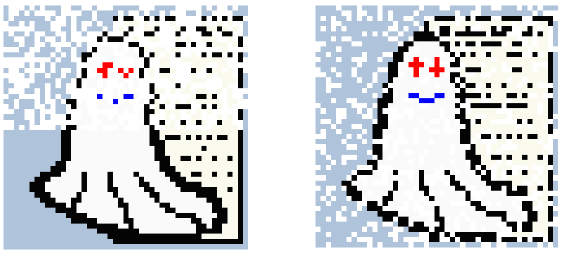
Thus a total image data of 2000 images are used for feature extraction, optimisation and classification. A few sample images used in this research are shown in Appendix.

3.2 Creation of Stego Images

The stego images created in this research work is based on the modified F5 algorithm. The original F5 algorithm used by Westfield [18] was based on permutation straddling and matrix encoding of the JPEG coefficients for increased embedding capacity. F5 algorithm considers the permutation of the image pixels to ensure uniform distribution of embedded information in an image at the cost of time complexity of Order $O(n)$. The F5 sequence is

$$F5 \leftarrow HFM\{EMB\{PERM_{key}\{JPEG[I_{n\times m}]\}\}\} \qquad \text{eqn. (1).}$$

Where $I_{n\times m}$ is the cover image of dimension n×m, PERM is permutation function according to key of a chosen password. The permuted sequence undergoes embedding and is finally delivered to Huffman coder. The uniform distribution of embedding changes due to permutation straddling is shown in Figure 1.



Continuous Embedding   Permuted Embedding

Figure1. Distribution of Embedding Changes Due to Continuous and Permuted Embedding

Matrix encoding for embedding two secret bits a_1 and a_2 in three image bits b_1, b_2, b_3 with at most one bit change may lead to four possibilities :

- $a_1 = b_1 \oplus b_3$, $a_2 = b_2 \oplus b_3$ implies no change ,
- $a_1 \neq b_1 \oplus b_3$, $a_2 = b_2 \oplus b_3$ implies change in a_1,
- $a_1 = b_1 \oplus b_3$, $a_2 \neq b_2 \oplus b_3$ implies change in b_2 ,
- $a_1 \neq b_1 \oplus b_3$, $a_2 \neq b_2 \oplus b_3$ implies change in b_3.

This assures only one bit change. Generalizing, for a code word b with n modifiable bits with k secret bits in x, matrix encoding gives b' for every b according to $x=f(b')$ such that the hamming distance $H_d(b, b') < H_{dmax}$. Thus F5 implements matrix encoding in terms of the triple (n, k, d_{max}) : to embed k bits, a code word with n places would be changed in less than d_{max} places. The embedding change density for code words with length $n= 2^k-1$,

$$S(k) = \frac{1}{n+1} = \frac{1}{2^k} \qquad \text{eqn. (2).}$$

Combined with embedding rate, $R(k) = \frac{k}{n} = \frac{k}{2^k-1}$ eqn.(3).

Gives an embedding efficiency (average number of bits per change),

$$\eta = \frac{R(k)}{S(k)} = \frac{k\,2^k}{2^k-1} \qquad \text{eqn. (4).}$$

When $d_{max} =1$, the embedding efficiency is always larger than k. F5 suffers from the problem of shrinkage where a coefficient may become zero after embedding. This demands embedding the data again in the next coefficient as the decoder will skip the zero coefficients. Thus the efficiency of F5 is reduced as 50% of the coefficients (+1 and -1) are skipped. The modified F5 algorithm used in this research applies the F5 algorithm to the DCT coefficients after syndrome coding where the sender encodes b bits of message $m \in$

$\{0,1\}$ b in n number of AC coefficients whose LSBs are $x\epsilon(0,1)n$. Only k out of n LSBs such that $|I| = k$ are non zero. For few embedding changes in sender side $x_i \epsilon I$, the modified LSBs $y\epsilon(0,1)n$ satisfy the criteria $m = yR$. R is a binary matrix known to both sender and receiver. Hence the embedding process involves finding the solution for $m = yR$ such that $x_i = y_i$ for $i\epsilon I$ and the hamming distance $x - y$ is minimum. The choice of a random matrix R provides a continuous family of codes with better adjustment of parameters for each specific payload. Thus the problem of shrinkage gets eliminated in this modified algorithm.

3.2 Image Feature Extraction

The aim of this research being universal steganalysis, changes due to embedding in all aspects needs to be identified. The feature extraction in our research builds a large number of individual feature vectors from the absolute values of the DCT coefficients. The 64 channels of DCT coefficients have dependencies in terms of frequency and space. The frequency dependencies correlates the DCT coefficients within the same 8×8 block and the spatial dependencies correlate the coefficients between two different 8×8 blocks. Most of the recent research work concentrate on the statistical relation between neighboring DCT coefficients and hence treat all DCT modes equally to get a reduced features set suitable for classification. The feature set extracted in our research consists of diverse relations between different DCT coefficients, the relation among coefficients in entire DCT plane and their differences in different directions. Considering an image of dimension M×N, the DCT coefficient $C_{P,Q}^{(I,J)}$ is the PQ^{th} coefficient of the IJ^{th} block where $(P,Q)\ \epsilon\ [0,1,\ldots 7]^2$, $I\epsilon\ [1,\ldots..M/8]$ and $J\epsilon[1,2,\ldots.N/8]$. The major groups of vectors chosen include the absolute values of the single DCT coefficients Z, the difference in coefficients among intra block coefficients Z_a, the difference in coefficients among inter block coefficients Z_e. Among the last two groups, differences in both horizontal (Z_{eh}, Z_{ah}) and vertical directions (Z_{ev}, Z_{av}) are considered.

$Z_{i,j} = C_{i,j}\ ;\ i\ \epsilon\ [1,2,\ldots M]$ and $j\ \epsilon[1,2,\ldots..N]$ eqn (5).
$Z_{eh} = C_{i,j} - C_{i,j+1}\ ;\ i\epsilon[1,2,\ldots M]$ and $j\ \epsilon[1,2,\ldots..N-1]$ eqn (6).
$Z_{ah} = C_{i,j} - C_{i,j+8}\ ;\ i\epsilon[1,2,\ldots M]$ and $j\ \epsilon[1,2,\ldots..N-8]$ eqn (7).
$Z_{ev} = C_{i,j} - C_{i+1,j}\ ;\ i\epsilon[1,2,\ldots M-1]$ and $j\ \epsilon[1,2,\ldots..N]$ eqn (8).
$Z_{av} = C_{i,j} - C_{i+8,j}\ ;\ i\epsilon[1,2,\ldots M-8]$ and $j\ \epsilon[1,2,\ldots..N]$ eqn (9).

The exact model development involves the computation of the co-occurrence matrices from the coefficients of the above matrices. Mathematically, these are represented as

$$CC_{kl}(p,q,\Delta p,\Delta q) = \frac{1}{R_0} + \sum_{i,j} \left|\left(C_{P,Q}^{(I,J)}\right)\right| \qquad \text{eqn (10).}$$

Where R_o is the normalization constant that maintains $\sum CC_{kl} = 1$ and $C_{P,Q}^{(I,J)}$ may be truncated by a sign function. The values of $(p, q, p+\Delta p, q+\Delta q)$ are not restricted to $\{0, 1, \ldots 7\}$ as the calculation of co-occurrences considers $C_{P+8,Q}^{(I,J)} = C_{P,Q}^{(I+1,J)}$. Due to the sign symmetry of the 8×8 DCT blocks, the matrices can be considered as

$$\overline{CC}_{kl}(p,q,\Delta p,\Delta q)= \frac{1}{2}\left(CC_{kl} + CC_{-k,-l}\right) \qquad \text{eqn. (11).}$$

Comparing the coefficients of the absolute components and that of the difference components, the bins of absolute components are zero due to non negative coefficients and hence its exact dimensionality is $(C+1)^2$ rather than $(2c+1)^2$. The redundant values of $\overline{CC}_{kl}$ have been reduced to get the dimensionality of $\frac{1}{2}(2C+1)^2 + \frac{1}{2}$ with compact models. The compact variables include 10 co-occurrences of the DCT absolute values (represented as Ab), 10 co-occurrences of the difference in DCT absolute values in horizontal direction (represented as Ch), 10 co-occurrences of the difference in DCT absolute values in vertical direction (represented as Cv) , 4 co-occurrences of the difference in DCT absolute values among inter block coefficients for spatial and frequency components in horizontal and vertical directions (represented as Ci). The various components are shown in Figure2.

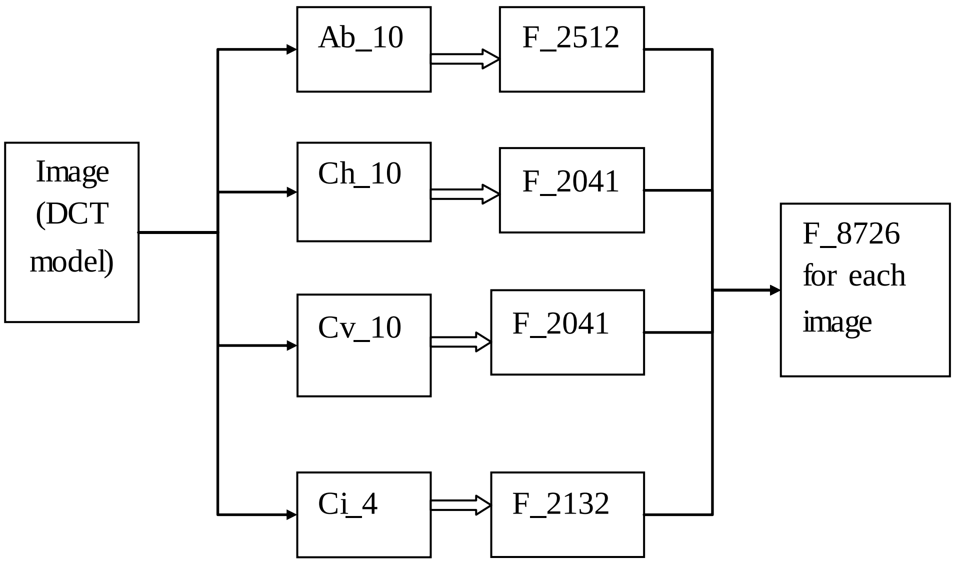


Figure 2. The Variables and the Features Extracted

This creates 34 groups each with features of variable lengths. For example, the co-occurrences of the DCT absolute values contribute 2512 features, co-occurrences of the difference in DCT absolute values in vertical direction contribute 2041 features. Each of these feature sets are due to individual feature components who are variable in length. The individual features of all these 34 groups are combined to give 8726 features for each image. Considering 1000 clean images and 1000 stego images, the extracted feature set is of size '2000 X 8726'.

3.3 Feature Set optimization

Universal image steganalysis being a two class classification problem, the features set extracted needs to be optimized to suit classification with less computational complexity. Increased computational time and lesser accuracy seems to be the curse of larger dimensional feature set [5]. This demands feature set optimization to choose the best features for Steganalysers. With this intuition, we introduce an unsupervised optimization algorithm to reduce the feature set size by a neural network. This NN based algorithm maps X dimensional vectors from S^X vector space to an optimal set $Y = \{yi\ ;\ i = 1, 2, \ldots.M\}$. Each y_i is one of the co- occurrence reduced from a variable dimension $\overline{CC}_{kl}$ to $\overline{CC}_y$. For each y_i there exists a nearest neighborhood, a partition of the entire vector space S^X, defined by

$$N_i = \{p \in S^X : ||p - y_i|| \leq ||p - y_j|| \ ;\ for\ all\ j \neq i\} \qquad \text{eqn. (12)}$$

The unsupervised network considers each partition as a cluster and assigns a centre. Each centre is implemented as a neuron of a SOM (Self Organizing Map). The extracted feature vectors dimension matches that of the input space of the SOM and their values act as the parameters of the NN. Training of the NN leads to the adaptation of the feature parameters to the cluster centre according to a minimum distance criterion. The chosen measure for the pair wise distances are Euclidean distance measure and Mahalanobis distance. The mean distance is maintained small for better data representation. The optimized features are exclusive features with no overlap. The unsupervised training tries to localize the feature by minimizing the mean distances between the chosen neurons and the feature value. The feature set with 8726 features has been reduced with a neural network comprising of 4 neurons and 8 neurons to give 392 and 784 features respectively. Figure3 shows the neighborhood and cluster formation before and after optimization.

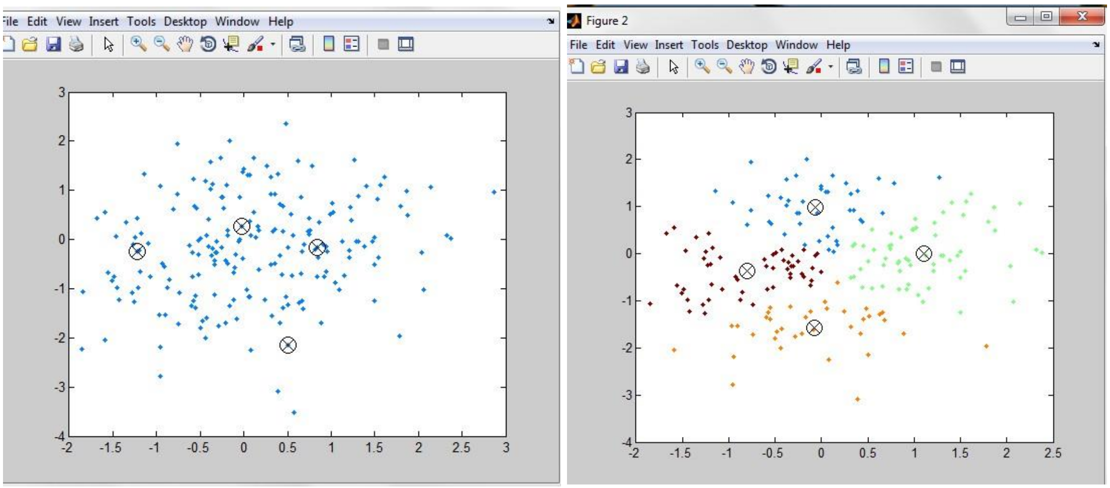


Figure 3. Neighborhood Formation Before and After Optimization.

The optimization of the feature set to 2000 x 392 features is implemented with 4 neurons and that with 2000 X 784 features is implemented with 8 neurons. This gives an optimization efficiency of 95.05% (4 neurons) and 91.01% (8 neurons).

3.4 Classification Scheme

After identifying the most appropriate feature set, the next step is to classify the images as clean and stego images. Classification is the most crucial part in steganalysis as it decides whether an image is true (clean) image or stego image. We have implemented two main classifiers SVM and MLP and fused versions of them. SVM (Support Vector Machine) is a powerful classifier and works under the principle of non-linear mapping, where the input vectors are mapped to high dimensional feature space [20]. The SVM works on linear data according to the function,

$$f(x) = \sum a_i y_i (x_i^T x) + b \qquad \text{eqn (13).}$$

The normalization of the weight vector and the margins in SVM can be chosen by designer. The normalization is chosen such that $w^T x_+ + b = +1$ and $w^T x_- + b = -1$ so that the margin for classification is

$$\frac{w}{||w||}(x^+ - x^-) = \frac{1}{||w||}(w^T x_+ - x\_) = \frac{2}{||w||} \qquad \text{eqn (14).}$$

MLP (Multi-Layer Perceptron) is another popular pattern recognition or classification network. MLP is a powerful classifier working under the principle of back propagation training [20]. MLP also has many free parameters so that the input associates with high output response. Owing to their high performance, we intend to choose SVM and MLP classifiers. We opted for fusing SVM and MLP classifiers with three different fusion schemes, namely Bayes method, Decision Template scheme and Dempster Schafer method. The accuracy of these fusion methods have been studied to identify the best method suitable for image steganalysis.

4. Investigation Outcome

The investigation in our research consists of two parts namely classification with Euclidean distance based optimization and classification with Mahalabonis distance based optimization. The research outcomes are enumerated in Table.1 and Table.2.

Algorithm

- Set the number of neurons to be chosen for optimization.
- For a specific payload (say 0.1) run the steganographic algorithm to create stego images.
- Extract features from both stego images and cover images to create a 2000 X 8726 feature vector.
- Run optimization algorithm to reduce feature set to 2000 X 392 (for 4 neurons).
- Repeat for different payloads (0.2, 0.3. 0.5).
- Repeat the above steps for 8 neurons to get reduced feature set (2000 X 784).

Table. 1. Accuracy of Classifiers for Optimization Based on Euclidean Distance

No. of neurons	Feature size	Payload (bpnzac)	SVM	MLP	Bayes	Decision Tree	Dempster Schafer
4	2000X392	0.1	0.453	0.3656	0.5470	0.4477	0.4542
		0.2	0.4126	0.3779	0.5874	0.5208	0.5208
		0.3	0.4126	0.3452	0.5874	0.4269	0.4269
		0.5	0.4126	0.3803	0.5874	0.4408	0.4408
8	2000X784	0.1	0.4404	0.3734	0.5596	0.4387	0.4191
		0.2	0.4469	0.4056	0.5531	0.4865	0.4865
		0.3	0.3938	0.4081	0.5062	0.4800	0.4800
		0.5	0.4404	0.3746	0.5596	0.4007	0.4007

The above table presents the experimental values obtained for Euclidean distance optimization with 4 neurons (392 features) and 8 neurons (784 features). The classifier accuracies are enumerated for individual classifiers (SVM and MLP) along with fusion classifiers (Bayes, Decision template and Dempster Schafer) for varying payloads.

Table. 2. Accuracy of Classifiers for Optimization Based on Mahalabonis Distance

No. of neurons	Feature size	Payload (bpnzac)	SVM	MLP	Bayes	Decision Tree	Dempster Schafer
4	2000X392	0.1	0.4126	0.4146	0.5874	0.4661	0.4616
		0.2	0.3934	0.3938	0.6066	0.5792	0.5400
		0.3	0.4126	0.3395	0.5874	0.4334	0.4334
		0.5	0.4126	0.3930	0.5874	0.4583	0.4583
8	2000X784	0.1	0.5241	0.4146	0.3607	0.4146	0.4146
		0.2	0.5000	0.4469	0.3480	0.4469	0.4469

		0.3	0.5507	0.4412	0.3546	0.4461	0.4461
		0.5	0.5343	0.4220	0.3611	0.4154	0.4154

The performance measure used to compare the classifiers is 'Classification Accuracy'. The percentage of correct prediction is called as classification accuracy and is calculated as, Accuracy = (TP + TN) / (TP + TN + FP + FN), where TP is number of True Positive, TN is number of True Negative, FP is number of false Positive and FN is number of False Negative.

5. Discussion

The performance of individual classifiers is good but fusion classifiers give better results. Comparison of the performance of SVM and MLP show that SVM is better as shown in Figure4. Among the three fusion schemes -Bayes, Dempster Schafer, Decision Template, Bayes fusion scheme gives the highest classification accuracy compared to others as shown in Figure 5.

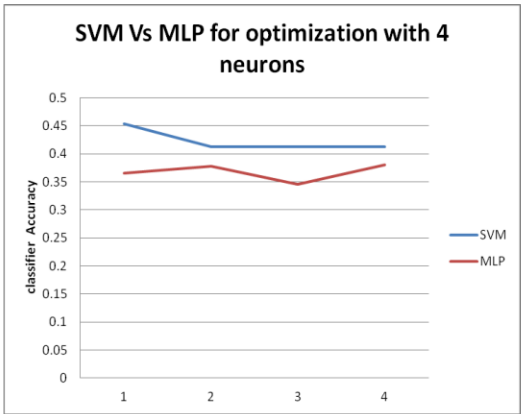


Figure 4. SVM Vs MLP

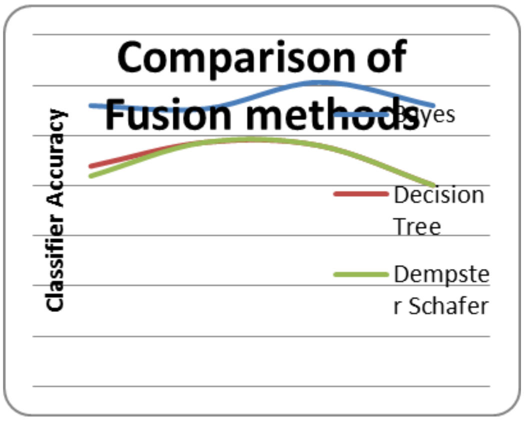


Figure 5. Comparison of Fusion Classifiers

Comparing individual classifier (SVM) with fusion classifier (Bayes), it can be seen that Bayes fusion classifier outperforms SVM classifier as in Figure6.

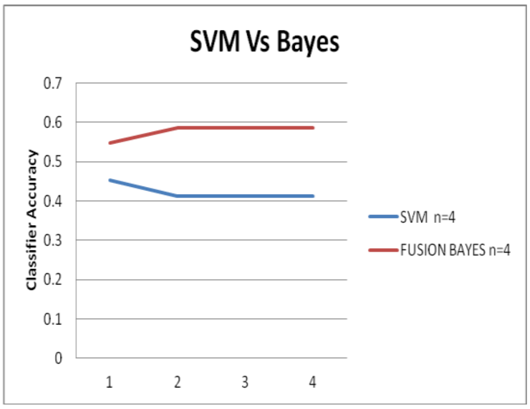


Figure 6. SVM Vs Bayes

Comparing the performance based on the number of neurons, the performance of individual classifiers (SVM and MLP) is more for 8 neurons (784 features) than 4 neurons (392 features) as shown in Figure 7. The performance of fusion classifiers - Bayes, Dempster Schafer, Decision Template is better for 4 neurons (392 features) than 8 neurons (784 features) as in Figure 8. This shows that fusion schemes work well with fewer features than large number of features.

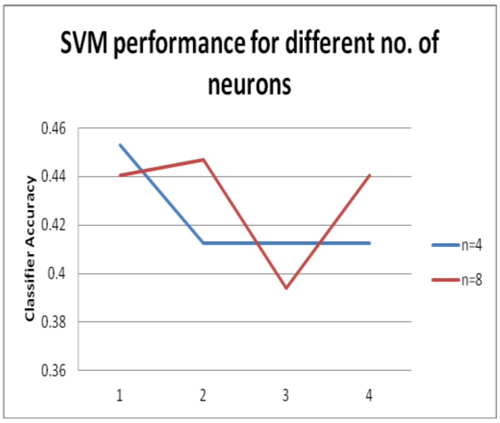


Figure 7. SVM for Different No. of Neurons

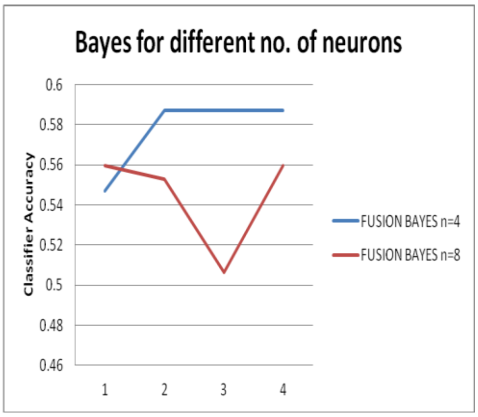


Figure 8. Bayes for Different No. of Neurons

Thus it has been identified that individual classifiers require more features for good classification accuracy, but fusion classifiers require fewer features for better accuracy. The importance of feature set optimization has been proved in this research and our optimization method has provided better classification accuracies compared to existing methods of image steganalysis.

6. Conclusion

As feature based steganalysis is the recent trend, extraction of important features has become a prime requisite. While features of cross domain parameters could give better steganalysis, the dimensionality of these features becomes a curse with respect to computational time and classification accuracy. Our research work has experimentally analyzed an unsupervised optimization technique for the chosen DCT features in JPEG domain. It has been found that optimization leads to better performance and classification accuracy. And the use of fusion classifiers has led to more accuracy in classification compared to individual classifiers. Our research has also analyzed the performance of classifiers for two different optimization metrics, namely Euclidean distance and Mahalanobis distance. It has been identified that SVM is good among the individual classifiers and Bayes fusion classifier is best among fusion classifiers. Also fusion classifiers give better classification accuracy for fewer features. This proves the importance of feature set optimization in image steganalysis and our optimization method has provided better classification accuracies compared to existing methods of image steganalysis.

References

[1] "Steganography software tools", http://members.tripod.com/steganography/stego/software.html Accessed, January 10, **(2015)**.

[2] A. Westfeld, "High capacity despite better steganalysis (F5 – a steganographic algorithm)", In I. S. Moskowitz, editor, Information Hiding, 4th International Workshop, volume 2137 of Lecture Notes in Computer Science, **(2001)**, pp. 289–302.

[3] Kelly J., "Terror Groups Hide Behind Web Encryption", USA Today, 2 May 2001. http://www.usatoday.com /life/cyber/tech/2001-02-05-binladen.htm. Accessed, January 30, **(2015)**.

[4] R. Din and A. Samsudin, "Digital Steganalysis: Computational Intelligence Approach", International Journal of Computers, vol. 3, no. 1, **(2009)**.

[5] J. Fridrich and J. Kodovsky, "Rich Models for Steganalysis of Digital Images", IEEE Transactions on Information Forensics and Security, **(2011)**.

[6] P. Sallee, "Model-based steganography", In T. Kalker, I. J. Cox, and Y. Man Ro, editors, Digital Watermarking, 2nd International Workshop, volume 2939 of Lecture Notes in Computer Science, **(2003)**, pp. 154–167.

[7] Y. Kim, Z. Duric and D. Richards, "Modified matrix encoding technique for minimal distortion steganography", In J. L. Camenisch, C. S. Collberg, N. F. Johnson, and P. Sallee, editors, Information Hiding, 8th International Workshop, volume 4437 of Lecture Notes in Computer Science, **(2006)**, pp. 314–327.

[8] A. Sarkar, K. Solanki and B. S. Manjunath, "Further study on YASS: Steganography based on randomized embedding to resist blind steganalysis", In E. J. Delp, P. W. Wong, J. Dittmann, and N. D. Memon, editors, Proceedings SPIE, Electronic Imaging, Security, Forensics, Steganography, and Watermarking of Multimedia Contents X, vol. 6819, **(2008)**, pp. 16–31.

[9] V. Sachnev, H. J. Kim and R. Zhang, "Less detectable JPEG steganography method based on heuristic optimization and BCH syndrome coding", In J. Dittmann, S. Craver, and J. Fridrich, editors, Proceedings of the 11th ACM Multimedia & Security Workshop, **(2009)**, pp. 131–140.

[10] Y. Shi, C. Chen and W. Chen, "A Markov process based approach to effective attacking JPEG steganography", Information Hiding, in: Proceedings of the 8th International Workshop, Springer, Berlin, **(2006)**, pp. 249–264.

[11] C. Chen and Y. Q. Shi, "JPEG image steganalysis utilizing both intra-block and inter-block correlations", In Circuits and Systems, ISCAS 2008. IEEE International Symposium on Circuits and Systems, **(2008)**, pp. 3029–3032.

[12] Q. Liu, "Steganalysis of DCT-embedding based adaptive steganography and YASS", In J. Dittmann, S. Craver, and C. Heitzenrater, editors, Proceedings of the 13th ACM Multimedia & Security Workshop, **(2011)**, pp. 77–86.

[13] Q. Liu, A. H. Sung, M. Qiao, Z. Chen and B. Ribeiro, "An improved approach to steganalysis of JPEG images", Information Sciences, vol. 180, no. 9, **(2010)**, pp. 1643–1655.

[14] J. Kodovský and J. Fridrich, "Steganalysis in high dimensions: Fusing classifiers built on random subspaces", In N. D. Memon, E. J. Delp, P. W. Wong, and J. Dittmann, editors, Proceedings SPIE, Electronic Imaging, Security and Forensics of Multimedia XIII, vol. 7880, pp. OL 1–13.
[15] Z. Sun, M. Hui and C. Guan, "Steganalysis Based on Co-occurrence Matrix of Differential Image", Intelligent Information Hiding and Multimedia Signal Processing, August **(2008)**, pp. 1097–1100.
[16] J. Kodovsky, J. Fridrich and V. Holub, "Ensemble classifier for steganalysis of digital media", IEEE Trans. Inf. Forensics Security, vol. 7, no. 2, April **(2012)**, pp. 432–444.
[17] P. Bas, T. Filler and T. Pevny, "Break Our Steganographic System the ins and outs of organizing BOSS", In proceedings of Information Hiding Conference, Prague, Available at http://exile.felk.cvut.cz/boss /BOSSFinal/index. php?mode=VIEW&tmpl=materials, **(2011)**.
[18] A. Westfeld, "F5-A steganographic algorithm: high capacity despite better steganalysis", Proceedings of Fourth International Workshop on Information Hiding, April **(2001)**, pp. 289–302.
[19] Fridrich J., Goljan M. and Soukal D., "Wet paper codes with improved embedding efficiency", Information Forensics and Security, IEEE Transactions on, vol. 1, no. 1, March **(2006)**, pp. 102,110.
[20] Windeatt T., "Accuracy/diversity and ensemble MLP classifier design", IEEE Transactions ON Neural Networks, ISSN 1045-9227, vol. 17, no. 5, **(2006)**, pp. 1194 – 1211.

Appendix

A. Sample raw images from BOSS database.

Authors

J. Anita Christaline, she was born in Dindigul, India in 1974. She comp leted her B.E. Degree in Electronics and Communication Engineering in 1996 from Bharathidasan University, Trichy. She is recipient of National level GATE score during 2003. She received her M. E. Degree in Applied Electronics in the year 2006 from Anna University, Chennai. She is currently working as Assistant Professor in SRM University, Chennai where she is pursuing her Doctoral Degree. Her area of interest is Steganography and Steganalysis by computational intelligence. She has published many paper in this field. She is member IEEE and IET.

R. Ramesh, He was born in Kanyakumari, India, 1976. He received the B. E. Degree in Electronics and Communication Engineering in 1998 and the M. E. degree in Communication Systems in 2000, both from Madurai Kamaraj University, India. He has been awarded Doctoral degree in the SRM University in the year 2009 for his research work on Testing the Stability of two dimensional recursive filters. He is currently working as a Professor in the Department of Electronics and Communication Engineering at Saveetha Engineering College, Chennai. India. His current research interests concern digital signal processing particularly to find the stability of two dimensional recursive digital filters.

D. Vaishali, was born in Pune, India, 1969.She received B. E. Degree in Electronics and Telecommunication Engineering in 1994 and the M.E. degree in Communication Systems in 2002, both from Pune University, India. She is a research scholar in the SRM University and her research work is progressing in the field of Image Processing with wavelet Transforms. She is currently working as an Assistant Professor in the Department of Electronics and Communication Engineering at SRM University, Chennai. India. She is member IEEE and IET.

