Blind Steganalysis for JPEG Images using SVM and SVM-PSO Classifiers

Deepa D.Shankar, Prabhat Kumar Upadhyay

***Abstract:** Blind steganalysis or the universal steganalysis helps to identify hidden information without previous knowledge of the content or the embedding technique. The Support Vector Machine (SVM) and SVM- Particle Swarm Optimization (SVM-PSO) classifiers are adopted for the proposed blind steganalysis. The important features of the JPEG images are extracted using Discrete Cosine Transform (DCT). The kernel functions used for the classifiers in the proposed work are the linear, epanechnikov, multi-quadratic, radial, ANOVA and polynomial. The proposed work uses linear, shuffle, stratified and automatic sampling techniques. The proposed work employs four techniques for image embedding namely, Least Significant Bit (LSB) Matching, LSB replacement, Pixel Value Differencing (PVD) and F5 and applies 25% embedding. The data to the classifier is split as 80:20 for training and testing and 10-fold cross validation is carried out.*

***Keywords—** Blind steganalysis, SVM, SVM-PSO, DCT, LSB, PVD, ANOVA1.0, Embedded Techniques, Cross validation.*

I. INTRODUCTION

In the modern world, personal data of individual are collected for many purposes, which include, banking transaction, government ID, military applications and employer enrolment and recognition. Though many techniques are being proposed and implemented, the security of information is at stake [1].In steganography, the user cannot recognise the existence of a covert message even though the decoding of messages is not much difficult in steganography [2].

Steganography literally means 'secret writing'. It adapts some procedures and methods to embed private information into another data so as to deny access for unapproved users from viewing and recognizing the hidden data [3][4]. The textual format used in the research contains text from embedded messages, whereas the audio format encrypts the audio files and graphical formats contains encrypted images and videos [5][6]. The power to identify the existence of steganography is referred as steganalysis, which can be categorized into targeted detectors and universal detectors. The former is intended to estimate objects of certain embedding tasks, and, the latter do not have previous information about a specific steganographic scheme [7].If the practice of steganalysis is capable of finding the existence of a message with a higher rate of success in comparison with that of random guess, then the related system is said to be broken [8].

In this paper, low embedding of 25 % is used with JPEG images. The novel idea behind the research is to consider six different kernels and four type of sampling for feature based steganalysis. The comparative study of classification is done using two separate classifiers –SVM and SVM-PSO.

II. RELATED WORK

The core of Steganalysis lies in the classification between stego and original images. The features extracted for this technique is vital for the classifiers [9]. The most imperative kinds of classifiers applied in steganalysis contain Ordinary Least Squares (OLS) regression, Fisher Linear Discriminant (FLD) type for fine-tuning the classifiers, Quadratic Discriminant Analysis (QDA) type , Support Vector Machines (SVM) to be used with uninterrupted features, Bayesian Belief Networks (BBNs) for features that are distinct or discretized, and Naive Bayes Classifiers (NBCs) for feature vectors that are assorted [10]–[15].

Steganography applying Bit-Plane Complexity Segmentation, BPCS, utilizes multifold bit planes and is a kind of digital method, in which there is a high possibility of embedding an enormous degree of data [16]. The embedded data in JPEG images can be modelled with high accuracy with the help of statistical model having an optimum number of discrete cosine transform (DCT) coefficients [17][18].The other variety of cataloguing that the steganalysis comprises is the spatial domain where the message embedding happens directly into the pixel intensity of the image [19].

An evolutionary algorithm, using fuzzy if-then rules has been employed in the Steganography Pattern Discovery (SPD) for the stegano image signature extraction [20]. The features, which are statistical in nature representing the 3D features, are mined from the stego as well as the cover image of 3D objects and are given as input to the classifiers [21]

Amongst the various steganographic methodologies, LSB replacement, Pixel Value Differencing, LSB matching and F5 are employed in this paper. Veena and Arivazhagan [22] presented a LSB centred algorithm with less training samples and low dimensions by combining the global as well as the local features by means of Discretized-All Condensed Nearest Neighbour model.

A learning dependent LSB matching steganalysis is suggested by Xia et al.[23] with SVM classifiers where the three significant histogram features of the images are used to train the classifier. The difference value concerning two successive pixels of a block adopted to find the amount of bits that can be inserted using Pixel Value Differencing (PVD) scheme, which is a transform domain based method. Wu et al.

Revised Manuscript Received on September 22, 2019.

Deepa D.Shankar, Research Scholar, Banasthali Vidyapith, Rajasthan, India, sudee99@gmail.com

Prabhat Kumar Upadhyay, Department Of Electrical and Electronics Engineering, Birla Institute of Technology, Mesra, Ranchi, India

[24] amalgamated LSB replacement and PVD methodologies to enrich the excellence of stego image. Malathi and Gireeshkumar [25] carried out research to incorporate a large payload with various techniques related to LSB and F5 algorithm. This paper thus included both the spatial (LSB techniques) and transform domains (PVD and F5) for steganalysis.

The success of statistical steganalysis relies on detecting the vital features, which are identical, and the statistical variations during the embedding of data [26]. The algorithms pertaining to steganalysis are designed to distinguish the changes with more precision and accuracy [27]. Choosing the appropriate types of features estimates the precision by curtailing the feature divergence for the different selected images and hence feature analysis becomes an integral part of blind steganalysis [28]. Feature extraction is done, the image was processed with Discrete Cosine Transform (DCT) and its factors are utilized for analysis [29].

AbdelQader & AlTamimi [30] have suggested a DCT approach for feature reduction and have applied these features to SVM linear classifier. The performance of SVM varies with the different type of kernel and the sampling process is not included in the work of the authors. The proposed work emphasizes the performance of SVM with six different kernels and four different samplings have been used in our work. In addition, the PSO algorithm is used with SVM in the proposed work.

Tanwar & Malhotrab [31] have detailed about the advantages and disadvantages of Genetic and Particle Swarm Optimization algorithms for the application of steganography. Our proposed work has provided the results of real experimentations. In addition, the various kernels and samplings have been used to analyze the optimum value and listed the results of experimentation.

Liu, et.al. [32] have carried out the steganalysis by different feature subsets. Each subset is given to a particular classifier model and the results of multi classifiers are fed to fusing SVM. The kernel function for the SVM is linear.

Deepa & Umarani [33] have discussed about the steganalysis by SVM classifiers with WOW stego images. They have considered only the spatial domain algorithms and not the transform domain algorithms. The proposed works have considered both the spatial, transform domain techniques, and has applied SVM and SVM-PSO classifiers. In addition, the SVM's performance is evaluated with regard to six various kernels.

III. METHODOLOGY

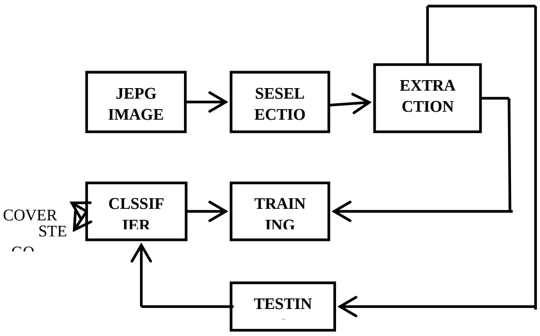
The recent literature emphasizes on the effect of machine learning techniques for the steganalysis applications but still only, a few literature has investigated this method, leaving an avenue to explore the research, based on machine learning. It is found that steganalysis using statistical approach would offer promising outcomes when the methods of machine learning for image classification is applied. The subsequent sections of this paper discuss the methodology and the experimental results for the proposed method.

The JPEG is highly preferred format for this research because of its great compression ratio and better optical quality of image. [34][35]. The transformation from DCT to frequency domain and the stages of quantization are considered lossy whereas DCT encoding with entropy is considered lossless compression [36].

The images are transformed using a Discrete Cosine Transform (DCT) to extract the conforming features of the image. These features are then given as input to the classifier. Support Vector Machine (SVM) and SVM-PSO are used in this research. JPEG with lossy compression is preferred in this research because of the fact that the DCT transforms the features into the low-frequency domain and the non-relevant data are deliberately discarded. The image segmentation is carried out through 8X8 blocks, which is followed by feature extraction. Normalisation is performed to improvise the efficacy of the algorithm.

The outline of steganalysis used is given as block diagram in figure 1.

Figure 1: Steganalysis System Architecture with SVM/SVM-PSO Classifier



3.1 Extraction of Features

The feature vectors are created with minimum dimensions without excluding any vital information pertaining to the image. This is performed by extracting the lower order features such as, first as well as second orders and the extended DCT as well as the Markov features, with final features summing up to 274. The single and dual histograms' standard deviation is accounted as first-order features.

A better perception of dispersal of coefficient values is obtained by the combination of individual and global histograms. The second order features are frequently inter block reliant features [37]. The statistical measures viz. co-occurrence, blockiness and variance are accounted for the second-order features.

3.2 Cross-validation and Classification

Cross-validation, which is otherwise referred to as rotation estimation, is a method of model authentication for evaluating the process of generalizing a dataset that is independent, from the outcomes of analysis through statistics. It is helpful in estimating the accuracy of predictive model performance [38].

When it is required to construct a machine learning design , the given data is split into training data and validation or test data[39]. The weightage of the partition which is allocated will vary the performance [40].To have consistency over the system performance, the training and testing are performed many times [41][42].

Classification is the process that is carried over after the feature extraction process. The images are classified related to their features, into different classes [43]. In steganalysis, the class labels may include, the cover image or image in which a message is embedded. That is, the image is either classified as a cover image or stego image [26]. In supervised machine learning, the algorithm is used to learn the mapping function between a given set of input and output variables [44]. The mapping function is approximated in such a way that the output data could be predicted with a new set of input variables [45]. After successful completion of training, an appropriate class label is predicted depending on the applied features. The next section gives a few details about the SVM classifier.

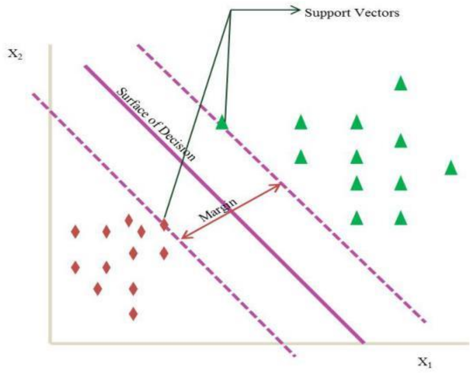


Figure 2: SVM Based Classification

The Support Vector Machine (SVM) aims to determine a hyperplane of the decision in a space with N-dimensions (N refers to the total number of features). The purpose of this hyperplane is the data point classification with accuracy and distinction. This is accomplished by increasing the margin between the given classes. The adjoining data points, that is, data points lying nearest to the decision surface are known to as support vectors. The distance between the surface of decision and the contiguous data point is referred to as margin. Though, many hyperplanes are possibly picked for splitting the classes, the goal of SVM is to identify a surface of decision with the highest margin. If the margin is less, noise will be introduced in the system and it hinders the suitable classification. Figure 2 represents the method of classification by using SVM classifiers. The SVM classifier performs with clarity as regard to the margin of separation, well suited to spaces of high dimensions, memory proficient, can be applied to binary as well as multi classifiers, and many kernels can be opted with these classifiers.

3.4 SVM – PSO

The optimization is carried out with the help of Particle Swarm Optimization (PSO) in SVM–PSO. Depending on the quality measure specified, the solution of the candidate (particle) is enhanced via successive iterations in PSO. Particle Swarm Optimization (PSO) is a calculation technique that is evolutionary based. Every possible solution is assumed as a particle in PSO and the particles are assigned with certain velocity to fly, covering the entire space of the problem. This paper utilizes the technique of PSO along with SVM in the classification.

3.5 Kernels

Kernels are helpful in calculating the feature mappings of large dimensions. The inner product of the transformed space is defined by the Kernel function to diminish the complication of determining the function of mapping. The various kernels used in this research paper include, the radial kernel, polynomial kernel, dot kernel, multi-quadric kernel, ANOVA kernel and Epanechnikov kernel.

The Radial Basis Function kernel, known to be the radial kernel, has the following kind of representation, $k(x,y) = exp\,(-g||x- y||^2\,)$,where g, is called gamma and is given by the gamma parameter of the kernel. The polynomial kernel, which is well-matched with normalized training data, is denoted mathematically as,

$k\ (x,\ y) = (x^*\ y\ +1)^p$, where the exponent p is the polynomial degree.

The dot kernel is described as

$k\ (x,y) = x^*y$

The dot kernel is nothing but the product of inner variables x and y.

The multiquadric kernel is defined by

$K\ (x,\ y) = (||\ x - y\ ||^2\ +c^2)^{0.5}$, where c is a constant.

The ANOVA kernel, whose performance is prominent in multidimensional problems, is defined as

$$k(x,y) = \sum_{k=1}^{n} \exp\left(-\sigma\left(x^k - y^k\right)^2\right)$$

Where σ can be derived from gamma, g; and d is the degree.

The Epanechnikov kernel, which is parabolic, is defined with the following equation,

$$k(u) = \frac{3}{4}(1-u^2) for\ |u| \leq 1$$

3.6 Standard Algorithms

Most of the algorithms applied for Image Steganography utilize an embedding technique called the Least Bit Embedding The LSB method incorporates a mild change in the least bit that changes the intensity value of pixels. This change is however not visible to the human eye or could not be identified by an individual.

3.6.1 Least Significant Bit (LSB) methods:

LSB practices utilize two types, which are replacement or substitution and the matching. In LSB replacement, the embedded data in its binary form is considered and the LSB in every byte is overwritten inside the image. On the other hand, for odd values of pixels, the value is decreased by 1. This method is found to be effective with gray scale images and referred to as LSB Matching or ±1 embedding.

3.6.2 Pixel Value Differencing (PVD):

It has been exposed that the images can hide more data than in smooth portions. This truth has led to Pixel Value Differencing (PVD) steganography. Initially, the image is segregated into various blocks with 1x 2 size. The difference or the variation among two pixels is calculated and this variation is changed with another fresh value by secreting a data inside it.

3.6.3 F5 Algorithm:

The F5 matrix embedding has three parameters viz. k, the number of bits to be embedded, n, the number of coefficients and c, and the number of variations to be performed for a group. At first, the PRNG seed is derived from user password to make a random walk across the cover images' DCT coefficients. The stream cipher encrypts k value with the help of PRNG, which is embedded with the message length at the start of stream of messages. The message body is implanted by means of matrix embedding, injecting k message bits to one clutch that contains (2k–1) constants by decreasing the total value by "one" in every group [46] [47] [48].

IV. EXPERIMENTAL RESULTS

4.1 Image Database

For the proposed blind steganalysis, two different image databases namely, the INIRA holidays dataset [49] and UCID image dataset [50] are considered. The details of the considered image database are provided before presenting the experimental setup. The details of the image dataset considered are elaborated here.

The details relevant to the experiment are listed in Table 1.

Table 1: Image Dataset details for Experiment

Name of Dataset	INRIA Holidays image dataset (for Training) UCID image dataset (for Testing/Validating)
Total Images Used	2300 standard images with different textures
Image Format	JPEG
Image Size	256 X 256 (Compressed)
No. of Training Images	1500
No. of Testing/validating images	800

4.2 Feature Extraction

The transformation phase is then followed by the extraction phase. The various statistical measures for first order, second order and special features decide the number of features to be extracted. The details of the extracted features are tabulated in Table 2.

Table 2: Table of Extracted Features

Type of Feature	Method	Total Extracted Features
First order	Individual Histogram	55
	Global Histogram	11
	Dual Histogram	99
Second order	Variance	01
	Blockiness	02
	Co-occurence	25
Markovian	-	81
Total Extracted Features		274

4.3 Training and Testing

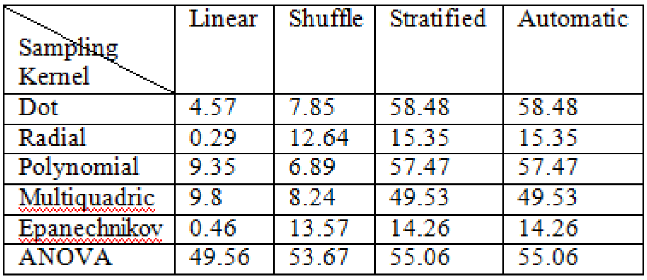
As pointed out in Table 1, the training dataset takes 1500 imageries from 2300 images and every image will contain 274 features. The classifier (SVM / SVM-PSO) is given a dataset comprising of labels either as stego image or as cover image. This helps in decreasing the count of false negatives as well as false positives that are utilized in the confusion matrix.

4.3.1 Results with SVM Classifier

SVM classifiers are adaptable with various kinds of kernels. Six different kernels have been tested with four different steganalysis techniques. A k-split cross-validation has been implemented for multiple split for training and testing with the value of split as 10 and the embedded percentage is 25% for all the experiments with the considered classifiers. The obtained results of various methods such as, LSB replacement, LSB Matching, F5 and PVD are produced in Tables 3-6.

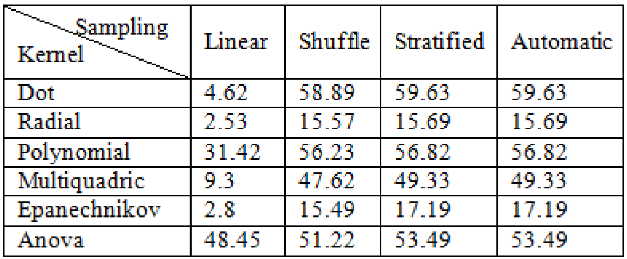
It is observed from Table 3, that the Dot kernel gives best results with respect to Shuffle, Stratifies and Automatic samplings. The subsequent better results are obtained with Polynomial and ANOVA kernels. It is also noted that the redial kernel is not suited for LSB replacement and ANOVA kernel provides a relatively good result with Linear.

Table 3: LSB Replacement Details with SVM



Sampling Kernel	Linear	Shuffle	Stratified	Automatic
Dot	4.57	7.85	58.48	58.48
Radial	0.29	12.64	15.35	15.35
Polynomial	9.35	6.89	57.47	57.47
Multiquadric	9.8	8.24	49.53	49.53
Epanechnikov	0.46	13.57	14.26	14.26
ANOVA	49.56	53.67	55.06	55.06

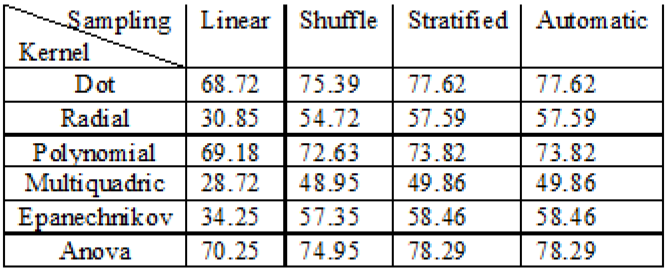
Table 4: LSB Matching Details with SVM



Sampling Kernel	Linear	Shuffle	Stratified	Automatic
Dot	4.62	58.89	59.63	59.63
Radial	2.53	15.57	15.69	15.69
Polynomial	31.42	56.23	56.82	56.82
Multiquadric	9.3	47.62	49.33	49.33
Epanechnikov	2.8	15.49	17.19	17.19
Anova	48.45	51.22	53.49	53.49

The experimental results from Table 4 reveal that the best results of LSB matching is achieved for all samplings except linear sampling, through Dot kernel. The ANOVA kernel yields a good result with linear sampling. As perceived in LSB replacement, the radial kernel is not suited for LSB matching technique also.

Table 5: F5 Details with SVM

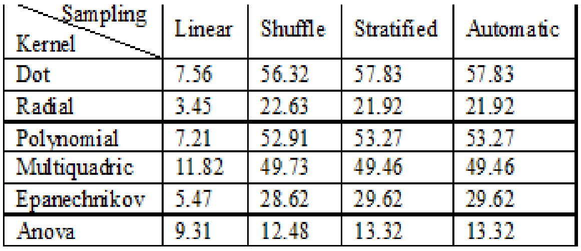


Sampling Kernel	Linear	Shuffle	Stratified	Automatic
Dot	68.72	75.39	77.62	77.62
Radial	30.85	54.72	57.59	57.59
Polynomial	69.18	72.63	73.82	73.82
Multiquadric	28.72	48.95	49.86	49.86
Epanechnikov	34.25	57.35	58.46	58.46
Anova	70.25	74.95	78.29	78.29

From Table 5, it is evident that the Dot kernel works well with samplings such as Shuffle, Stratified and Automatic. The ANOVA kernel performs well with respect to linear sampling. Unlike the previous two methods, the Radial kernel gives a nominal result for most of the samplings.

It can be noticed from Table 6 that Dot kernel is again performing well with all samplings except linear sampling. The other best results are observed with Polynomial and multiquadric kernels. No better result is observed with linear sampling and the results of the Radial kernel are not appealing.

Table 6: PVD Details with SVM



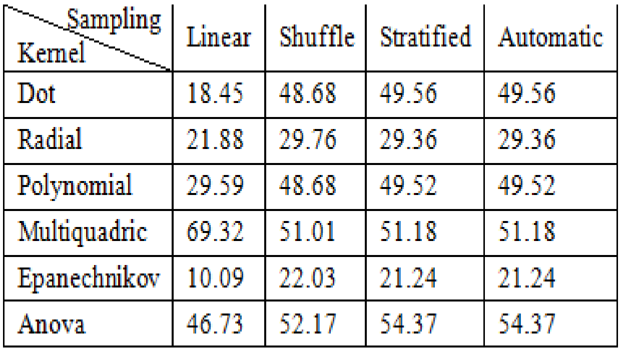
Sampling Kernel	Linear	Shuffle	Stratified	Automatic
Dot	7.56	56.32	57.83	57.83
Radial	3.45	22.63	21.92	21.92
Polynomial	7.21	52.91	53.27	53.27
Multiquadric	11.82	49.73	49.46	49.46
Epanechnikov	5.47	28.62	29.62	29.62
Anova	9.31	12.48	13.32	13.32

It is obvious from Tables 3-6 that the Dot kernel is performing extremely well irrespective of the method of steganalysis except for linear sampling.

4.3.2 Results with SVM Classifier with PSO (SVM-PSO)

Though the SVM classifier is one of the profound model found in steganalysis, there exist some issues related to picking the parameters of SVM, which in turn results in finding the optimal hyperplane. To address this issue of SVM, the particle swarm optimization algorithm is executed to choose the parameters of support vector machine. Tables 7-10 display the numerical results with 10 fold cross-validation and 25% embedding.

Table 7: LSB Replacement Details with SVM-PSO

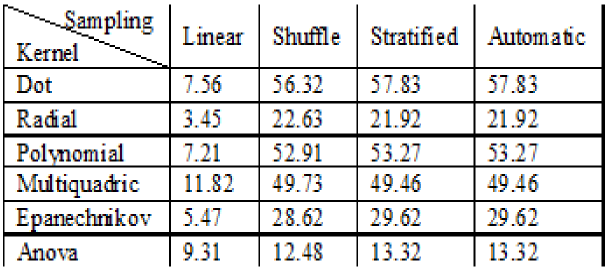


Sampling Kernel	Linear	Shuffle	Stratified	Automatic
Dot	18.45	48.68	49.56	49.56
Radial	21.88	29.76	29.36	29.36
Polynomial	29.59	48.68	49.52	49.52
Multiquadric	69.32	51.01	51.18	51.18
Epanechnikov	10.09	22.03	21.24	21.24
Anova	46.73	52.17	54.37	54.37

From Table 5, it is evident that the Dot kernel works well with samplings such as Shuffle, Stratified and Automatic. The ANOVA kernel performs well with respect to linear sampling. Unlike the previous two methods, the Radial kernel gives a nominal result for most of the samplings.

It can be noticed from Table 6 that Dot kernel is again performing well with all samplings except linear sampling. The other best results are observed with Polynomial and multiquadric kernels. No better result is observed with linear sampling and the results of the Radial kernel are not appealing.

Table 6: PVD Details with SVM



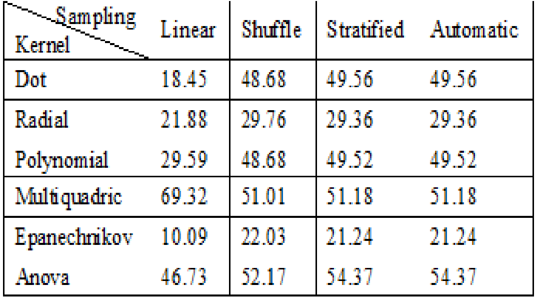
Sampling / Kernel	Linear	Shuffle	Stratified	Automatic
Dot	7.56	56.32	57.83	57.83
Radial	3.45	22.63	21.92	21.92
Polynomial	7.21	52.91	53.27	53.27
Multiquadric	11.82	49.73	49.46	49.46
Epanechnikov	5.47	28.62	29.62	29.62
Anova	9.31	12.48	13.32	13.32

It is obvious from Tables 3-6 that the Dot kernel is performing extremely well irrespective of the method of steganalysis except for linear sampling.

4.3.2 Results with SVM Classifier with PSO (SVM-PSO)

Though the SVM classifier is one of the profound model found in steganalysis, there exist some issues related to picking the parameters of SVM, which in turn results in finding the optimal hyperplane. To address this issue of SVM, the particle swarm optimization algorithm is executed to choose the parameters of support vector machine. Tables 7-10 display the numerical results with 10 fold cross-validation and 25% embedding.

Table 7: LSB Replacement Details with SVM-PSO

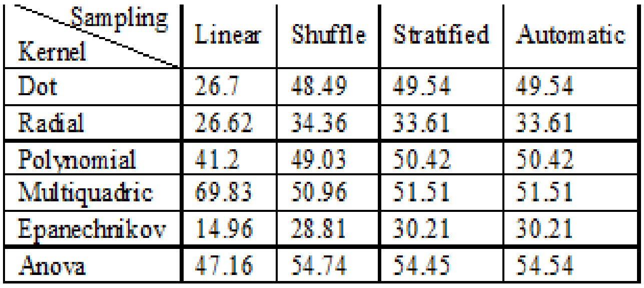


Sampling / Kernel	Linear	Shuffle	Stratified	Automatic
Dot	18.45	48.68	49.56	49.56
Radial	21.88	29.76	29.36	29.36
Polynomial	29.59	48.68	49.52	49.52
Multiquadric	69.32	51.01	51.18	51.18
Epanechnikov	10.09	22.03	21.24	21.24
Anova	46.73	52.17	54.37	54.37

Through the acquired results, the ANOVA kernel is found to be a good choice with regard to shuffle, stratified and automatic samplings. Altogether, a promising and reasonable performance is offered by multiquadric kernel irrespective of the samplings. All the other kernels performed nominally in comparison to SVM classifier (without PSO).

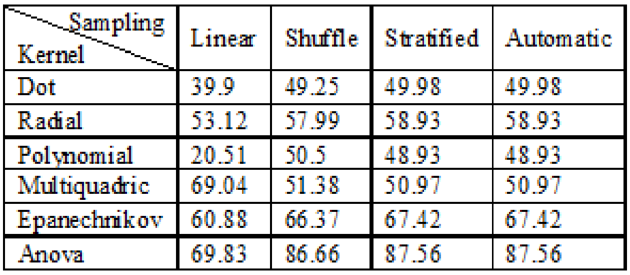
The shuffle, stratified and automatic samplings give good results in ANOVA kernel as shown in Table 8. The subsequent better results are obtained with Multiquadric kernel for the samplings except linear. The linear sampling gives better result with Multiquadric kernel. In total, it is evidenced that a nominal value is observed for all the samplings with Multiquadric kernel.

Table 8: LSB Matching Details with SVM-PSO



Sampling / Kernel	Linear	Shuffle	Stratified	Automatic
Dot	26.7	48.49	49.54	49.54
Radial	26.62	34.36	33.61	33.61
Polynomial	41.2	49.03	50.42	50.42
Multiquadric	69.83	50.96	51.51	51.51
Epanechnikov	14.96	28.81	30.21	30.21
Anova	47.16	54.74	54.45	54.54

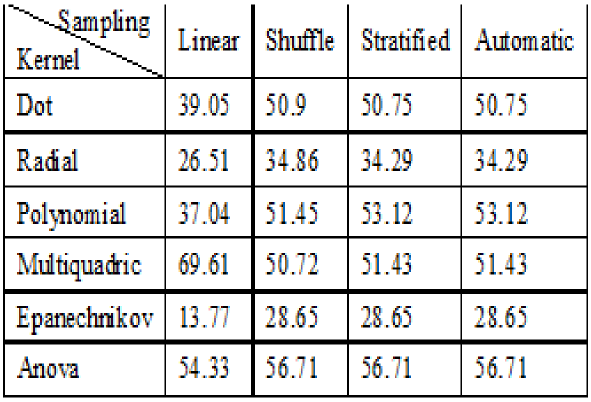
Table 9: F5 Details with SVM-PSO



Sampling / Kernel	Linear	Shuffle	Stratified	Automatic
Dot	39.9	49.25	49.98	49.98
Radial	53.12	57.99	58.93	58.93
Polynomial	20.51	50.5	48.93	48.93
Multiquadric	69.04	51.38	50.97	50.97
Epanechnikov	60.88	66.37	67.42	67.42
Anova	69.83	86.66	87.56	87.56

Better results are derived with ANOVA kernel for all the four samplings considered. The next better results are attained with Epanechnikov kernel except for linear sampling, whereas the Multiquadric kernel works well with linear sampling next to ANOVA kernel. As it is observed from Table 9, the ANOVA kernel outperforms all other kernels with respect to F5 with all four samplings.

Table 10: PVD Details with SVM-PSO



Sampling / Kernel	Linear	Shuffle	Stratified	Automatic
Dot	39.05	50.9	50.75	50.75
Radial	26.51	34.86	34.29	34.29
Polynomial	37.04	51.45	53.12	53.12
Multiquadric	69.61	50.72	51.43	51.43
Epanechnikov	13.77	28.65	28.65	28.65
Anova	54.33	56.71	56.71	56.71

Table 10 details the values of PVD with regard to 6 different kernels and four different samplings. Like the previous methods with SVM-PSO, the ANOVA kernel is well suited for Shuffle, Stratified and Automatic samplings. The better result for linear sampling is obtained through the Multiquadric kernel. The subsequent better results for all samplings except linear sampling are achieved with Polynomial and Multiquadric kernels respectively. Analyzing the Tables 7-10, the ANOVA kernel is the well-suited environment with SVM-PSO.

V. RESULT AND CONCLUSION

The method of feature extraction and the type of classifier that are employed decide the result of steganalysis. The spatial domain method comprising LSB replacement and LSB matching as well as the transform domain techniques entailing the F5 and PVD methods are considered in this paper. The machine learning algorithm, which is, the core of steganalysis is performed with 1500 training images and 800 validation images. This decision of selecting the number of training and testing images is aided with the help 10 fold cross-validation.The issues in the selection of SVM factors are addressed with the PSO algorithm.

The numerical results for both SVM and SVM-PSO classifiers for the four steganographic methods under 6 kernels and for four samplings have been provided. It is seen from the results that the classifiers with PSO provide better results across all the kernels and samplings. Superior performance has been observed with Dot kernel in case of SVM without PSO and with ANOVA in case of SVM-PSO.

REFERENCES

1) R. Ibrahim and T. Suk Kuan, Steganography Algorithm to Hide Secret Message inside an Image. 2011.
2) B. R. Kumar and P. R. . Murti, "Data Security and Authentication Using Steganography," Int. J. Comput. Sci. Inf. Technol., vol. 2, no. 4, pp. 1453–1456, 2011.
3) Nagaraj, V. Vijayalakshmi, and G. Zayaraz, "Modulo based Image Steganography Technique against Statistical and Histogram Analysis," IJCA Spec. Issue Netw. Secur. Cryptogr., pp. 34–39, 2011.
4) J.-S. Kang, Y. You, and M. Y. Sung, "Steganography using block-based adaptive threshold," in 2007 22nd international symposium on computer and information sciences, 2007, pp. 1–7.
5) F. A. Jassim, "A Novel Steganography Algorithm for Hiding Text in Image using Five Modulus Method," Int. J. Comput. Appl., vol. 72, no. 17, pp. 39–44, 2013.
6) M. Jiang, N. Memon, E. Wong, and X. Wu, "Quantitative steganalysis of binary images," in 2004 International Conference on Image Processing, 2004. ICIP '04., 2004, vol. 1, pp. 29–32.
7) M. U. Celik, G. Sharma, and A. M. Tekalp, "Universal image steganalysis using rate-distortion curves," in SPIE 5306, Security, Steganography, and Watermarking of Multimedia Contents, 2004, p. 467.
8) J. A. Christaline, R. Ramesh, and D. Vaishali, "STEGANALYSIS WITH CLASSIFIER COMBINATIONS," 2014.
9) Y. Miche, P. Bas, C. Jutten, A. Lendasse, and O. Simula, Extracting relevant features of steganographic schemes by feature selection techniques. 2007.
10) E. Elbaşı and A. M. Eskicioglu, Naïve Bayes Classifier Based Watermark Detection in Wavelet Transform. 2006.
11) Y. Miche, P. Bas, and A. Lendasse, "Using multiple re-embeddings for quantitative steganalysis and image reliability estimation," Aalto University School of Science and Technology, 2010.
12) J. J. Harmsen and W. A. Pearlman, "Kernel Fisher discriminant for steganalysis of JPEG hiding methods," in The International Society for Optical Engineering, 2004, p. 13.
13) G. Rajput, R. Agrawal, and N. Aggrawal, "Performance Evaluation of Exponential Discriminant Analysis with Feature Selection for Steganalysis," Def. Sci. J., vol. 62, no. 1, pp. 19–24, Jan. 2012.
14) H. Bhat, S. Krishna, P. D. Shenoy, K. R. Venugopal, and L. M. Patnaik, "HUBFIRE — A multi-class SVM based JPEG steganalysis using HBCL statistics and Fr Index," in 2010 International Conference on Security and Cryptography (SECRYPT), 2010, pp. 1–6.
15) X. Y. Yu and A. Wang, "Steganalysis Based on Regression Model and Bayesion Network," in 2009 International Conference on Multimedia Information Networking and Security, 2009, pp. 41–44.
16) E. Kawaguchi, "BPCS-Steganography – Principle and Applications," in Knowledge-Based Intelligent Information and Engineering Systems, 2005, pp. 289–299.
17) T. H. Thai, R. Cogranne, and F. Retraint, "Statistical Model of Quantized DCT Coefficients: Application in the Steganalysis of Jsteg Algorithm," IEEE Trans. Image Process., vol. 23, no. 5, pp. 1980–1993, May 2014.
18) A. Attaby, M. F. M. Mursi Ahmed, and A. K. Alsammak, "Data hiding inside JPEG images with high resistance to steganalysis using a novel technique: DCT-M3," Ain Shams Eng. J., vol. 9, no. 4, pp. 1965–1974, Dec. 2018.
19 M. Kalita and T. Tuithung, "A Comparative Study of Steganography Algorithms of Spatial and Transform Domain," IJCA Proc. Natl. Conf. Recent Trends Inf. Technol., pp. 9–14, 2015.
20 H. Sajedi, "Steganalysis based on steganography pattern discovery," J. Inf. Secur. Appl., vol. 30, pp. 3–14, Oct. 2016.
21 Z. Li and A. G. Bors, "Steganalysis of 3D objects using statistics of local feature sets," Inf. Sci. (Ny)., vol. 415–416, pp. 85–99, Nov. 2017.
22 S. T. Veena and S. Arivazhagan, "Quantitative steganalysis of spatial LSB based stego images using reduced instances and features," Pattern Recognit. Lett., vol. 105, pp. 39–49, Apr. 2018.
23 Z. Xia, L. Yang, S. Xingming, W. Liang, D. Sun, and Z. Ruan, A Learning-Based Steganalytic Method against LSB Matching Steganography, vol. 20. 2011.
24 H.-C. Wu, N.-I. Wu, C.-S. Tsai, and M.-S. Hwang, "Image steganographic scheme based on pixel-value differencing and LSB replacement methods," IEE Proc. - Vision, Image, Signal Process., vol. 152, no. 5, p. 611, 2005.
25 P. Malathi and T. Gireeshkumar, "Relating the Embedding Efficiency of LSB Steganography Techniques in Spatial and Transform Domains," Procedia Comput. Sci., vol. 93, pp. 878–885, 2016.
26 S. S. Chaeikar and A. Ahmadi, "Ensemble SW image steganalysis: A low dimension method for LSBR detection," Signal Process. Image Commun., vol. 70, pp. 233–245, Feb. 2019.
27 Wu, G. Feng, X. Zhang, and Y. Ren, "Unbalanced JPEG image steganalysis via multiview data match," J. Vis. Commun. Image Represent., vol. 34, pp. 103–107, Jan. 2016.
28 L. Wang, Y. Xu, L. Zhai, Y. Ren, and B. Du, "A posterior evaluation algorithm of steganalysis accuracy inspired by residual co-occurrence probability," Pattern Recognit., vol. 87, pp. 106–117, Mar. 2019.
29 R. Mehta and N. Agarwal, "Splicing Detection for Combined DCT, DWT and Spatial Markov-Features Using Ensemble Classifier," Procedia Comput. Sci., vol. 132, pp. 1695–1705, 2018.
30 AbdelQader and F. AlTamimi, "A Novel Image Steganography Approach Using Multi-Layers DCT Features Based on Support Vector Machine Classifier," Int. J. Multimed. Its Appl., vol. 9, no. 1, pp. 1–10, Feb. 2017.
31 R. Tanwar and S. Malhotrab, "Scope of Support Vector Machine in Steganography," IOP Conf. Ser. Mater. Sci. Eng., vol. 225, p. 012077, Aug. 2017.
32 P. Liu, F. Liu, C. Yang, and X. Song, "Improving Steganalysis by Fusing SVM Classifiers for JPEG Images," in 2015 International Conference on Computer Science and Mechanical Automation (CSMA), 2015, pp. 185–190.
33 Deepa and Umarani, "Steganalysis on images based on the classificationof image feature sets using svm classifier," Int. Acad. Sci. Eng. Technol., vol. 5, no. 5, pp. 16–24, 2016.
34 Bhasin and P. Bedi, "Steganalysis for JPEG Images Using Extreme Learning Machine," in Proceedings - 2013 IEEE International Conference on Systems, Man, and Cybernetics, SMC 2013, 2013, pp. 1361–1366.
35 J. Fridrich, T. Pevný, and J. Kodovský, "Statistically undetectable jpeg steganography," in Proceedings of the 9th workshop on Multimedia & security - MM&Sec '07, 2007, p. 3.
36 Y.-F. Tan, W.-N. Tan, and X. Guo, "Integrated lossy and lossless compression with LSB Insertion technique in steganography," 2013, p. 88781R.
37 Ashu and R. Chhikara, "Performance Evaluation of First and Second Order Features for Steganalysis," Int. J. Comput. Appl., vol. 92, no. 16, pp. 17–22, Apr. 2014.
38 D. Berrar, "Cross-Validation," in Encyclopedia of Bioinformatics and Computational Biology, Elsevier, 2019, pp. 542–545.
39 Y. Kokkinos and K. G. Margaritis, "Managing the computational cost of model selection and cross-validation in extreme learning machines via Cholesky, SVD, QR and eigen decompositions," Neurocomputing, vol. 295, pp. 29–45, Jun. 2018.
40 C. Bergmeir, R. J. Hyndman, and B. Koo, "A note on the validity of cross-validation for evaluating autoregressive time series prediction," Comput. Stat. Data Anal., vol. 120, pp. 70–83, Apr. 2018.
41 L. Xu et al., "Representative splitting cross validation," Chemom. Intell. Lab. Syst., vol. 183, pp. 29–35, Dec. 2018.
42 G. Jiang and W. Wang, "Error estimation based on variance analysis of k -fold cross-validation," Pattern Recognit., vol. 69, pp. 94–106, Sep. 2017.
43 X. Hou, T. Zhang, L. Ji, and Y. Wu, "Combating highly imbalanced steganalysis with small training samples using feature selection ," J. Vis. Commun. Image Represent., vol. 49, pp. 243–256, Nov. 2017.
44 D. C. G. Pedronette, Y. Weng, A. Baldassin, and C. Hou, "Semi-Supervised and Active Learning through Manifold Reciprocal kNN Graph for Image Retrieval," Neurocomputing, Feb. 2019.
45 M. Castelli, L. Vanneschi, and Á. R. Largo, "Supervised Learning: Classification," in Encyclopedia of Bioinformatics and Computational Biology, Elsevier, 2019, pp. 342–349.
46 M. Tang, M. Fan, and G. Wang, "An Extential Steganalysis of Information Hiding for F5," in 2010 2nd International Workshop on Intelligent Systems and Applications, 2010, pp. 1–4.

47 J. Fridrich, M. Goljan, and D. Hogea, “Steganalysis of JPEG Images: Breaking the F5 Algorithm,” in Information Hiding, 2003, pp. 310–323.

48 J. A. Briffa, H. G. Schaathun, and A. W. A. Wahab, “Has F5 really been broken?,” in 3rd International Conference on Imaging for Crime Detection and Prevention (ICDP 2009), 2009, pp. P17–P17.

49 H. Jegou, M. Douze, and C. Schmid, “Hamming Embedding and Weak Geometric Consistency for Large Scale Image Search,” in Proceeding ECCV ’08 Proceedings of the 10th European Conference on Computer Vision: Part I, 2008, pp. 304–317.

50 G. Schaefer and M. Stich, “UCID - An uncompressed colour image database,” Proc. SPIE - Int. Soc. Opt. Eng., vol. 5307, pp. 472–480, 2004.