10.48047/jocaaa.2019.27.05.02

Enhancing Image Steganalysis Performance with Local Binary Pattern for Feature Extraction

[1]Reeta Bhardwaj, [2]Rajeev Kumar*, [3]B. T. Ramaiah
[1,2]Assistant Professor, DAV Institute of Engineering & Technology, Jalandhar-144008, Punjab, India.
[3]Research Scholar, I.K.Gujral Punjab Technical University, Kapurthala-144603, Punjab, India.
[1]er.reeta@gmail.com, [2]rajeev.daviet@gmail.com, [3]bridgehead17@gmail.com.
*Corresponding author: rajeev.daviet@gmail.com

***Abstract*- Steganography techniques typically seek to conceal more sensitive data under cover photos. While many of these techniques ensure that the visual quality of the resulting stego image remains undetectable, some can also preserve the global structure of the original cover image. However, these methods often have a limited capacity for embedding secret information. Steganography is become a major barrier for digital forensics. However, the primary purposes of steganalysis are information extraction and stego detection from the covers. The increasing complexity of steganography combined with the increasing power of steganalysis algorithms has made it more difficult to build methods that perform significantly better. Then, an artificial bee colony-based optimization approach is employed to choose the best attributes. The DCT-extracted features are sensitive to changes in the direction of light, and the DCT coefficient's value is not spatially invariant. Every pixel in LBP is represented as a collection of adaptive local binary pattern histograms that are computed across a circle encircling the pixel. The suggested method utilizes local binary patterns to conceal the secret image's bits, ensuring that the resulting stego image maintains the local relationships present in the cover image. The performance of this steganography approach has been evaluated on various image types to demonstrate its robustness. A comparison with state-of-the-art LSB-based steganography methods illustrates the effectiveness of feature-based image steganography.**

Keywords **– steganography, steganalysis, stego, artificial bee colony, discrete cosine transform, local binary pattern.**

1. INTRODUCTION

Information System Security (encryption, watermarking, and steganography) is a discipline that ensures the confidentiality, integrity, and availability of information and services. Steganalysis is a form of attack, that always tries to break the security. Steganography's ultimate goals and key elements distinguish it from comparable techniques like cryptography and watermarking. Steganography conceals the existence of the message itself, making it impossible for an observer to determine where it is.

While steganalysis is the art of figuring out the existence of the secret communication, steganography is the science of secure communication in which the communication's presence cannot be discovered. Most of the time, processing a large volume of data requires a lot of computational resources and execution time. Consequently, pre-processing must be used, since this may control the execution duration and computing resources. In this work, we provide a novel feature-based blind steganalysis technique for separating stego pictures from JPEG-formatted cover (clean) images. To this end, we introduce an enhanced Artificial Bee Colony (ABC) as the

10.48047/jocaaa.2019.27.05.02

basis for our feature selection method. The social behaviour of honeybees in their quest for the ideal food source serves as an inspiration for the ABC algorithm. The performance of the classifier and the chosen feature vector's dimension are dependent on the use of wrapper-based techniques in the suggested approach. Two JPEG photos are used for the tests. Comparing the suggested steganalysis procedure to other methods now in use, experimental findings show that it is successful.

The initial stage of steganalysis is called blind steganalysis, or the identification of stego pictures. Instead of attempting to identify a specific concealing technique, blind steganalysis correctly identifies things that may contain concealed information. Therefore, blind steganalysis is more useful than particular steganalysis approaches. It uses fuzzy if-then rules and an evolutionary strategy to extract the signature of stego pictures versus clean images. Using the newly acquired knowledge, appropriate trained steganalysis models may be used, resulting in very accurate stego picture detection. Swarm intelligence (SI) has recently attracted a lot of interest in related domains as well.

2. LITERATURE WORK

The analysis of sentiment has been the subject of several studies and research projects in recent years. Numerous scholarly articles on sentiment analysis utilizing lexical, hybrid, and machine learning approaches have been published. Below is a brief analysis of the numerous papers that have been cited.:

Digital media files can be used as cover objects in a system proposed by Abhilasha et al. [1] for the detection of least significant bit (LSB) steganography. BMP pictures are used in LSB in its most basic version because they provide lossless compression. Regretfully, a very huge cover picture would be needed in order to conceal a hidden message inside a BMP file. 800 x 600-pixel BMP pictures are no longer often seen online and may raise red flags. LSB steganography has evolved to work with several image file formats as a result. Steganography is the process of embedding data into a vessel or container such that it seems to be empty. There are several carrier boats to choose from, including executable files, music snippets, and digital pictures. Every steganographic method now in use has a finite amount of information concealing ability. This method embeds secret information in the bit- planes of the vessel using an image frame from a movie as the vessel data. It is possible to replace all of the "noise-like" areas in the bit-planes of the main image frame of video with secret data without sacrificing the quality of the picture.

According to Amira and colleagues [2], Computed Tomography (CT) is a crucial medical imaging technique used to assess a variety of illnesses, including vascular lesions and malignancies. However, speckle noise distorts the CT scans and casts doubt on the interpretation of the clinical data. Therefore, noise reduction and sharp/clear pictures are essential for proper diagnosis, and this requires medical image enhancement. In this study, a log transform-based optimization framework approach for improving medical images is suggested. To attain optimization, the best parameter values for the log transform are found using a well- known meta-heuristic technique called the Cuckoo search (CS) algorithm. The suggested method's effectiveness is examined using a dataset of poor contrast CT images. In addition, the outcomes unequivocally demonstrate that the CS-based strategy outperforms PSO in terms of convergence and fitness values, with the CS achieving a quicker convergence rate, demonstrating the effectiveness of the CS-based methodology. Lastly, Image Quality Analysis (IQA) validates the suggested augmentation technique's resilience.

10.48047/jocaaa.2019.27.05.02

An overview of the Cuckoo Search optimization method, created by Yang and Deb in 2009, is provided by Venkata et al. [3]. It is utilized in addressing optimization issues. The cuckoo bird species, which deposits its eggs in the nests of other host birds, served as its inspiration. The primary driving force behind the creation of a novel optimization algorithm is the laying and breeding of cuckoo eggs. The convergence rate, accuracy, and efficiency are all increased by this optimization process. Numerous applications and categories of the Cuckoo search are examined.

A thorough overview of feature selection techniques to maximize the performance of available models can be found in Saurav et al. [4]. These are the general classifications that are frequently applied to feature selection. In order to improve and streamline model performance, the study explores a number of subset feature space approaches and strategies. Machine learning algorithms have little influence on the feature selection process. Rather, characteristics are chosen based on how well they correlate with the end variable according to the results of several statistical tests. The article provides an explanation of the wrapper, embedded, and filter techniques.

A novel method for picture steganalysis employing Artificial Bee Colony for feature selection (FS) is put forth by Hedieh et al. [5]. While steganalysis is the art of figuring out the existence of the secret communication, steganography is the science of secure communication in which the communication's presence cannot be discovered. Most of the time, processing a large volume of data requires a lot of computational resources and execution time. Consequently, a pretreatment step must be used, which can control the execution duration and computing resources. In this work, we provide a novel feature-based blind steganalysis technique for separating stego pictures from JPEG format cover (clean) images. To this end, we introduce an enhanced Artificial Bee Colony (ABC) as the basis for our feature selection method. The social behavior of honeybees in their quest for the ideal food source serves as an inspiration for the ABC algorithm. The performance of the classifier and the chosen feature vector's dimension are dependent on the use of wrapper-based techniques in the suggested approach. Two sizable datasets of JPEG pictures are used in the studies.

Goel et al. [6], The assault on data concealing problem is presented in the study. Targeted Steganalytic Attacks are the goal of the first strategy. The primary focus of the study is on targeted attacks based on first order data. Two techniques that can maintain an image's first-order statistics after embedding have been introduced. The suggested technique increases the security of the algorithms against targeted assaults by retaining the picture statistics, as demonstrated by the experimental results. The second strategy seeks to thwart blind steganalytic attacks, particularly those that use calibration to attempt to infer a cover picture model from the stego image.

Meenpaa et al. [7], The local binary pattern operator, which converts an image into an array or picture of integer labels characterizing the image's small-scale look, is introduced in this chapter as an image operator. Subsequent picture analysis uses these labels or their statistics, most often the histogram. The operator was initially intended for monochrome still photos, but it has since been expanded to include color (multi-channel) images, movies, and volumetric data as well. The several iterations of the real LBP operator in the spatial domain are covered in this chapter. The 3 × 3 pixel block of a picture is where the original version of the local binary pattern operator operates. To determine a label for the center pixel, the pixels in this block are thresholded by the value of the center pixel, multiplied by powers of two, and then added together. Given that the neighborhood has eight pixels, a total of 28 = 256 distinct labels may be produced based on how the center and the neighborhood's pixels compare in terms of grey values.

Chen et al. [8], The study presents a new and very fast JPEG steganalysis technique that makes use of the JPEG coefficients' intrablock and interblock correlations. Steganalysis of JPEG images

10.48047/jocaaa.2019.27.05.02

has gained more and more interest lately. Each difference JPEG 2-D array's transition probability matrix is computed to take use of the intrablock correlation; similarly, the "averaged" transition probability matrices for those difference mode 2-D arrays are computed to take advantage of the interblock correlation. These matrices' constituents are all utilized as steganalysis characteristics.
In Babu et al. [9], A Cuckoo Search Algorithmin with Morphological Operation picture enhancement technique is presented coupled this study. Digital pictures are created these days for use in several image processing applications. A few industries that use image processing applications include manufacturing, computer interfaces, machine vision, compression for storage, and more. The results of the experiments show that the suggested method produces original colour photographs free of noise and uses an adaptive procedure to improve image quality.
Ladha et al. [10] When it comes to large dimensional datasets in particular, feature selection is a crucial subject in data mining. Subsets of the features that are accessible from the data are chosen for the application of a learning algorithm in a process called feature selection, which is also referred to as subset selection in machine learning. The subset that maximizes accuracy while having the fewest dimensions is the optimal one. We eliminate the remaining, superfluous dimensions. One of the two strategies to escape the curse of dimensionality is to do this crucial pre-processing step (the other being feature extraction). Forward selection and reverse selection are the two methods used in feature selection. In the fields of pattern recognition, statistics, and data mining, feature selection has been a busy study topic.

3. PROPOSED WORK

The suggested method is a novel blind image steganography approach that uses the cover image's binary statistics to efficiently disguise the payload while maintaining the local statistical structure. The payload is first divided into 8 bits and the Local Binary Pattern (LBP) is taken out of the cover image's immediate vicinity in order to embed it. The final binary string is created by pairwise shuffling the X-ORed value, which is obtained by XORing these two binary values. To maintain the local structure of the cover in the stego picture, this binary string is inserted into the cover and each transformed pixel is synchronized using.
The proposed method contains 3 major steps

- Generation of dataset
- Feature extraction
- Feature Optimization

Step 1: Generation of dataset

The dataset must first be created using cover pictures, and then the Least Significant Bit (LSB) approach must be used to embed the secret message into the cover image in order to produce the stego images. The second stage, or feature extraction, is then given these stego pictures. The suggested steganography technique uses LBP to encode the cover picture while investigating the local relationships between its pixels. LBP is a binary pattern that is calculated by contrasting eight nearby pixels with the centre pixel of a 3 × 3 mask. The visual properties of the image are represented by this 8-bit pattern, which varies in response to non-uniform environmental changes. Fig. 1(a) displays the embedding process flow diagram. The size N × N cover picture, represented by C, is composed of nonoverlapping 3 × 3 pieces.

10.48047/jocaaa.2019.27.05.02

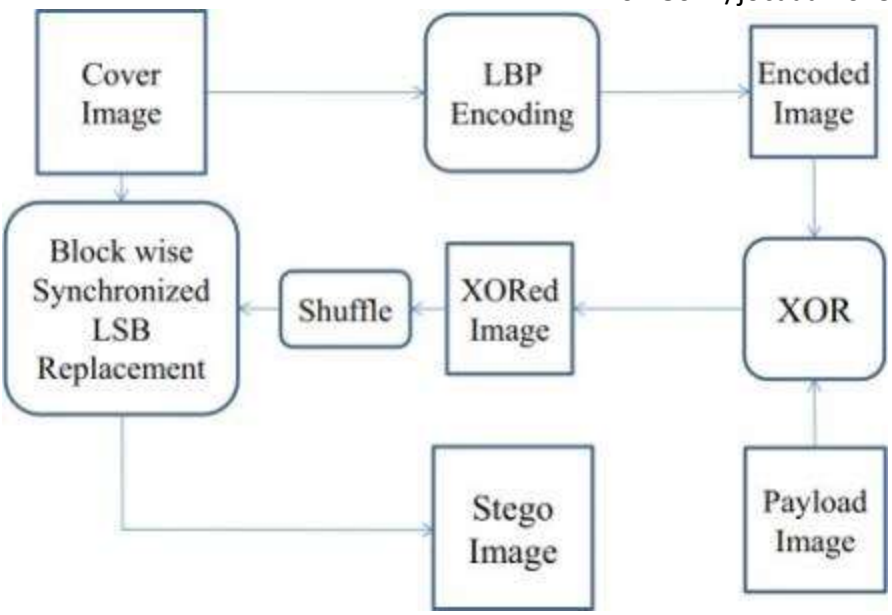

Figure 1: Proposed method of Generation of datasets

Step 2: Feature Extraction Using LBP

In computer vision, a kind of visual descriptor called Local Binary Patterns (LBP) is utilized for categorization. The specific instance of the suggested Texture Spectrum
model is LBP. Since then, it has been discovered to be a potent feature for texture classification; moreover, it has been discovered that combining LBP with the Histogram of Oriented Gradients (HOG) descriptor significantly enhances the detection performance on certain datasets. A comparison of the original LBP's several enhancements in the area of background subtraction.

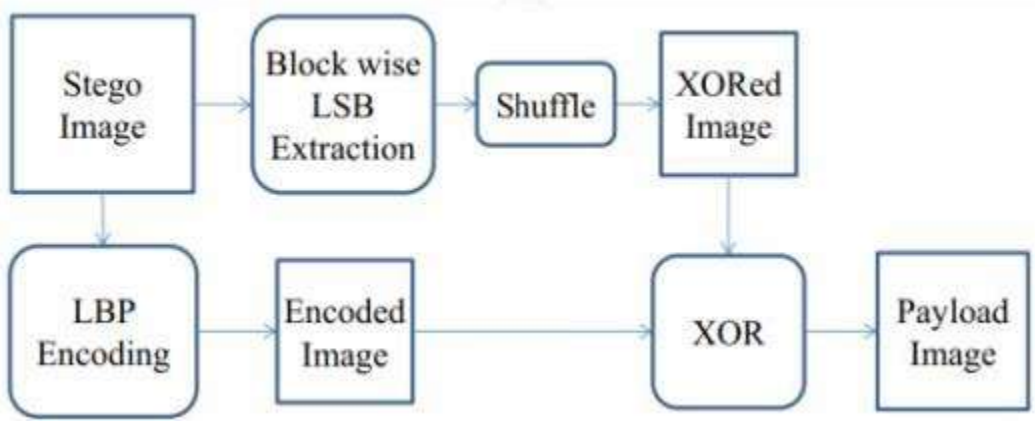

Figure 2: Proposed method Extraction process

10.48047/jocaaa.2019.27.05.02

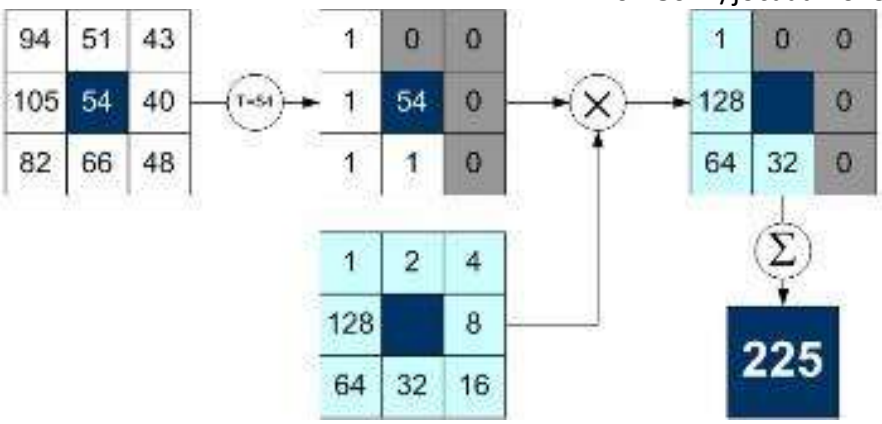

Figure 3: The stages of LBP calculation

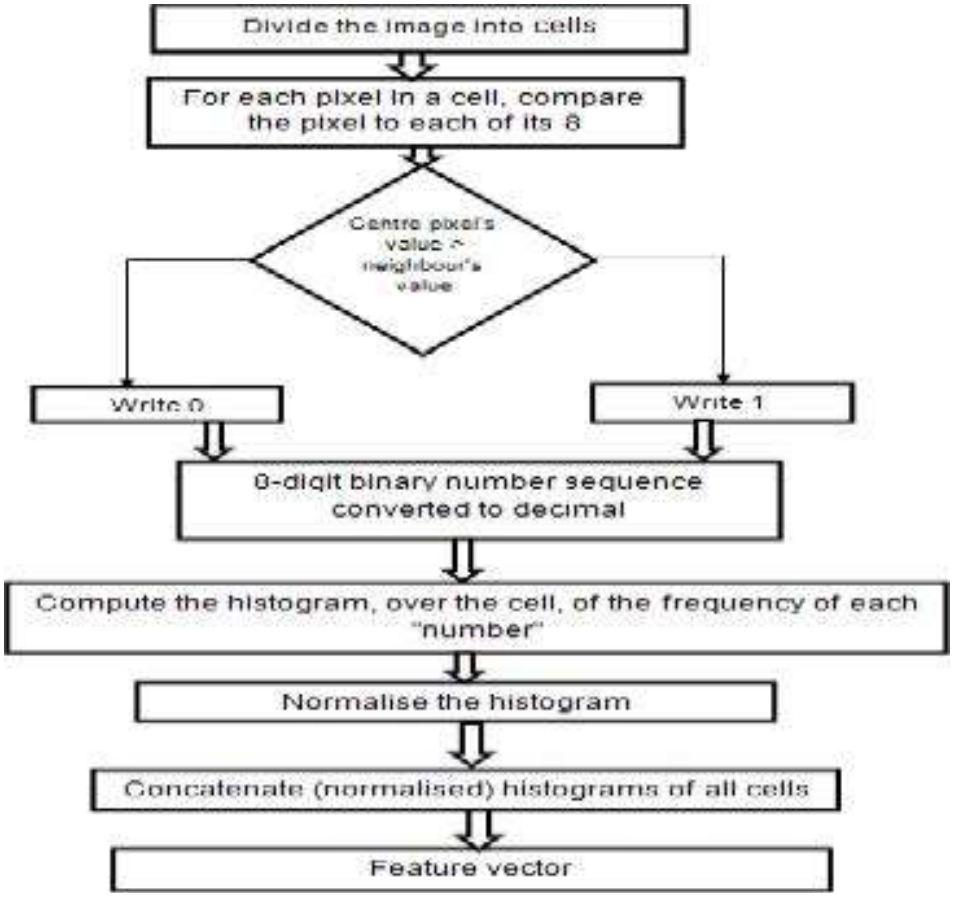

Figure 4: Flow Chart of LBP Vector

The LBP feature vector, in its simplest form, is created in the following manner:

- Acellularized the window under examination (e.g., 16x16 pixels for each cell).

10.48047/jocaaa.2019.27.05.02

- Compare each pixel in a cell to each of its eight neighbours, starting from the left-top, moving left- middle, moving left-bottom, moving right-top, etc. Trace the pixels in a circle, either in a clockwise or counterclockwise direction.
- Write "0" in the cases when the value of the center pixel is higher than that of the neighbour. If not, type "1". This yields an 8-digit binary number, which is conveniently translated to decimal.
- Compute the frequency histogram of each "number" that occurs over the cell (that is, any combination of pixels that are smaller or bigger than the centre).
- You may think of this histogram as a 256- dimensional feature vector.
- You can choose to normalize the histogram.
- Concatenate each cell's normalized histograms. A feature vector for the full window is therefore obtained.

Step 3: Feature Optimization using Artificial Bee Colony

Inspired by the clever actions of honey bees, Dervis Karaboga developed one of the most modern algorithms, Artificial Bee Colony (ABC), in 2005. It employs just standard control parameters, including colony size and maximum cycle number, and is as straightforward as Particle Swarm Optimization (PSO) and Differential Evolution (DE) algorithms.
Artificial bees in the ABC system fly about in a multidimensional search space, and some (employed and observer bees) make decisions about where to get food based on their own and their nest members' experiences. They also modify their locations. Some (scouts) fly and select food sources at random without consulting their prior knowledge. They remember the new position and forget the old one if the nectar amount of a new source is greater than that of the prior one.
Three processes make up each search cycle: locating the scout bees and guiding them toward potential food sources; placing the worker and observer bees onto the food sources and estimating their nectar quantities.

I. One potential fix for the issue that has to be optimized is a food source position.
II. The quality of the solution is correlated with the quantity of nectar a food source produce.
III. A probability-based selection procedure is used to position observers in relation to the food sources.
IV. The likelihood that a food source will be favored by observers rises in tandem with the nectar content of the food source.
V. The scouts are distinguished by poor average food source quality and low search expenses.
VI. A control parameter named "limit" governs the selection.
VII. A food source is abandoned and the hired bee becomes a scout if a solution representing it is not improved after a predefined number of tries.

4. EXPERIMENTAL RESULTS

A series of steganalysis tests were conducted in order to evaluate the effectiveness of the suggested FS algorithm and compare it with other established FS techniques as well as the MBEGA method. In experiments, we make use of the breaking out steganography system (BOSS) 1.01 together with a database of grayscale images. 0.4 hidden text embeddings per pixel is the default setting. Ten thousand cover pictures and ten thousand stego images are in this database. Out of the 20,000 photos on the BOSBASE site, 900 were used.

10.48047/jocaaa.2019.27.05.02

The suggested technique uses eight comparisons to generate LBP code in order to incorporate a single payload pixel. To calculate the updated payload pixel, eight bit-wise X-OR operations are required, and to compute the shuffled pixel, four circular shifts of length two bits are needed. To produce the intermediate stego pixel intensities, the LSB of the cover must be replaced with a maximum of eight additions and subtractions.

The suggested approach calls for a maximum of 8 additions and subtractions to synchronize the stego pixels. Therefore, the suggested technique may encapsulate a single payload pixel with a maximum of 28 basic operations.

The local structure of the picture is significantly altered by the steganography techniques, which embed the secret bits into the cover image. Local associations depend on the image's visual qualities and features, including backdrop, object form, and lighting.

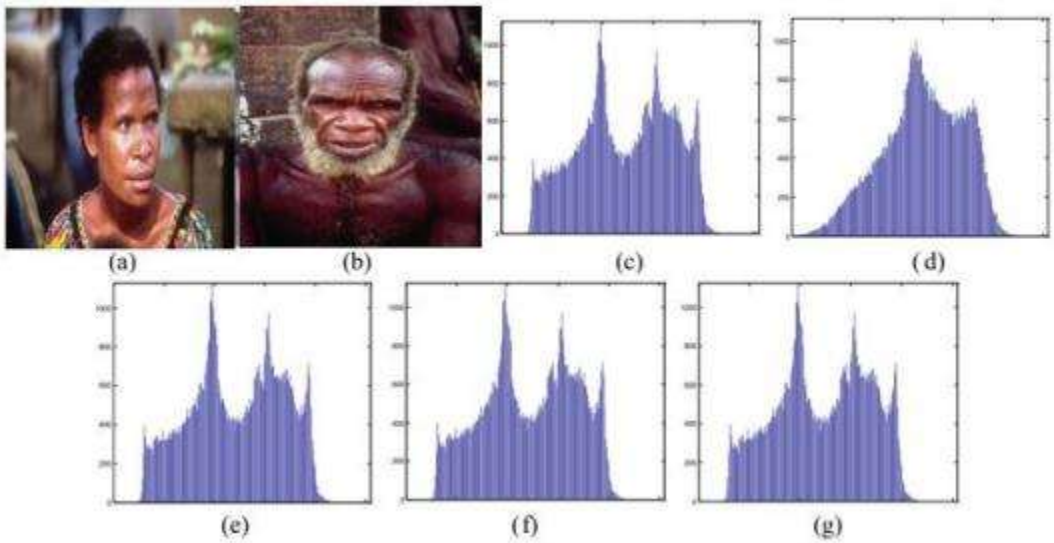

Figure 5: a) The cover image; b) The payload; c) The cover histogram; d) The payload histogram; e) The stego image histogram with 30% embedding; f) The stego image histogram with 40% embedding; g) The stego image histogram with 50% embedding

The addition of the hidden bits changes these properties. Robust feature analyses are employed to detect these modifications, which is capable of quickly identifying the existence of classified material in the cover photo. We have not yet encountered any steganography techniques that maintain the local structures that correspond to the features of the cover picture. The suggested technique calculates one such binary structure and preserves this local binary structure while embedding the secret data in the cover picture.

4.1 Performance Analysis

Three separate kinds of analyses have been conducted to demonstrate the robustness of the suggested approach. By calculating the histogram, the visual quality of the stego picture has been examined. A quantitative metric like PSNR has also been employed to demonstrate that the suggested approach preserves the cover image's visual quality upon embedding. The simulation parameters used in this paper as given in table 1.

10.48047/jocaaa.2019.27.05.02

Table 1. Simulation Parameters

Parameter	Value
Population size	2* Number of features in data set (SPAM = 668, CC-Pev=600
Food source	Number of features in data set (SPAM = 668, CC-Pev = 600)
Feature Dimension (D)	85
Lower Bound	1
Upper Bound	N = Number of features in the data set

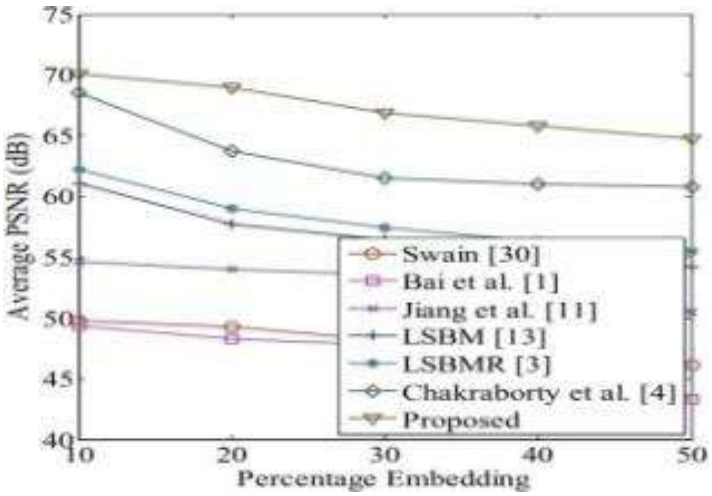

Figure 6. Average PSNR of several techniques at various rates of embedding

The histogram is the most basic statistical component of any image. Fig. 4 displays the original cover histogram as well as the associated stego image histogram with various embedding rates. Even with 50% embedding, the stego image's histogram resembles the cover image's histogram. This demonstrates how well the suggested technique maintains the cover image's statistical characteristics. This section examines the effectiveness of the suggested approach with several parameter configurations.

Figures 6 and 10 show the impact of threshold on the chosen subset of characteristics. It's evident that 250 is the ideal amount for CC-Pev and 80 for SPAM for this parameter. This number suggests that in order to get the highest level of fitness, we should use this subset of attributes. These figures allow us to demonstrate the ability of our suggested strategy to increase SPAM and CC-Pev accuracy.

The impact of various limit types in the suggested strategy is shown in Figure 7. It is clear that the ideal limit value to establish is 100.

10.48047/jocaaa.2019.27.05.02

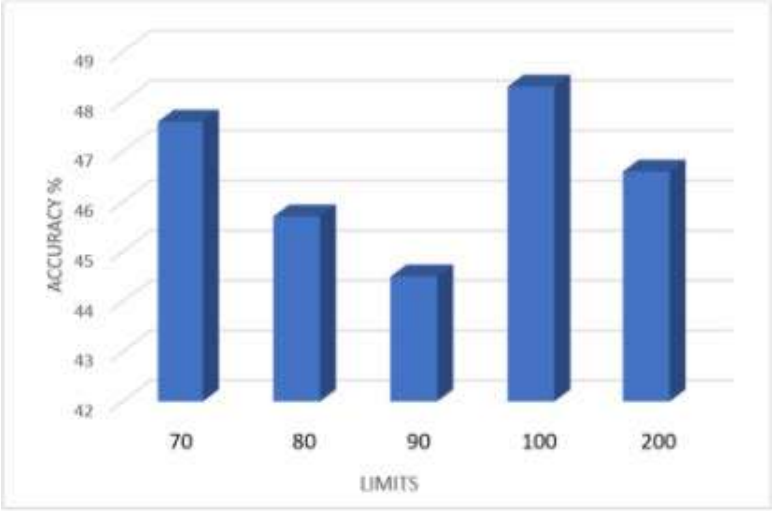

Figure 7. Impact of various limits on accuracy of SPAM

Figure 8. A comparison of the differential steganalysis method's classification accuracy

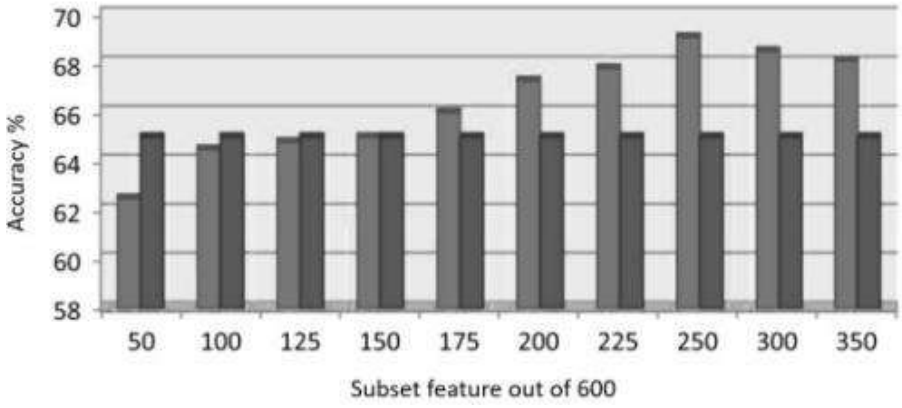

Figure 9. Comparing the many types of feature dimensions on CC-PEV

10.48047/jocaaa.2019.27.05.02

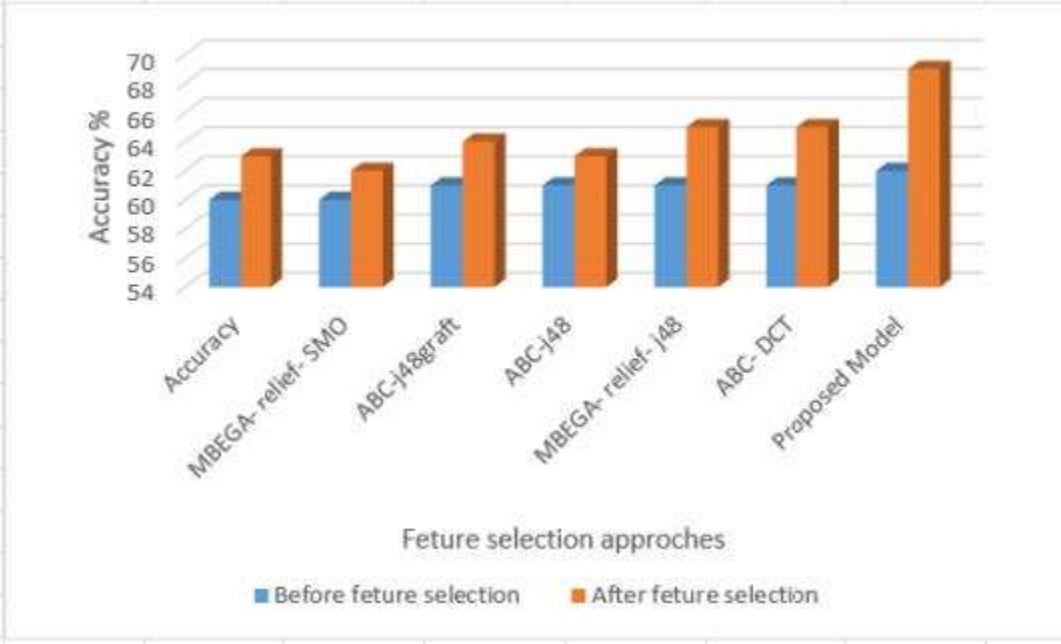

Figure 10. A comparison of the differential steganalysis method's classification

5. CONCLUSION

The study's primary goal is to offer a better technique for reducing data dimensionality. This work proposes a unique Image steganalysis LBP method based on KNN and ABC- LBP. We will use the LBP approach in the suggested method, which makes use of the wrapper strategy for categorization. The suggested method is a distinctive feature-based resilient steganography methodology that maintains the cover's local structure in the final stego picture. The suggested technique outperforms the majority of current state-of-the-art steganographic schemes, as demonstrated by a number of performance evaluation trials. The local structure of the cover is broken by state-of-the-art steganography, allowing for the powerful statistical feature-based steganalysis that detects concealed data. ABC-LBP proposed model performance assessment. Our suggested technique outperforms the other methods, as demonstrated by Releif-J48, SMO, and other types of ABC-DCTbased algorithms.
The trials demonstrate how well the suggested strategy maintains local structure and how resilient it is to feature- based steganalysis.

REFERENCES

[1] Abhilasha Ramdas Bhagat & Prof Ashish B Dhembhare, " An Efficient and Secure Data Hiding Technique – Steganography", International Journal of Innovative Research in Computer & Communication Engineering, Feb 2015.
[2] Amira S. Ashour, Sourav Samanta, Nilanjan Dey, Noreen Kausar, Wahiba Ben Abdessalemkaraa, Aboul Ella Hassanien, "Computed Tomography Image Enhancement Using

10.48047/jocaaa.2019.27.05.02

Cuckoo Search: A Log Transform Based Approach", Journal of Signal and Information Processing, vol 6, pp 244-257, Aug 2015.
[3] Venkata Vijaya Geeta, Pentapalli, Ravi Kiran Varma, "Cuckoo Search Optimization and its Applications: A Review", International Journal of Advanced Research in Computer and Communication Engineering, vol 5, issue 11, Nov 2016.
[4] Saurav Kaushik, " Introduction to Feature Selection Methods (how to select right variables)". In : Analytics Vidya retrieval on the web, Dec 2016.
[5] Hedieh Sajedi , "Image Steganalysis Using Artificial Bee Colony Algorithm", Journal of Experimental & Theoretical Artificial Intelligence, Jan 2017.
[6] Piyush Goel, "Data Hiding in Digital Images : A Steganographic Paradigm", proceedings of theses, IIT Kharagpur, May 2008.
[7] Meenpaa, T., Ojala, T., Pietikainen, M., Soriano, " Robust Texture Classification by Subsets of Local Binary Patterns. In: Proc. 15th International Conference on Pattern Recognition, vol. 3,pp. 947–950, 2010.
[8] Chen, C., & Shi, Y. Q. "JPEG Image Steganalysis Utilizing both Intrablock and Interblock Correlations. In IEEE International, Symposium on Circuits and Systems, ISCAS (pp. 3029–3032). Seattle, WA, 2008.
[9] Ratna Babu, K. and K.V.N. Sunitha, "Enhancing Digital Images Through Cuckoo Search Algorithm in Combination with Morphological Operation", Journal of Computer Science, July 2014.
[10] L ladha & T Deepa, "Feature Selection Method and Algorithms", International Journal on Computer Science & Engineering (IJCSE), May 2011.