International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 4 Issue 04, April-2015

# DWT Feature based Blind Image Steganalysis using Neural Network Classifier

Manisha Saini
Student, M.Tech (CSE)
Department of CSE/IT
ITM University, Gurgaon, Haryana, India

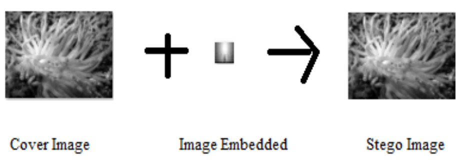
Rita Chhikara
Assistant Professor
Department of CSE/IT
ITM University, Gurgaon, Haryana, India

***Abstract*-The objective of forensic steganalysis is to identify the existence of embedded message and to ultimately retrieve the secret message. In this paper, we have extracted Histogram, Markov and Co-occurrence features from wavelet domain and compared with existing farid 72 DWT features. The performance metrics are MSE (Mean square error) and classification accuracy for blind image steganalysis using two steganography algorithms outguess and nsF5. Neural network back propagation classifier has been used to classify images into stego images and clean images. Experimental results show that the proposed features considerably outperform existing farid 72 DWT features.**

***Keywords—Discrete wavelet transformation, Neural network, Outguess, Steganography, Steganalysis, nsF5.***

## I. INTRODUCTION

Steganography[1] is a process in which the message is embedded in a cover medium to create a stego medium with the help of steganography tool that results in changes in the statistical properties of cover medium where medium could be video, text, image, audio etc. Now a day's wide variety of steganographic tools are freely and widely available on the internet such as StegHide [2], Outguess [3] etc. Statistical un-detectability [4] is the necessity of steganography process, which means that it is difficult or impossible for attacker to judge whether an image is the stego or cover based on the statistics.




Steganography hides data in [5] Spatial domain and Transform domain or Frequency domain. Transform Domain is further divided into three categories (a) DCT (Discrete Cosine Transformation technique), (b) DWT (Discrete Wavelet transformation technique), (c) DFT (Discrete Fourier transformation technique).In transform domain, images are initially transformed and then the message is embedded while in spatial domain message is embedded directly inside the pixel.

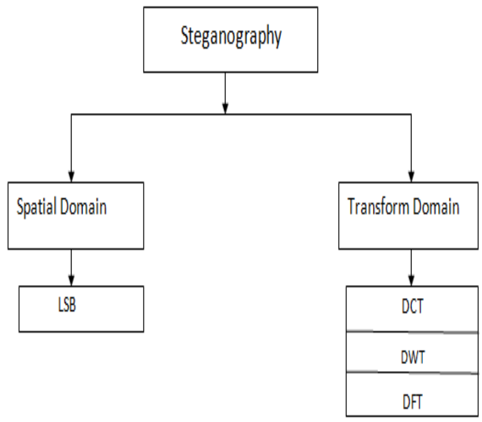



Fig. 1 Data Hiding Methods

Reverse of steganography is steganalysis [6]. The goal of steganalysis is recovering the embedded message from cover medium.Current forensics explore has resulted in a number of steganalysis detection techniques.Steganlysis process can be divided into the two categories [6] (a) Specific steganalyis (b) Blind steganalayis. Specific steganalyis is used to target a particular steganographic algorithm whereas Blind/Universal steganlysis is one in which particular steganography algorithm is not known prior.

The rest of this paper is divided in the following sections. Section 2 discusses about two steganography algorithm Outguess and nsF5. Section 3 describes the general framework for steganalysis using neural networks. In section 4 proposed feature extraction method is explained. Section 5 discusses in detail about neural network classifier and its importance in steganalyis process. In section 6 and 7 Experimental results and analysis is shown in detail. Finally the paper is concluded in section 8.

## II. STEGANOGRAPHY ALGORITHMS

### *A. Outguess*

Outguess [3] is a universal steganographic tool that allows the addition of hidden information into the redundant bits of data sources.Outguess was designed by Niels provos. Outguess is an advanced variant of Jsteg. It uses pseudo-random number generator (PRNG) based scattering to complicate steganalysis.

The embedding and extraction functions therefore each require a seed as additional parameter to initialize the PRNG. Outguess algorithm is as shown below [7]

```
initialize PRNG with shared secret
while data left to embed do
        get pseudo-random DCT coefficient from cover
        image
        if DCT ≠ 0 and DCT ≠ 1 then
        get next LSB from message
        replace DCT LSB with message LSB
        end if
        insert DCT into stego image
end while
```

### *B. nsF5*

nsF5(no shrinkage F5) is enhanced version of F5,proposed in 2007.This algorithm was developed to improve the problem which exist in F5 algorithm.F5 is popular steganography algorithm and is undetectable using chi-square test.F5 uses matrix encoding and permutation straddling to hide the data. On the other hand shrinkage problem exists in F5 algorithm which results in low embedding efficiency. In 2007 nsF5 was introduced to overcome the negative effect of shrinkage by combining F5 algorithm with wet paper codes (WPC). Theoretical bound for the embedding efficiency is [8]-

$$e = \frac{\alpha}{H^{-1}(\alpha)}$$

Where $H^{-1}$ is inverse of entropy and alpha is the number of non zero AC DCT coefficients exist in the image.

## III. GENERAL FRAMEWORK

Steps of Proposed work are:-

(A) An image dataset containing of both cover and stego images is obtained by using two steganography algorithms outguess and nsF5 on cover images.

(B) The image dataset is divided into training and testing set.

(C) Wavelet image decomposition is performed, where each image from dataset is divided into 12 subbands

(D) The following set of features are extracted- (i) proposed Histogram 130 feature set (ii) proposed Co-occurrence 52 feature set (iii) Markov 81 feature set (4) existing farid 72 DTW feature set from image & subands obtained after image decomposition.

(E) In Training phase, features extracted from training image dataset are fed to Neural Network Back propagation classifier to train the classifier and finally trained classifier will make the decision whether image is clean or stego image.

(F) In testing phase, features are extracted from test image dataset similar to the way as done in training phase. These are fed to trained neural network back propagation classifier and classifier based on the training decides final classification decision

(F) Detection results for both steganography algorithm Outguess and nsF5 is compared and analyzed in terms of classification accuracy and MSE (mean square error) for all sets of features extracted.

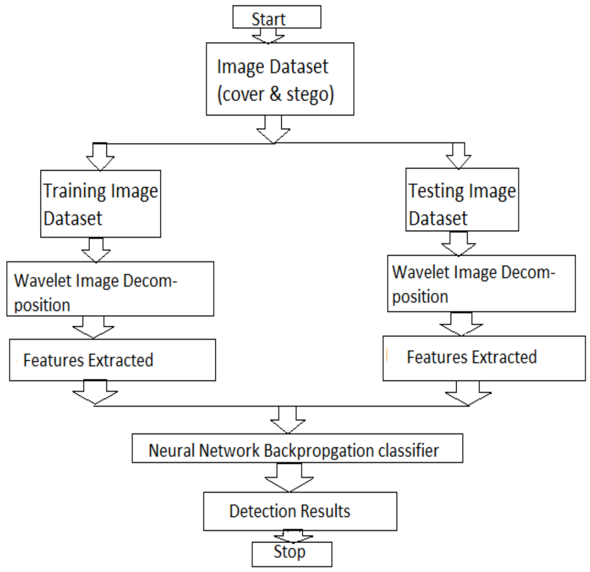



Fig 2 General Framework

## IV. PROPOSED DWT FEATURE EXTRACTION

On applying Discrete wavelet transformation [9] on image we obtain 4 different subbands at level 1 (a)LL1 (b)HL1 (c)LH1 (d) HH1 as shown in fig2, where LL1 subband is low frequency component containing maximum information also known as approximation coefficient and other three subbands (HL1,LH1,HH1) are high frequency component containing least information also known as detailed coefficients. Maximum embedding takes place in LH and HL subbands. Main objective of DWT is space localization, it can forecast at which space we have low frequency component and high frequency component.

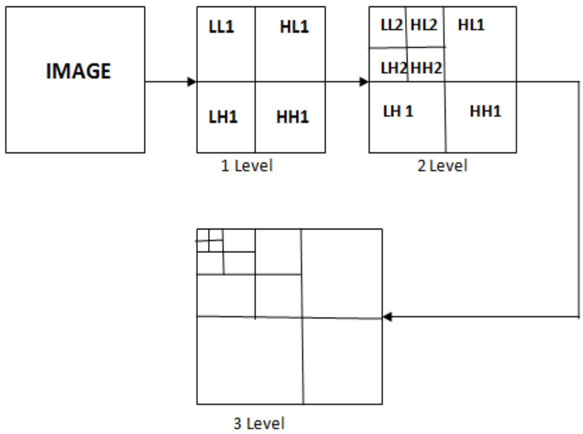



Fig 3 Three-level wavelet decomposition

We have divided each image from image dataset into 4 subbands at each level using Haar wavelet decomposition (up to 3 level decomposition is done) and obtained 12 subbands as shown below.

cA1,cH1,cV1,cD1 –Level 1
cA2,cH2,cV2,cD2 –Level 2
cA3,cH3,cV3,cD3 –Level 3

All the features are extracted from the complete image and 12 subbands obtained after decomposition of the image with DWT. The various features extracted are as given in sections below.

*A. Histogram*

Histogram [10] is a approach of summarizing data that are calculated on an interval scale (either discrete or continuous). It partitions the range of feasible values in a dataset into groups. It is used in dealing with large dataset and helps us to detect outlier values in the dataset.

The 130 features of histogram are extracted by binning the elements of all subbands into 10 evenly spaced containers and returning the number of elements in each container.The values are thereafter normalized by dividng each element by sum of all elements.

$$\mathrm{H\_0} = \mathrm{H\_0}/\mathrm{sum}(\mathrm{H\_0})$$

We have extracted 130 features calculated from 1 Image+12 subands.

*B.Markov*

Markov features show the difference between absolute values of neighboring coefficients.These features have been extracted in DCT domain by [12] but we have extracted Markov features in DWT domain. Calculated four difference array along horizontal, vertical and diagonal direction using formula:-

$$\mathrm{F_h}(u,v) = \mathrm{F}(u,v) - \mathrm{F}(u+1,v)$$

$$\mathrm{F_v}(u,v) = \mathrm{F}(u,v) - \mathrm{F}(u,v+1)$$

$$\mathrm{F_d}(u,v) = \mathrm{F}(u,v) - \mathrm{F}(u+1,v+1)$$

$$\mathrm{F_m}(u,v) = \mathrm{F}(u+1,v) - \mathrm{F}(u,v+1)$$

From difference array calculated 4 TPM (Transition Probability matrix) in horizontal, vertical, diagonal direction using formula

$$M_h(i,j) = \frac{\sum_{u=1}^{S_u-2}\sum_{v=1}^{S_v}\delta\left(F_h(u,v)=i, F_h(u+1,v)=j\right)}{\sum_{u=1}^{S_u-1}\sum_{v=1}^{S_v}\delta\left(F_h(u,v)=i\right)}$$

$$M_v(i,j) = \frac{\sum_{u=1}^{S_u}\sum_{v=1}^{S_{v-2}}\delta\left(F_v(u,v)=i, F_v(u,v+1)=j\right)}{\sum_{u=1}^{S_u}\sum_{v=1}^{sv-1}\delta\left(F_v(u,v)=i\right)}$$

$$M_d(i,j) = \frac{\sum_{u=1}^{S_u-2}\sum_{v=1}^{S_v-2}\delta\left(F_d(u,v)=i, F_d(u+1,v+1)=j\right)}{\sum_{u=1}^{S_u-1}\sum_{v=1}^{sv-1}\delta\left(F_d(u,v)=i\right)}$$

$$M_m(i,j) = \frac{\sum_{u=1}^{S_u-2}\sum_{v=1}^{S_v-2}\delta\left(F_m(u+1,v)=i, F_m(u,v+1)=j\right)}{\sum_{u=1}^{S_u-1}\sum_{v=1}^{sv-1}\delta\left(F_m(u,v)=i\right)}$$

We have extracted features from HL and HL subbands and averaged the four TPM (Transition Probability matrix) to obtain 81 features. Where threshold value applied is -4 to 4.

*C. Co-occurrences*

Co-occurrence matrix[13] depict the correlation between wavelet coefficients. We have extracted contrast, correlation, energy and homogeneity [18] features from Co-occurrence matrix from 1Image+12 subbands of DWT of an image obtaining 52 features.

**(i) Contrast:** it is a measure of the amount of local variations present in an image. It returns an amount of the intensity contrast .

**(ii) Energy**-returns sum of square elements in the Co-occurrence matrix.

**(iii) Correlation**-Correlation measures the linear dependency of grey levels of neighboring pixels.

**(iv)Homogeneity**:-Returns a value that measures then closeness of the distribution of elements in the Co-occurrence matrix to the Co-occurrence matrix diagonal.

TABLE 1 .Contrast, Energy, correlation and Homogeneity Formula

| | |
|---|---|
| Contrast | $\sum_{i,j=0}^{N-1} \mathrm{P_{i,j}}\ (i-j)^2$ |
| Energy | $\sum_{i,j=0}^{N-1} \mathrm{P_{i,j}^2}$ |
| Correlation | $\sum_{i,j=0}^{N-1} \frac{(i-\mu i)(j-\mu j)P(i,j)}{\sigma_i\sigma_j}$ |
| Homogeneity | $\sum_{i,j=0}^{N-1} \frac{P_{i,j}}{1+(i-j)^2}$ |

where i and j are rows and columns of Co-occurrence matrix and range 0 to N-1 is defined.

*D. Existing Farid features*

Farid proposed 72 feature set [14] [15], 4 features that are extracted from subands are mean, variance, skewness and kurtosis as shown below [16].

$$E(x) = \frac{1}{\mathrm{n}}\sum_{k=1}^{n} \mathrm{x_k}$$

$$Var(x) = \frac{1}{\mathrm{n}-1}\sum_{i=1}^{n} (\mathrm{x_i} - \mathrm{E(x)})^2$$

$$S(x) = \mathrm{E}\left[\left(\frac{\mathrm{x} - \mathrm{E(x)}}{\sqrt{(\mathrm{Var(x)})}}\right)^3\right]$$

$$K(x) = \mathrm{E}[(\frac{\mathrm{x} - \mathrm{E(x)}}{\sqrt{(\mathrm{Var(x)})}})^4]$$

The image decomposition used in farid paper, is based on separable quadrature mirror filters (QMFs) contain vertical, horizontal and diagonal subbands.

## V. NEURAL NETWORK CLASSIFIER

Another key element in steganalysis is selection of a good classifier so as to maximize the classification rate and increase the accuracy and minimize the error rate. Various classifiers are present, neural network is one of the classifier. Some of the advantages of neural networks [17] are adaptive learning, self organisation, fault tolerance, real time operation, high tolerances to noisy data, ability to replicate any nonlinear relation,ability to generalize and fast learning that makes neural network classifiers as a proficient approach to categorize or classify patterns.

We have used Neural Network back propagation classifier to classify images into stego images and cover images [18].Neural network back propogation classifier has three layers (1) Input layer (2) Hidden layer (3) Output layer. Features extracted are given to input layer. In this experiment we have considered cases for both 20 hidden neurons and 30 hidden neurons. Range of output layer neurons is (0,1).Neural network will help us to select input data, create and train a network, and evaluate its performance via mean square error, and percentage error (classification accuracy=100-percentage error).

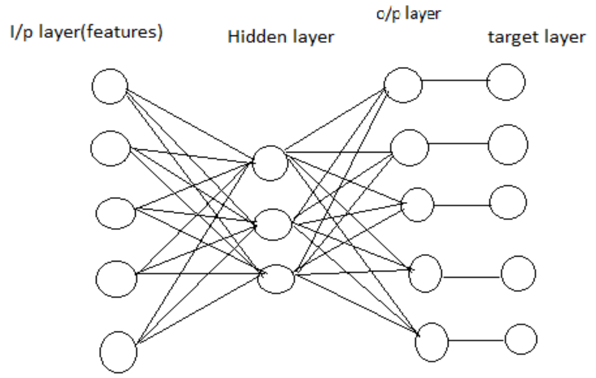



Fig 4: BP Neural Network

## VI. EXPERIMENTAL RESULTS

We have created image dataset consisting of 2000 cover images downloaded from various sites and photographs. Two steganography algorithm (1) Outguess (2) nsF5 have been applied to obtain two set of 2,000 stego images.Various embedding capacities have been used to generate the dataset as shown in TABLE 2. So total 4,000 images (cover plus stego) have been generated in each case. These 4000 images have been divided into training (2,800 samples), Validation (600 samples) and testing (600 samples) sets.The various embedding capacities are as shown in table 2.

TABLE 2 .embedding capacities

| Outguess | nsF5 |
|---|---|
| 16X 16 | 10 % |
| 32X32 | 25 % |
| 48X48 | 50% |

After dividing the dataset into training, validation and testing, features are extracted from each image present in image dataset. This is then fed to Neural network back propagation classifier.Performance measures accuracy and MSE (mean square error) is calculated, Percent Error indicates the portion of samples which are misclassified and accuracy is 100-Percent Error and Mean Squared Error (MSE) is the average squared difference between outputs and targets.

## VII. ANALYSIS

TABLE 3 and TABLE 4, shows the comparisons between proposed features and existing Farid 72 features on the basis of parameter MSE (Mean square error) using 20 hidden neurons and 30 hidden neurons respectively in outguess steganography algorithm. From the experiment results of TABLE 3 and TABLE 4, we found that Histogram 130 features gives less mean square error in comparison to Markov 81 features, Co-occurrence 52 features and existing Farid 72 features. When we increase the number of hidden neurons to 30, we didn't notice any significant change in the results.

TABLE 5 and TABLE 6 shows the comparison between proposed features and existing Farid 72 features on the basis of parameter MSE(mean square error) using 20 hidden neurons and 30 hidden neurons respectively in nsF5 steganography algorithm. From the experiment results of TABLE 5 and TABLE 6, we found that Markov 81 features gives less mean square error in comparison to Histogram 130 features,Co-occurrence 52 features and existing Farid 72 features and also when we increase the number of Hidden neurons from 20 to 30, didn't show any considerable change in the results.

As shown in Fig 5 we conclude that Histogram 130 features give best accuracy in comparison to Markov 81 features and Co-occurrence 52 features and existing Farid 72 features and even Markov 81 features give equivalent performance as existing Farid 72 features in outguess steganography algorithm.

From Fig 6 we conclude that Markov 81 features give best accuracy in comparison to Histogram 130 features and Co-occurrence 52 features and existing Farid 72 features in nsF5 steganography algorithm.

This indicates that outguess algorithm is more sensitive to histogram statistical features and nsF5 can be detected with markov statistical features.

## VIII. CONCLUSION

This paper highlights the process known as steganography and steganalyis technique. In steganalyis process feature extraction and classifier plays an important role. We have created image dataset consisting of both cover and stego images, where stego images are obtained by applying two stegonagraphy algorithms Outguess and nsF5. Then we have extracted Features from wavelet domain (i) Histogram 130 feature set,(ii) Markov 81 feature set, (iii)Co-occurrence 52 feature set and compared it with existing Farid 72 DWT features . These features extracted are fed to neural network back propagation classifier for classification. The performance is analyzed on the basis of parameters classification accuracy and mean square error (MSE).Experimental results show that (i) proposed histogram features give best performance in Outguess steganography algorithm and (ii) Markov features give best performance in nsF5 steganography algorithm. Future work will be to use other steganography algorithm and analyze the performance

TABLE 3. Experimental Result with 20 Hidden neurons in Outguess

| Outguess | | | | |
|---|---|---|---|---|
| 20 Hidden Neurons | | | | |
| | Co-occurrence (52 features) | Markov(81 features) | Histogram(130 features) | Farid(72 features) |
| | MSE | MSE | MSE | MSE |
| Training | 2.02877 $e^{-1}$ | 1.81086 $e^{-1}$ | 1.29693 $e^{-2}$ | 1.99075 $e^{-1}$ |
| Validation | 2.03 527 $e^{-1}$ | 1.97902 $e^{-1}$ | 1.28365 $e^{-2}$ | 1.97495 $e^{-1}$ |
| Testing | 1.99548 $e^{-1}$ | 1.99411 $e^{-1}$ | 2.16521 $e^{-2}$ | 2.02965 $e^{-1}$ |

TABLE 4. Experimental Result with 30 Hidden neurons in Outguess

| Outguess | | | | |
|---|---|---|---|---|
| 30 Hidden Neurons | | | | |
| | Co-occurrence (52 features) | Markov(81 features) | Histogram(130 features) | Farid(72 features) |
| | MSE | MSE | MSE | MSE |
| Training | 2.0089 $e^{-1}$ | 1.933379 $e^{-1}$ | 1.72811 $e^{-2}$ | 1.94656 $e^{-1}$ |
| Validation | 2.01432 $e^{-1}$ | 2.14 229 $e^{-1}$ | 2.75580 $e^{-2}$ | 2.06669 $e^{-1}$ |
| Testing | 2.0254 $e^{-1}$ | 2.19434 $e^{-1}$ | 2.31695 $e^{-2}$ | 2.022833 $e^{-1}$ |

TABLE 5. Experimental Result with 20 Hidden neurons in nsF5

| | | | | |
|---|---|---|---|---|
| 20 Hidden Neurons | | | | |
| | Co-occurrence (52 features) | Markov(81 features) | Histogram(130 features) | Farid(72 features) |
| | MSE | MSE | MSE | MSE |
| Training | 2.49333 $e^{-1}$ | 5.68279 $e^{-2}$ | 2.49820 $e^{-1}$ | 2.25286 $e^{-1}$ |
| Validation | 2.4499 $e^{-1}$ | 6.08998 $e^{-2}$ | 2.49956 $e^{-1}$ | 2.30938 $e^{-1}$ |
| Testing | 2.50006 $e^{-1}$ | 6.60 275 $e^{-2}$ | 2.51608 $e^{-1}$ | 2.35782 $e^{-1}$ |

TABLE 6. Experimental Result with 30 Hidden neurons in nsF5

| nsF5 | | | | |
|---|---|---|---|---|
| 30 Hidden Neurons | | | | |
| | Co-occurrence (52 features) | Markov(81 features) | Histogram(130 features) | Farid(72 features) |
| | MSE | MSE | MSE | MSE |
| Training | 2.4633 $e^{-1}$ | 5.079418$e^{-2}$ | 2.49046 $e^{-1}$ | 2.27077 $e^{-1}$ |
| Validation | 2.4625 $e^{-1}$ | 5.69214 $e^{-2}$ | 2.51257 $e^{-1}$ | 2.26128 $e^{-1}$ |
| Testing | 2.4728 $e^{-1}$ | 6.17991 $e^{-2}$ | 2.50756 $e^{-1}$ | 2.42707 $e^{-1}$ |

Fig 5. Accuracy in % of proposed features with existing features for Outguess

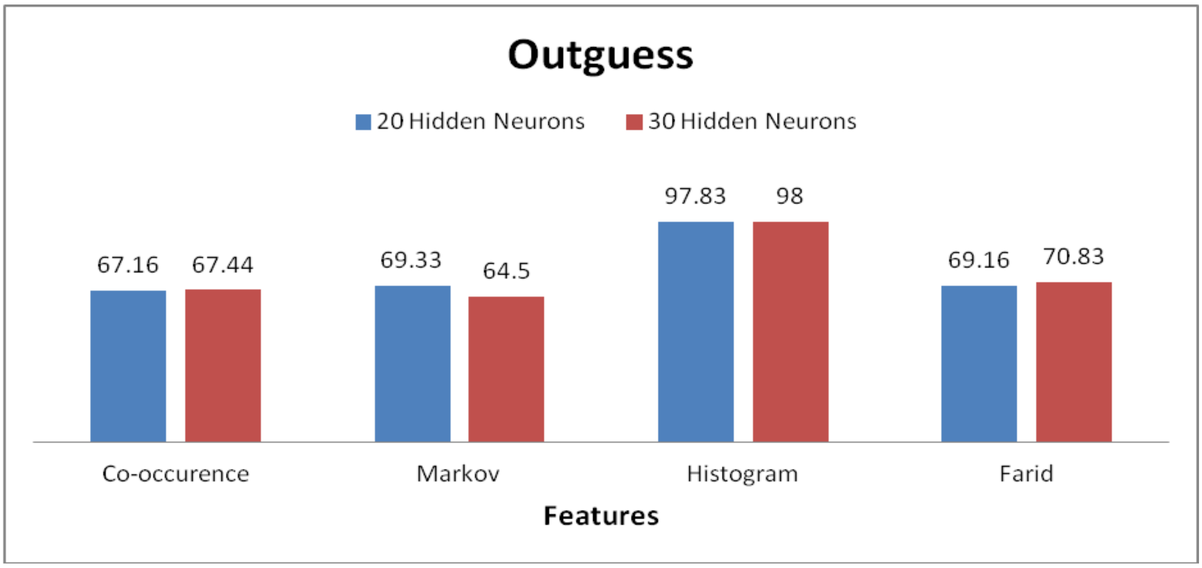
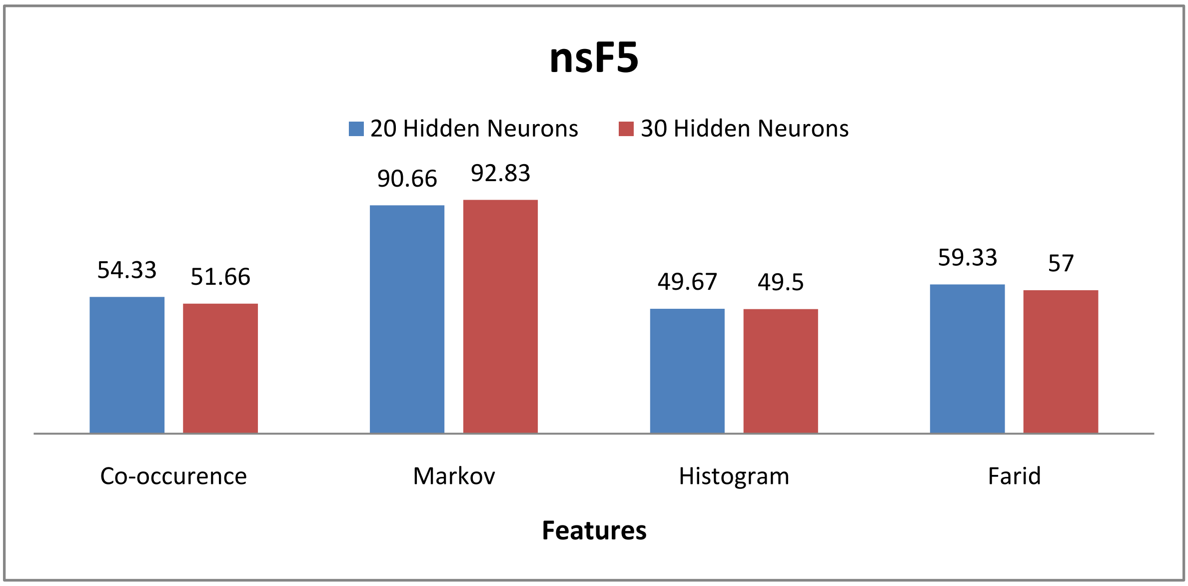



Fig 6. Accuracy in % of proposed features with existing features for nsF5



## REFERENCES

[1] Abbas Cheddad, Joan Condell, Kevin Curran, Paul Mc Kevitt,“Digital image steganography: Survey and analysis of current methods”, Signal Processing90 (2010)727–752.

[2] http://steghide.sourceforge.net/

[3] Niels provos, www.outguess.org

[4] Yong Wang, Jiufen Liu, Weiming Zhang and Shiguo Lian,“ Reliable JPEG steganalysis based on multi- directional correlations”**,** Signal Processing: Image Communication 25 (2010)577–587.

[5] Preeti Parashar and Rajeev Kumar Singh,“ A Survey: Digital Image Watermarking Techniques”, International Journal of Signal Processing, Image Processing and Pattern Recognition, Vol. 7, No. 6 (2014), pp. 111-124

[6] Arooj Nissar,A.H. Mir, “Classification of steganalysis techniques: Astudy”, Digital Signal Processing 20 (2010) 1758–1770.

[7] Jessica Fridrich, Miroslav Goljan and Dorin Hogea,“Attacking the OutGuess”, in Proc. of the ACM Workshop on Multimedia and Security 2002, France, December 6, 2002

[8] Jessica Fridrich, Tomáš Pevný and Jan Kodovský,“Statistically Undetectable JPEG Steganography: Dead Ends, Challenges, and Opportunities”, in Proc. of the ACM workshop on Multimedia & security, September 20-21,2007,pp. 3-14.

[9] Shaker K. Ali , Zou Beijie,“Analysis and Classification of Remote Sensing by using Wavelet Transform and Neural Network”, IEEE 2008 International Conference on Computer Science and Software Engineering, 12-14 Dec. 2008,pp.963–966, doi:10.1109/CSSE.2008.464

[10] Shimazaki, H.; Shinomoto, S. , "A method for selecting the bin size of a time histogram". Neural Computation 19 (6): 1503–1527, 2007.

[11] Archana Deshlahra,G.S.Shirnewar ,Dr. A.K. Sahoo,“A Comparative Study of DCT, DWT & Hybrid (DCT-DWT) Transform”, International Conference on Emerging Trends in Computer and Image Processing(ICETCIP), 24 February, 2013, Bangalore

[12] Tom´aˇs Pevn´y, Jessica Fridrich, “Merging Markov and DCT Features for Multi-Class JPEG Steganalysis”,The International Society for Optical Engineering, volume 6505, pp.28-40, 2007.

[13]Ahd Aljarf, Saad Amin, John Filippas, and James Shuttelworth, “Develop a Detection System for Grey and ColourStego Images”, International Journal of Modeling and Optimization, Vol. 3, No. 5, October 2013

[14] Hany Farid, “Detecting Hidden Messages Using Higher-order Statistical Models”, Proc. IEEE Symp. Int'l Conf. on Image Processing (ICIP 2000), IEEE Press, Sep.2002,pp.905-908,doi:10.1109/ICIP.2002 1040098.

[15] S. Lyu, and H. Farid, “Steganalysis Using Higher-order Image Statistics”,Trans. Information Forensics and Security,vol.1,Jan.2006,pp.111-119,doi:10.1.1.116.1418.

[16] Yuan Liu, Li Huang, Ping Wang, Guodong Wang, “A Blind Image Steganalysis Based on Features from Three Domain”, Proc. IEEE Int'l Conf. on Control and Decision,2008,pp.2933-2936.

[17] Yun Q. Shi, Guorong Xuan, Dekun Zou, Jianjiong Gao,Chengyun Yang, Zhenping Zhang, Peiqi Chai, Wen Chen and Chunhua Chen,“Image Steganalysis Based on Moments of Characteristic Functions Using Wavelet Decomposition, Prediction-Error Image, and Neural Network” in Proc. IEEE International Conference on Multimedia and Expo (ICME05), Amsterdam, Netherlands, July, 2005.

[18] H.B.Kekre, A.A.Athawale & S.A.Patki,“ Steganalysis of LSB Embedded Images Using Gray Level Co-Occurrence Matrix”, International Journal of Image Processing (IJIP), Volume (5) : Issue (1) 36-45, 2011.

[19] M. Egmont-Petersen, D. de Ridder, H. Handels. "Image processing with neural networks - a review," Pattern Recognition, Vol. 35, No. 10, pp. 2279-2301, 2002.