

## A SYSTEMATIZED APPROACH TO DATA HIDING USING IMPROVED LSB

Ganesh R

NMSSVN College, Madurai-625009, Tamilnadu, India

Thabasu Kannan S

Pannai College of Engg& Tech, Sivagangai – 630 561, Tamilnadu, India

Email: thabasukannan@gmail.com

Received: 10<sup>th</sup> October 2022, Accepted: 3<sup>rd</sup> November 2022 and Published: 3<sup>rd</sup> November 2022

### ABSTRACT

**Aim:** This paper mainly focused on introducing a new framework for image and text hiding using a simple Least Significant Bit (LSB) substitution method.

**Results:** Six 24-bit images are used as reference images. XOR operation is used in the stego key to generate new bit planes of the stego images. An RGB color image is used as a secret cover image. The proposed method dramatically increases the embedding capacity without significantly decreasing the Peak Signal-to-Noise Ratio (PSNR) value.

**Conclusion:** This method increased the compression ratio, embedding ratio, PSNR and Space 24%, 53%, 36% and 47%, respectively. The result shows that the newly developed method improved the value of accuracy by 30% and sensitivity by 16% for the hidden images.

**Keywords:** LSB, PSNR, FNR (False Negative Rate), Sensitivity, MSE (Mean Square Error), Specificity

### HIGHLIGHTS:

1. The newly developed method with simple Least Significant Bit (LSB) improved the accuracy value by 30%, and sensitivity by 16% for the hidden image.
2. The improved LSB method increased security and preserved the quality of hidden images.

### INTRODUCTION

A simple method adopted in this study is to hide data directly in the LSB of each pixel in an image. Then based on the LSB, an algorithm for a 24-bit colour image is developed, thus improving the hidden image's quality. This method has broad applications to maintain digital copyrights, protect information and conceal secrets.

Image hiding deals with the concealing of an image on an image. This section's main intention is to formulate a new framework called Triple framework for Image hiding, which is used to optimize the level of quality

The procedure for the triple framework is as follows:

```
{
Input :      24 bit color image (Cover) and 8 bit color image (Secret)
Output :     Stego Image
{
Segregate the cover image into RGB Panels
{
1MSB = R * 254; // Last one LSB to zero 2&3MSB = G*252; // Last two LSB to zero
4, 5 & 6MSB = B*248; // Last three LSB to zero
}
Stego_image=1MSB+2&3MSB+ 5&6MSB;
}
}
```

### RESULTS

The above framework improves the percentage of Robustness, imperceptivity and hiding capacity. The performance of the above triple framework for image hiding is evaluated by using four measurements viz compression ratio, embedding ratio, Space save and PSNR.

Table 1 summarizes the improvement of the newly developed framework by considering the above four measurements compared with the existing system of LSB and MSB. We have investigated the above-said measurements individually based on the maximum hike. Our new Triple Framework improved CR by 26.29% when compared with LSB and 20.68% when compared with MSB and for ER, the framework attained a hike of 73.69% when compared with LSB and 64.39% when compared with MSB. For PSNR, there is an increase of 36.28% compared with LSB and 27.35% compared to MSB. The Space save also reached an increase of 64.44% when compared with LSB and 55.49% when compared with MSB

Table 2 shows the average results of various measurements compared with variant and invariant types. We have found that the compression ratio for the variant case is increased by 21.5% compared to the invariant case. The embedding ratio for the variant case is increased by 30.93% compared to the invariant case. Furthermore, the Space Save for an invariant case is increased by 48.39% compared to a variant case. The MSE for a variant case is optimized by 73.33% compared to an invariant case. The PSNR for a variant case is increased by 3.27% as compared with an invariant case. Hence, the Variant case's efficiency is superior to the invariant case for the measurements of CR, ER, MSE and PSNR. However, the SS for a variant case is inferior to the Invariant case.

The text Hiding method nowadays plays an essential role in providing security. It is defined as the concealment of text on an image. We have developed a new framework, which hides text using the simple LSB method. Here we have noticed some of its demerits. The level of security is found to be low. The level of distortion is HIGH for an even smaller ratio. The level of imperceptibility is also not advisable. We have proposed a framework for text hiding to overcome the above drawbacks.

#### Procedure for text hiding using the triple framework

```
{
Input: Cover image, Text File and Stego Key
Output : Stego_Image
{
A.    Read the characters from the Text_file one by one until EOF
B.    Convert to ASCII and again Convert to its equivalent binary value and stored
in an array
C.    XORing the Stego_key and used to store in the above array and sent to the
last LSB of the Red Pixel
D.    Generate random key and Permute then resize the above array to get the
stego_image.
}
```

The above framework provides tight security for the Stego images because it provides 3-tier security for our data. The 1-tier is in XORing the Stego\_key, the 2-tier during the generation of Random\_key and the 3-tier provided while the permutation. The above framework's efficiency has been evaluated using some measurements, namely CR, ER, SS, MSE and PSNR. Table 3 clearly shows that the newly developed framework for text hiding proves to be superior compared to the existing system of the Simple LSB method for text hiding. Here our new framework has attained an improvement of 45% in Compression Ratio, 16% in PSNR, 18% in Embedding Ratio, 16% in Space Save and got 100% optimum in MSE.

The study focuses on improving the quality in terms of accuracy for the image-hiding process. To improve the quality, we have used the latest measurements viz Precision, Recall, Accuracy, PSNR, Prevalence, F1-Score, FPR, Balanced Accuracy and Specificity. We have utilised the above nine statistical measurements, which can be used to improve the present level of accuracy in a refined manner. The above measurements can be calculated from the confusion matrix derived from the visibility-based image hiding outcome. Here all the above measurements have been compared with the performance of the visibility-based and confusion matrix-based methods. The results are summarised in Table 4.

Here the accuracy value of the existing method is very close to the value of 0.56 compared to the new method, i.e. 0.71. Hence the deviation closeness is 0.64. The precision of the existing method is very close to the value of 0.56 compared with the new method's results of 0.65. The new method provides higher precision and high accuracy.

## CONCLUSION

This paper proposes effective image-hiding and text-hiding methods, resulting in a higher embedding ratio, compression ratio, higher Space save, or less MSE. A generalised framework increases the quality and reduces the error rate. In addition to the above, the third problem is quality enhancements in accuracy, sensitivity and specificity. In this paper, we used the LSB and Random key generation Techniques to get a secure hidden image. Our result shows that the PSNR of our new method is higher than the simple method. The image quality does not change while hiding the data in the image, and the image is protected with a random key generation [1]. The same approach could be extended to the cloud concepts to improve security [2].

## REFERENCES

1. <https://dl.acm.org/doi/abs/10.5555/1378408.1378425>
2. <https://doi.org/10.51611/iars.irj.v5i1.2015.44>

## TABLES & FIGURES:

Table 1: Comparison of Various measurements with the improvements

Measure	On Max basis		On Individual basis		Average for	
	Hike %-L	Hike %-M	Hike %-L	Hike %-M	LSB	MSB
CR	25.5	20.13	27.08	21.23	26.29	20.68
PSNR	36.28	25.55	36.28	29.14	36.28	27.35
ER	52.63	42.11	94.74	86.67	73.69	64.39
SS	47.06	38.24	81.82	72.73	64.44	55.49

Table 2: Comparison of Variant and Invariant with maximum measurements

Measure	Invariant	Variant	Percentage Deviation
CR	2	2.43	21.50
ER	2.36	3.09	30.93
SS	62	32	48.39
MSE	0.15	0.04	73.33
PSNR	61.53	63.54	3.27

Table 3: Comparison of the performance of test hiding

Method	Max CR	Max PSNR	Max ER	Max SS	Min MSE
Simple LSB	1.23	75	0.28	45	0.02
Modified LSB	1.79	87	0.33	52	0
Hike %	<b>45.53</b>	<b>16.00</b>	<b>17.86</b>	<b>15.56</b>	<b>100.00</b>
On individual basis	<b>31.36</b>	<b>21.15</b>	<b>27.27</b>	<b>22.22</b>	<b>100</b>
Range	<b>30-45</b>	<b>15-20</b>	<b>18-27</b>	<b>16-22</b>	<b>100</b>

Table 4: Comparison of various statistical measurements

Method	Accu	Prec	Recall	F1-Score	Prevalence	Bal Acc	Specificity	FPR	PSNR
Vicinity	0.61	0.68	0.92	0.68	0.37	0.74	0.87	0.55	43.19
New	0.78	0.82	1.00	0.84	0.45	0.83	0.95	0.67	52.45
Hike %	<b>27.87</b>	<b>20.59</b>	<b>8.70</b>	<b>23.53</b>	<b>21.62</b>	<b>12.16</b>	<b>9.20</b>	<b>21.82</b>	<b>21.44</b>
individual	30.95	20.59	22.3	23.94	25.35	20.68	8.81	23.63	25.76
Range	28-31	21	9-22	24	22-25	12-21	8.8 - 9.2	22-24	22-26
Average	<b>29.5</b>	<b>21</b>	<b>15.5</b>	<b>24</b>	<b>23.5</b>	<b>17</b>	<b>9</b>	<b>23</b>	<b>24</b>

## SUPPLEMENTARY DATA

### Introduction and Discussion

Data act as a main role for now-a-days effective communication. The information hiding method guarantees for sending the information to the intended destination in secured manner. In digital era, there are various methods available for concealing information in any medium for effective communication. Here digital images are used as a medium for hiding and the information in the form of text, digital image, video or audio file may be used as secret message. This information hiding method can be used for a number of applications like military and intelligence agencies which needs secret communications. For 24 bit color image each pixel has three color components: Red, Green and Blue (RGB). Each pixel is represented with three bytes to indicate the intensity of these three colors (RGB). Previously the hiding methods are used for still images only. A simple method is directly hiding secret data into the LSB of each pixel in an image. Then based on the LSB, an algorithm for 24 bit color image is developed, which improves the quality of hidden-image. It is used to maintain digital copyrights, protect information and conceal secrets. The information hiding mainly depends on three facts i.e. capacity, security and robustness. Capacity means the media on which the data is to be hidden, so that the complexity of the medium should not be disturbed. Security means the hiding method is secure if the hidden information cannot be removed beyond reliable detection by targeted attacks. Finally, robustness means the amount of manipulation a cover image can handle without drawing any attention that a change has taken place. Here, Dove image and King Image are classified as perfect recall. Hence the recall value for Dove and King is absolutely 1.00. Every result retrieved was relevant. The existing method value is closed to 0.66 and new method is closed to 0.75, hence the difference is just 0.71. It is near to both the method. Here the classes of Dove, King, Polarbear and Logo return high recall for both the method of image hiding. It means that the method returns most of the relevant results. The difference between hike % and maximum of individual percentage are very closure to each other, i.e its deviation is just 0-2, except Recall measurements. Here the F1-Score for existing method is MODERATE (68%) and the new method gains HIGH (84%). It indicates there is a perfect precision (78%) and recall (82%). It does not take TNs into account. Here the precision value and F1-Score value are more or less same with the above two methods. The range of deviation and the average lies on the same position for the measurements are Precision and F1-Score. Hence these two measurements are closely correlated with the two methods. Prevalence is the proportion of a particular population with a condition at a specific time. Here the prevalence rate of all classes except Dove and Polarbear for the existing method is lower, which means less common and the Dove and Polarbear contains higher prevalence, which means more common. Just like this the prevalence rate of all classes for the new method is higher, which means more common. Here 20% to 25% of the common elements are shared. Balanced accuracy is used to evaluate how good the classifier method is. It is especially useful when the classes are imbalanced. This still seems a higher (0.74 - existing method & 0.83 - new method), because it has identified the more positives and some classes did not take care about negatives. Hence our new method proved to be a better than the naive "all negatives" classifier. So our new method is able to pick out some positives but not many of them. There is a lot of deviation between the average of Accuracy and Balanced Accuracy. Balances The difference between Accuracy & Balanced Accuracy for the existing method is 8%, but the deviation for the confusion matrix based method is just 1% only. Hence the new method reveals a very good performance in terms of Accuracy. Specificity is the proportion of negatives in the binary classification which are correctly identified. Here the deviation of Spcificity for two methods are 9.20%; but the maximum of individual hike is 8.8%; hence the range of the above two cases is 8.8 – 9.2 and its average is 9%; Here the outcome reveals higher specificity, which means a lower the type-1 error rate i.e the many FPs. The specificity is higher for the Doll & TigerTaj classes, because the FPs are higher. FPR is the probability of falsely rejecting the null hypothesis and is calculated as the ratio between the numbers of negative events wrongly categorized as positive (FPs) and the total number of actual negative events. All the measurements its average starts from 21 and ends with 25 except Accuracy and Recall. The PSNR value deviation between the two methods is 24%.