Image Steganography Using K-Means and DES Algorithm

Sampritha S. Shetty[1*], K. Athmaranjan[2], Shambhavi[3], Shreya D. Rai[4], Soujanya R. Shetty[5]

[1,3,4,5]*B.E. Student, Department of Information Science and Engineering, Srinivas Institute of Technology, Mangaluru, India*

[2]*Associate Professor, Department of Information Science and Engineering, Srinivas Institute of Technology, Mangaluru, India*

*Corresponding author: sampishetty1906@gmail.com

***Abstract*: Steganography is that the process which involves hiding of image, text or any sensitive information inside another image, video or audio in such some way that an attacker won't be ready to detect its presence. Steganography hides the info so nothing appears out of ordinary. it's done to extend the protection against various malicious attacks. Image steganography uses a picture because the cover media to cover the key message. This project uses image steganography method which clusters the image into various segments and hides data in each of the segment. K-means clustering algorithm is employed for image segmentation. Segmentation involves huge set of knowledge within the style of pixels, where each pixel further has three components namely red, green and blue. K-means clustering technique is employed to induce the accurate leads to small period. Segmented images are used for hiding the data using DES algorithm.**

***Keywords*: Stegnography, K-means clustering, DES algorithm.**

1. Introduction

One of the explanations that intruders will be successful is that the most of the knowledge they acquire from a system is during a form that they will read and comprehend. Intruders may reveal the knowledge to others, modify it to misrepresent a private or organization, or use it to launch an attack. One solution to the current problem is, through the utilization of steganography. Steganography may be a technique of hiding information in digital media. In contrast to cryptography, it's to not keep others from knowing the hidden information but it's to stay others from thinking that the knowledge even exists. Steganography become more important as more people join the cyberspace revolution. Steganography is that the art of concealing information in ways in which prevents the detection of hidden messages. Steganography include an array of secret communication methods that hide the message from being seen or discovered. Besides hiding data for confidentiality, this approach of data hiding will be extended to copyright protection for digital media: audio, video and pictures. Therefore, the confidentiality and data integrity are requires to guard against unauthorized access and use. This has resulted in an explosive growth of the sphere of data hiding. Information hiding is an emerging research area, which encompasses applications like copyright protection for digital media, watermarking, fingerprinting, and Steganography hide the secrete message within the host data set and presence imperceptible and is to be reliably communicated to a receiver.

2. Problem Statement

The former consists of linguistic or language sort of hidden writing. The later, like invisible ink, attempt to hide messages physically. Drawback of this linguistic steganography is that user must equip them to own an honest knowledge of linguistry. In recent years everything is trending towards digitalization and with the event of internet technology digital media will be transmitted conveniently over the network. Messages will be secretly carried by digital media using the steganography techniques.

3. Literature Survey

Literature survey is that the documentation of a comprehensive review of the published and unpublished work from secondary sources data within the areas of specific interest to the researcher. it's important for gathering the secondary data for the research which could be proved very helpful within the research. The literature survey may be conducted for several reasons. The literature review may be in any area of the business. Below are the few papers which are referred.

Enhancing the protection and Quality of LSB Based Image Steganography It is a variation of plain LSB (Least Significant Bit) algorithm. The stego-image quality has been improved by using bit-inversion technique. LSB method improves the PSNR of stego image. Through storing the bit patterns or which LSBs are inverted, image could also be obtained correctly. For the improving the robustness of steganography, RC4 algorithm had been implemented to attain the randomization doggo message image bits into cover image pixels rather than storing them sequentially. This method randomly disperses the bits of the message within the cover image and thus, harder for unauthorized people to extract the initial message. The presented method shows good enhancement to Least

Significant Bit technique in consideration to security in addition as image quality [1].

A unique Steganography Method for Image supported Huffman Encoding Novel technique for image steganography supported Huffman Encoding. Two 8-bit gray level image of size M X N and P X Q are used as cover image and secret image respectively. Huffman Encoding is performed over the key image/message before embedding and every little bit of Huffman code of secret image/message is embedded inside the quilt image by altering the smallest amount significant bit (LSB) of every of the pixel's intensities of canopy image. The dimensions of Huffman encoded bit stream and Huffman Table also are embedded inside the quilt image, so as that the Stego-Image becomes standalone information to the receiver. Results show that the algorithm encompasses a high capacity and a decent invisibility. Moreover, Peak Signal to Noise Ratio (PSNR) of stego image with cover image shows better lead to comparison with other existing steganography approaches. The satisfactory security is maintained during this research [2].

A Secure Image Steganography supported RSA Algorithm and Hash-LSB Technique. The problem statement consists of embedding the key message within the LSB of every RGB pixels value of the quilt image. Before embedding the key message must be converted to cipher text using RSA algorithm to boost the secrecy of the message. during this approach we implemented a method called Hash-LSB derived from LSB insertion on images. during this Hash-LSB, we are employing a hash function to judge the positions where to cover the info bits or to be embedded. it's a challenging process which is able to lead us to mix the 2 technologies; one amongst them is RSA algorithm from cryptography and other is Hash-LSB from steganography. Our research has focused on providing an answer for transferring and sharing important data with none compromise in security. All the reputed organizations while sending business documents over the net always use encryption of the info to guard leakage of knowledge about their organization from their rivals or intruders. we've used Hash-LSB and RSA algorithm to form a secure steganography algorithm which is much safer than many systems being employed for the aim of secretly sending the info. Cover Image and Secret Message in our proposed system. First of all we select a real color image of size 512 x 512 for to that as a canopy image and a secret message which is able to be embedded within the cover image [3].

New Image Steganography Method supported K-means Clustering. In this paper, a brand new image steganography method supported k-means clustering for embedding secret messages into a gray image is proposed. The proposed methodology may be a combination of two techniques, image clustering and Least Significant Bits replacement. within the embedding process, a canopy image is segmented into clusters using the K-means clustering algorithm. Each cluster is partitioned into two regions, smooth and complicated, only smooth regions may contain the key data. during this manner, degradation of the Stego image quality is imperceptible to the human eye. For better protection of secret message, a pseudo-random key mechanism is coupled to the implementation of the LSB method. the key data may be recovered directly from the image Stego without respect to cover-image. The experimental results show that the proposed method encompasses a high image quality and a decent embedding capacity [4].

An Approach to Steganography using Local Binary Pattern on CIELAB based K-Means Clustering. This paper presents an approach to steganography using LBP on a cluster formed by the CIELAB (Informally called Lab) based k-means clustering. Image Steganography embeds the key message into the photographs exclusively. Basically Image Steganography uses LSB of pixels to embed the message, but this lone technique is well detectable, so LBP is that the technique which is employed within the approach to embed data into image. It uses the pattern classification characteristics to change the values of pixels in such some way that the modification yields the message requirements and aids the extraction process. But using pure LBP is additionally detectable and messages may be retrieved. Even to secure the situation of LBP hidden message within the image, segmentation is employed for hiding the using only desired areas of the image. For segmentation various clustering techniques may be implemented. For the aim k-means clustering is employed, which provides better accuracy for huge data sets. the photographs have huge number of pixels and each pixel further has 3components namely red, green and blue for color images. So k-means clustering serves the aim even concerning the speed. For k-means clustering using Lab color space adds its own advantages, as Lab color space approximates human vision [5].

4. Implementation

System implementation is that the stage where the theoretical design is converted into a working system, the new system could also be totally new, replacing an existing manual, or automate system or it's going to be a serious modification to an existing system. The system is implemented using Visual studio 2015.

A. *Visual Studio2015*

Microsoft Visual Studio is an integrated development environment (IDE) from Microsoft. it's accustomed develop computer programs, similarly as websites, web apps, web services and mobile apps. It can produce both native code and managed code.

Visual Studio includes a code editor supporting Intelligentsias well as code refactoring. Other built-in tools include a code profiler, designer for building GUI applications, web designer, class designer, and database schema designer. The integrated debugger works both as a source-level debugger and a machine-level debugger. It accepts plug-ins that enhance the functionality at almost every level including adding support for source control systems and adding new toolsets like editors and visual designers for domain-specific languages or toolsets for other aspects of the software development lifecycle.

B. C#

C# may be a general-purpose, multi-paradigm programming language encompassing strong typing, imperative, declarative, functional, generic, object-oriented (class-based) and component-oriented programming disciplines. it absolutely was developed around 2000 by Microsoft as a part of its .NET initiative, and later approved as a world standard by Ecma (ECMA-334) and ISO (ISO/IEC 23270:2018). C# is one amongst the programming languages designed for the Common Language Infrastructure (CLI). Mono is that the name of the free and open-source project to develop a compiler and runtime for the language.

C# was designed by Anders Hejlsberg, and its development team is currently led by Mads Torgersen. the foremost recent version is 8.0, which was released in 2019 alongside Visual Studio 2019 version16.3.

The design goals for C# listed by Ecma are as follows:

- The language is supposed to be a straightforward, modern, general-purpose, object-oriented programing language.
- The language, and implementations should provide support for software engineering principles like strong type checking, array bounds checking, detection of attempts to use uninitialized variables, and automatic pickup. Software robustness, durability, and programmer productivity are important.
- C# is meant to be suitable for writing applications for both hosted and embedded systems, starting from the very large that use sophisticated operating systems, all the way down to the very small having dedicated functions.
- The language is meant to be used in developing software components suitable for deployment in distributed environments.

C. .NET framework

.NET may be a software framework which is meant and developed by Microsoft. within the year 2002 the primary version of the .Net framework 1.0 was released. it's a virtual machine for compiling and executing programs written in numerous languages like C#, VB.Net etc.

It is accustomed develop Form-based applications, Web-based applications, and Web services. there's a spread of programming languages available on the .Net platform, VB.Net and C# being the foremost common ones. it's accustomed build applications for Windows, phone, web, etc. It provides plenty of functionalities and also supports industry standards.

.NET Framework supports quite 60 programming languages during which 11 programming languages are designed and developed by Microsoft. The remaining Non-Microsoft Languages which are supported by .NET Framework but not designed and developed by Microsoft.

The main Components of .NET Framework

- Common Language Runtime (CLR): CLR is that the basic and Virtual Machine component of the .NET Framework. it's the run-time environment within the .NET Framework that runs the codes and helps in making the event process easier by providing the varied services like remoting, thread management, type-safety, memory management, robustness, etc. it's accountable for managing the execution of .NET programs no matter any .NET programing language. It also helps within the management of code, as code that targets the runtime is understood because the Managed Code and code doesn't target to runtime is understood as Unmanaged code.
- Framework Class Library (FCL): it's the gathering of reusable, object-oriented class libraries and methods, etc which will be integrated with CLR. Also called the Assemblies. it's a bit like the header files in C/C++ and packages within the java.

D. Algorithmic steps

1) K-means clustering

Let X = be the set of knowledge points and V = be the set of centers.

1) Randomly select 'c' cluster centers.
2) Calculate the gap between each information and cluster centers.
3) Assign the information point to the cluster center whose distance from the cluster center is minimum of all the cluster centers.
4) Recalculate the new cluster center using:

$$vi =(1/ci)\sum_{(j==0)}^{ci}xi.$$

where, 'ci' represents the amount of knowledge points in ith cluster.
5) Recalculate the gap between each information and new obtained cluster centers.
6) If no information was reassigned then stop, otherwise repeat from step 3).

2) DES algorithm

- Steps for generating keys

There are total of 16 rounds of encryption within the algorithm, each round uses different key. The keys are generated as follows.

1) Compress and transpose the given 64-bit key into a 48-bit key
2) Divide the result into two equal parts: C and D.
3) C and D are left-shifted circularly. For encryption rounds 1, 2, 9, and 16 they're left shifted circularly by 1 bit; for all of the opposite rounds, they're left-circularly shifted by 2.
4) The result's compressed to 48 bits

The results of step 3 is that the input for the subsequent round of key generation.

- Steps for encryption

1) Transpose the bits within the 64-block
2) Divide the result into equal parts: left plain text (1-32 bits) and right plain text (33-64 bits)
3) The resulting parts undergo 16 rounds of encryption

in each round.

The right plain text is expanded

4) The expanded right plain text now consists of 48 bits and is XORed with the 48-bit key.
5) The results of the previous step is split into 8 boxes. Each box contains 6 bits. After prying the eight substitution boxes, each box is reduced from 8 bits to six bits. the primary and last little bit of each box provides the row index, and therefore the remaining bits provide the column index. These indices are wont to look-up values in an exceedingly substitution box. A substitution box has 4 rows, 16 columns, and contains numbers from 0 to fifteen.
6) The result's transposed
7) XOR the left half with the result from the above step. Store this within the right plain text.
8) Store the initial right plain text within the left plain text.
9) These halves are inputs for the subsequent round. There are different keys for every round.
10) After the 16 rounds of encryption, swap the left plain text and therefore the right plain text.
11) Finally, apply the inverse permutation (inverse of the initial permutation) and therefore the cipher text are generated.

- Steps for decryption

The order of the 16 48-bit keys is reversed such key 16 becomes key 1, and so on. Then, the steps for encryption. are applied to the cipher text.

E. Flow Chart

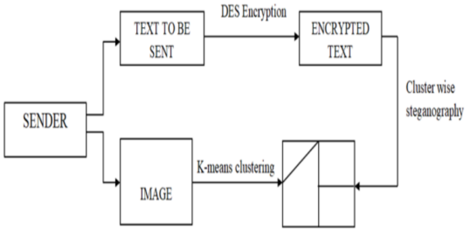


Fig. 1. An overview

5. Proposed System

We design image steganography method which clusters the image into various segments and hides data in each of the segment. K-means clustering algorithm is employed for image segmentation. Segmentation involves huge set of information within the type of pixels, where each pixel further has three components namely red, green and blue. K-means clustering technique is employed to urge the accurate leads to small fundamental measure. Segmented images are used for hiding the knowledge using DES algorithm.

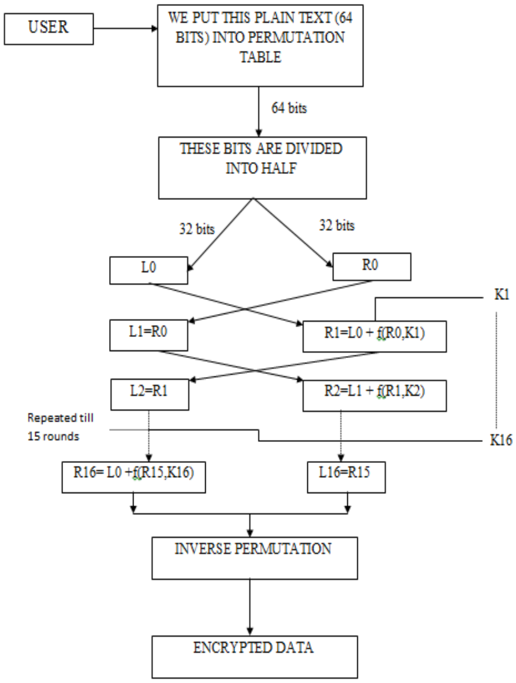


Fig. 2. DES-Structure

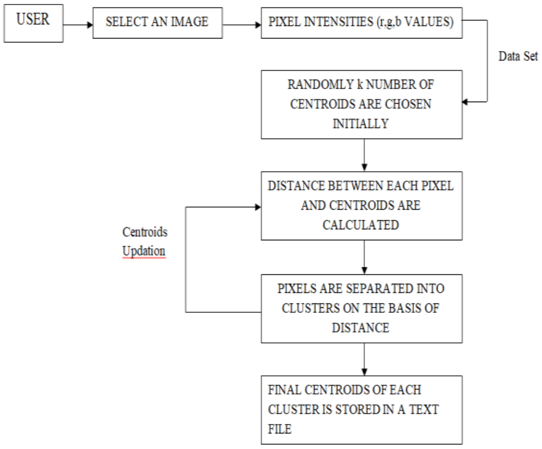
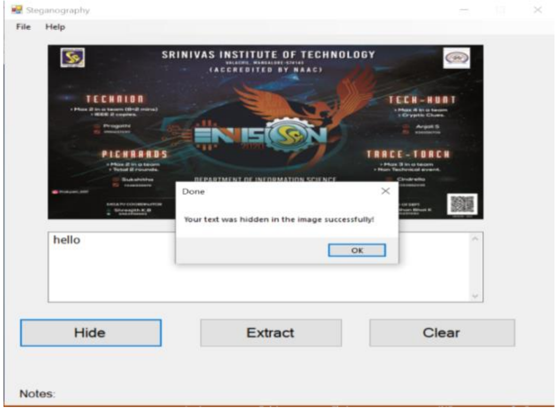

USER
SELECT AN IMAGE
PIXEL INTENSITIES (r,g,b VALUES)
Data Set
RANDOMLY k NUMBER OF CENTROIDS ARE CHOSEN INITIALLY
DISTANCE BETWEEN EACH PIXEL AND CENTROIDS ARE CALCULATED
Centroids Updation
PIXELS ARE SEPARATED INTO CLUSTERS ON THE BASIS OF DISTANCE
FINAL CENTROIDS OF EACH CLUSTER IS STORED IN A TEXT FILE

Fig. 3. K-means algorithm



Done

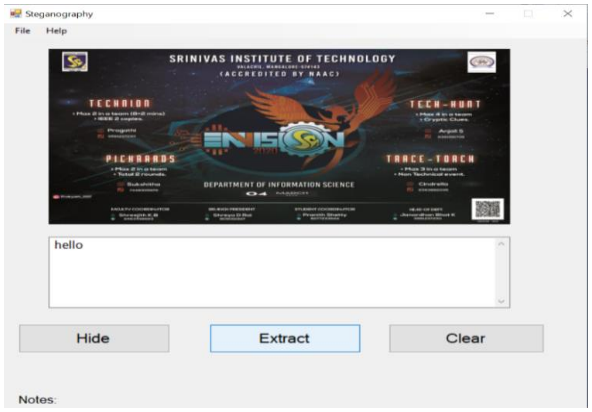
Your text was hidden in the image successfully!

OK

hello

Notes:

Fig. 4. Encryption



hello

Fig. 5. Decryption

6. Conclusion

Steganography may be used for hidden communication. this can be the image steganography system using k-means and DES approach to supply a way of secure communication. This steganography application software provided for the aim to a way to use any style of image formats to hiding any style of files inside them. This project uses image steganography method which clusters the image into various segments and hides data in each of the segment. K-means clustering algorithm is employed for image segmentation. Segmentation involves huge set of information within the variety of pixels, where each pixel further has three components namely red, green and blue. K-means clustering technique is employed to urge the accurate ends up in small fundamental measure. Bit map images are used for hiding the knowledge using DES algorithm. The master work of this application is in supporting any style of pictures without have to convert to bitmap, and lower limitation on file size to cover, due to using maximum memory space in pictures to cover the file.

References

[1] Akhtar N, Johri P, Khan S, "Enhancing the Security and Quality of LSB Based Image Steganography," Computational Intelligence and Communication Networks (CICN), 2013 5th International Conference on, vol., no., pp. 385-390, 27-29, Sept. 2013.

[2] Das R, Tuithung T,"A novel steganography method for image based on Huffman Encoding," Emerging Trends and Applications in Computer Science (NCETACS), 2012 3rd National Conference on, pp. 14-18, 30-31 March 2012.

[3] Shridevi Shetty, "A secure image steganography based on RSA algorithm and hash- LSB technique," Information and Communication Technologies (WICT), 2015 World Congress on, pp. 755-758, Oct. 30 2012-Nov. 2 2015.

[4] Ismail Kich, El Bachir Ameur, Abdelghani Souhar, "New image steganography method based on k-means clustering," BDCA’17 Proceedings of the 2017 Second International Conference on Big data, Cloud and Applications, 29 March 2017.

[5] Diljjeet Singh, "An approach to steganography using local binary pattern on CIELAB based k-means clustering," Computing Communication & Networking Technologies (ICCCNT), 2015 Third International Conference on, pp. 1-11, 26-28 July 2015.