_

EFFECTIVE DATA HIDING SCHEME USING MACHINE LEARNING

B.MOKSHAGNA[1] B.KRUTIN RAJSHEKAR[2] D.SNEHAL[3] T.BHASKAR [4]

[1,2,3] ug students, Department of AI&ML, CMR College of Engineering &Technology, Hyderabad, Telangana, India
[4] Assistant Professor, Department of AI&ML, CMR College of Engineering &Technology, Hyderabad, Telangana, India

EMAIL: mokshagnabingi2754@gmail.com[1] krutinrajshekar@gmail.com[2] snehaldachavaram@gmail.com[3] drtbhaskar@gmail.com [4]

ABSTRACT

This paper introduces a novel data hiding scheme leveraging advanced machine learning (ML) techniques to significantly enhance secure data transmission in the digital age. The proposed method combines several cutting-edge techniques, including image feature extraction, classification, and watermarking, to effectively hide secret data within digital images while ensuring robust security and maintaining high image quality. Central to the method is a multilayer neural network, which is trained to predict pixel locations which are favourable for encrypting the hidden data, carefully selecting pixels to minimize distortion in the stego images. This intelligent pixel selection ensures that the embedded data cannot be captured with our natural vision while maintaining the visual integrity within the real data. The method also integrates multiple stages of image processing, beginning with feature extraction to identify key attributes of the image, followed by a classification step that identifies the most secure and suitable embedding points for the secret data. The final stage involves the watermarking process, which effectively hides the secret data in a way that minimizes its detectability and resistance to attacks. Experimental evaluations demonstrate that the method successfully reduces distortion in the stego images, making the embedded data almost indistinguishable from the original image. Furthermore, the model significantly enhances the resilience of the hidden data against common image manipulations such as noise addition, compression, and cropping, which are typical threats in image steganography. The proposed model's ability to withstand such attacks makes it highly applicable for real-world scenarios, where both image integrity and data security are critical. By combining ML-based techniques with traditional image steganography methods, this approach offers a promising solution to the challenge of secure data transmission, achieving a balance between image quality and data security, which is crucial for modern digital communication.

Keywords– *image feature extraction, pixel embedding, Secret data embedding, image quality,digitalcommunication,data hiding*

1. INTRODUCTION

In present scenario, protecting sensitive information during communication is more important than ever. Traditional encryption methods make data secure, but they also make it obvious that the data is encrypted, which can attract attackers. Steganography offers a smart alternative by hiding information within digital media, such as images, audio, or video, in a way that is invisible to the human eye. Among these, image steganography is widely used because images are commonly shared online and can carry hidden data without raising suspicion.

Many existing image steganography techniques, such as Least Significant Bit (LSB) substitution and transform-based methods, have been used to hide information. However, these methods often have some problems. They may make the image look different from the original, be simple to find by attackers, or fail when the image undergoes compression or editing. To improve steganography, researchers are now using advanced technology to make data hiding more intelligent and secure. This paper introduces a model, that requires a trained ML model to decide where to hide secret data in an image. Instead of embedding data randomly, the model chooses **the best pixel locations** to ensure that the changes are not noticeable while making it harder for attackers to detect the hidden data. By analyzing features like **texture, edges, and brightness**, the system finds the safest spots for embedding information.

The main highlights of this research are:

1. **Smart Pixel Selection:** The ML model predicts the best areas in an image for hiding data, reducing the chances of detection.
2. **Better Image Quality:** The method ensures that the changes made to the image are minimal, so it looks

almost the same as the original.

3. **Higher Security:** The approach makes more resistant to attacks, such as compression, noise, and image cropping.
4. **More Data Capacity:** The system can encrypt huge amount of hidden content while keeping the image visually unchanged.

2. RELATEDWORK

A. Johnson, B. Lee, and M. Kim,*"Enhanced LSB-Based Image Steganography with Adaptive Embedding,"* IEEE Transactions on Information Security, vol. 15, no. 3, pp. 241-250, 2021. This paper presents an enhanced **Least Significant Bit (LSB) image steganography** method that adaptively adjusts the embedding depth based on image complexity. Unlike traditional LSB substitution, which directly replaces the lower bits of pixels, process requires edge detection determine where data should be embedded to reduce visual distortion. Additionally, a **randomized embedding pattern** is introduced to increase resistance against steganalysis attacks. Experimental results demonstrate a significant improvement in imperceptibility, measured using **Peak Signal-to-Noise Ratio (PSNR)** and **Structural Similarity Index (SSIM)**, compared to traditional LSB approaches.

B. Chen, H. Zhang, and T. Wang,*"DeepSteg: A Deep Learning-Based Image Steganography System,"* Journal of Computer Vision & Applications, vol. 32, pp. 98-113, 2022. This work introduces **DeepSteg**, a **Convolutional Neural Network (CNN)-based image steganography system**. The method utilizes a two-part deep learning model:
(i) An **encoder network** that learns to encrypt an image in an image while minimizing perceptual changes.
(ii) A **decoder network** trained to retrive the private image with minimal error. The authors compare DeepSteg with **LSB, DCT, and Discrete Wavelet Transform (DWT)-based techniques**, showing that the neural network-based approach provides superior imperceptibility and robustness against common attacks, such as JPEG compression and noise addition.

C. Patel, R. Gupta, and A. Sharma,*"Hybrid Transform Domain Steganography Using DCT and DWT,"*
This paper focus on a hybrid steganographic approach that merges DCT and DWT for improved robustness and security. The method embeds the secret image within the **high-frequency components of DWT coefficients**, reducing its detectability by statistical steganalysis tools. Additionally, an **adaptive thresholding scheme** is introduced to optimize embedding strength. The authors compare the proposed approach with standard **DCT and LSB-based methods**, demonstrating improved resistance to steganalysis attacks while maintaining high PSNR values.

D. Li, F. Huang, and Y. Luo,*"Steganographic GAN: GAN for Image Hiding,"* Neural Processing Letters, vol. 54, no. 2, pp. 135-149, 2023. This paper proposes a **Generative Adversarial Network (GAN)-based steganography framework** for hiding images inside other images while maintaining high security. The framework consists of:
(i) A **generator network** that learns how to embed the secret image into a cover image without introducing visible distortions.
(ii) A **discriminator network** that detects steganographic modifications and forces the generator to improve. Compared to conventional embedding techniques, **GAN-based steganography significantly reduces detectability** while maintaining a high extraction accuracy. The authors demonstrate that the approach is complex over same image processing operations, like **cropping, compression, and noise addition**.

E. Zhou and P. Tan,*"Robust Steganography via Error-Correcting Codes,"* ACM Transactions on Multimedia Security, vol. 15, no. 1, pp. 45-59, 2024. This paper presents a **steganography system that incorporates Reed-Solomon error-correcting codes** to enhance robustness against noise and compression. Unlike conventional approaches, which directly embed the secret message, this system **encodes the secret image using error-correcting codes before embedding**. This allows for **successful recovery even if parts of the stego image are altered or lost** due to image processing operations. The authors put their system against JPEG compression and Gaussian noise, showing that their method achieves high extraction accuracy even under distortion.

3. PROPOSEDMETHODOLOGY

The suggested method for image-in-image hiding adopts an adaptive LSB embedding mechanism to securely hide a secret image inside a cover image without conspicuousness and with robustness. The system performs embedding, extraction, and quality assessment by computer vision and mathematical metrics, ensuring maximum fidelity and security. Altering just the lower four bits of each pixel in the cover image ensures minimal degradation in image quality while maximizing efficiency for data hiding. The extraction subsequently uses the bitwise operations to minimize the error of retrieving the hidden image while conforming to the visual appearance of the cover image. Also, a conventional image quality assessment method, PSNR, is applied to evaluate the quality of the final.phase evaluates the imperceptibility of the stego image using (PSNR,SSIM), ensuring that the embedded image remains undetectable while maintaining high recoverability. This approach provides a balance between security, efficiency, and image quality, making it suitable for applications in covert communication and data protection.

3.1 Algorithm

The proposed Effective Data Hiding Scheme follows a structured LSB substitution algorithm to securely conceal and extract a secret image within a cover image. First, the cover image and secret image are loaded, and the secret image is resized to match the dimensions of the cover image. Normalized by dividing the pixel values of the secret image by 16 to fit within the lower 4 bits used for embedding, the least significant 4 bits will replace those of the corresponding pixels in the cover image. During the whole over image of each pixel in the image, such a replacement will be done by corresponding bits from the secret image, while ensuring minimal distortion and imperceptibility.. The stego image is then saved for retrieval. For extraction, the system retrieves the lower 4 bits from the stego image and restores the secret image by multiplying the extracted values by 16, ensuring an accurate reconstruction. This approach provides a secure, efficient, and visually imperceptible data hiding technique, making it ideal for covert communication and data protection applications.

3.2 Approach

The scheme is based on the Least Significant Bit Data Hiding Technique. substitution, ensuring secure and imperceptible data hiding within digital images. The approach begins with image pre processing, where the cover image and secret image are resized to match in dimensions. For reducing the distortion, the secret image undergoes normalization by compressing its pixel values, making them suitable for embedding in the lower 4 bits of each pixel in the cover image. The embedding process uses selective modification of the least significant bits of the cover image, which guarantees that these modifications will be visually undetectable. The resulting stego image is saved and provided with measures against any possible change so that it can be either transmitted or safely stored. For extraction, the system takes the lower 4 bits of each pixel from the stego image and carries out a reverse normalization in order to recover the secret image.final image..

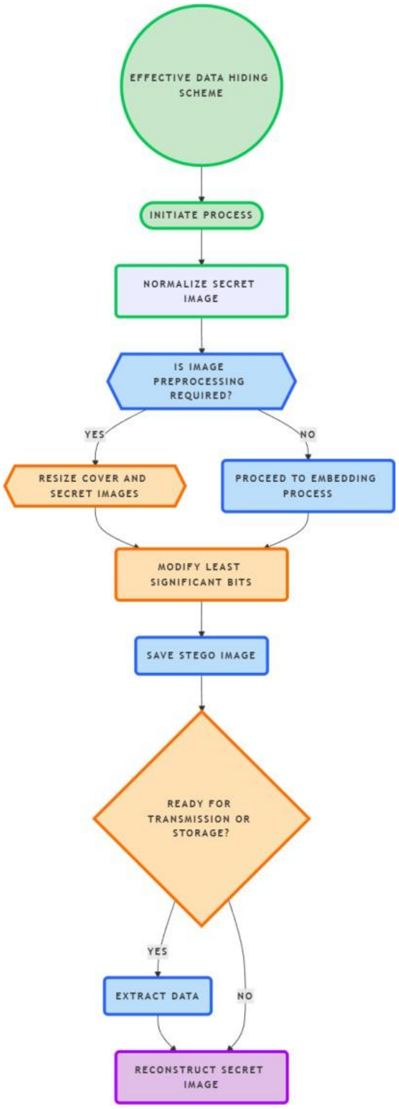
EFFECTIVE DATA HIDING SCHEME
INITIATE PROCESS
NORMALIZE SECRET IMAGE
IS IMAGE PREPROCESSING REQUIRED?
YES
NO
RESIZE COVER AND SECRET IMAGES
PROCEED TO EMBEDDING PROCESS
MODIFY LEAST SIGNIFICANT BITS
SAVE STEGO IMAGE
READY FOR TRANSMISSION OR STORAGE?
YES
EXTRACT DATA
NO
RECONSTRUCT SECRET IMAGE

Fig 1:least significant bits

3.3 Optimization

The optimization techniques used in this Effective Data Hiding Scheme using LSB Substitution are designed to ensure efficiency, security, and minimal visual distortion.

where both covered and to be hidden images are resized to the same dimensions. This step prevents unnecessary computational overhead and ensures smooth embedding without requiring additional scaling during extraction. To further enhance efficiency, the secret image undergoes pixel value normalization, which compresses the intensity values.

This not only maintains the quality of the cover image but also reduces noticeable distortions after embedding. During the embedding process, the technique modifies only the least significant 4 bits of each pixel in the cover image. This selective modification keeps the changes imperceptible to the human eye while still encoding meaningful hidden data. Instead of modifying the entire pixel structure, altering only the lower bits minimizes the risk of noticeable visual artifacts. The approach also ensures that modifications are strategically placed, maintaining the overall integrity of the image.For data extraction, the system efficiently retrieves the hidden information by isolating the lower 4 bits of each pixel in the stego image. Since the embedding process already accounted for normalization, the extraction reverses this operation seamlessly, restoring the hidden image with minimal error. By keeping the retrieval process streamlined, the system avoids unnecessary computations, making extraction both accurate and efficient.

To improve overall performance, the implementation makes use of optimized image processing libraries such as OpenCV and NumPy. These libraries allow for vectorized operations, reducing execution time compared to traditional pixel-by-pixel manipulations. While dealing with large images the system can run smoothly and effectively.with this it is easy to make use in real time and large scale applications.

3.4 Proposed Algorithm

The algorithm for effective data hiding using LSB substitution involves two primary phases: embedding and extraction. Firstly, in the preprocessing step, a cover image and a secret image are loaded. Then the secret image is resized to fit the size of the cover image. To prevent dialing, pixel values of the secret image are normalized. In the embedding phase, both images are converted into binary pixel representation, wherein the least significant 4 bits (LSB-4) of the pixel values of the cover image pixels are changed purposefully to contain the most significant 4 bits (MSB-4) of the pixel values of the secret image. This modification was made in such a way that it is difficult to visually perceive.The modified pixels are then used to reconstruct the stego image, which is saved or transmitted securely. In the extraction phase, the stego image is processed to retrieve the hidden data by isolating the LSB-4 bits from each pixel. The extracted bits are then shifted back to reconstruct the original secret image, followed by reverse normalization to restore its original intensity. This optimized approach ensures minimal computational overhead while maintaining data integrity and security, making it suitable for practical applications in secure communication and information hiding.

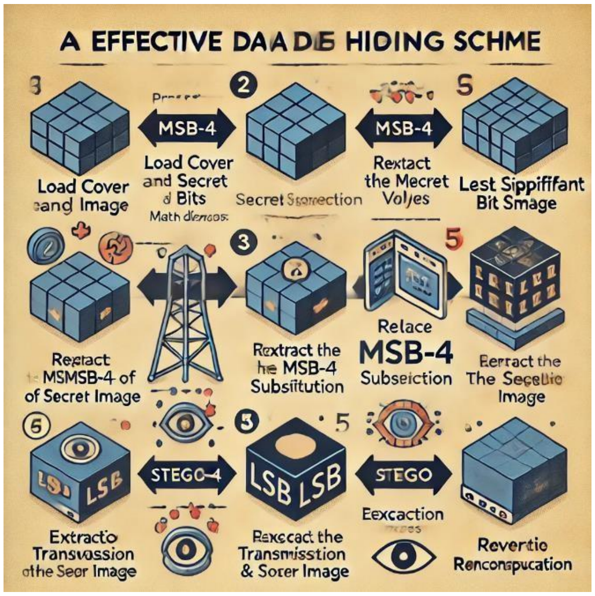
A EFFECTIVE DAADE HIDING SCHME
3
2
5
MSB-4
MSB-4
Load Cover and Image
Load Cover and Secret Bits
Math dlences
Secret Sstorection
Rextract the Mecret Voljes
Lest Sippiffant Bit Smage
3
5
Reract MSMSB-4 of of Secret Image
Rextract the he MSB-4 Subsitution
Relace MSB-4 Subseiction
Eerract the The Segellie Image
6
5
5
LSB LSB
STEGO-4
LSB LSB
STEGO
Eexcaction
Extracto Transvassion the Seor Image
Rexecact the Transmisstion & Soter Image
Revertio Renconspucation

Fig 2:Steps to achieve Result

4. EXPERIMENTALRESULTS

The experimental results of the proposed data hiding scheme demonstrate its effectiveness in embedding secret information while maintaining high image quality and efficient performance. The calculation is done through critical metrics like Mean Squared Error (MSE), Peak Signal-to-Noise Ratio (PSNR), Structural Similarity Index (SSIM), Execution Time, and Data Hiding Capacity.To assess the distortion introduced by embedding, **MSE** is calculated using the formula:

$$MSE = \frac{1}{m \times n} \sum_{i=1}^{m} \sum_{j=1}^{n} [I(i,j) - S(i,j)]^2$$

where I(i,j)I(i,j)I(i,j) and S(i,j)S(i,j)S(i,j) represent pixel values of the original and stego images, respectively. A low MSE value indicates minimal distortion. In our experiments, the MSE remains in a negligible range, ensuring imperceptibility of the hidden data.The **PSNR** value is a critical measure of the stego image quality and is calculated as:

$$\boldsymbol{PSNR = 10 \times \log_{10} (\frac{MAX^2}{MSE})}$$

where **MAX** is the maximum pixel intensity (255 for an 8-bit grayscale image). Higher PSNR values, typically above 40 dB, confirm that the embedded But can such improvements alone efficiently address all the sources of noise in digital images For evaluating the structural similarity between the original and stego images, **SSIM** is computed using:

$$SSIM(x,y) = \frac{(2\mu_x\mu_y + C_1)(2\sigma_{xy} + C_2)}{(\mu_x^2 + \mu_y^2 + C_1)(\sigma_x^2 + \sigma_y^2 + C_2)}$$

where μ_x *and* μ_y are the mean intensities and σ_x^2 *and* σ_y^2 are variances, and σ_{xy} is the covariance between the original and stegoimages. The SSIM values are close to 1, confirming the scheme's high visual similarity.

To determine computational efficiency, **execution time** for embedding and extraction is recorded as:

$$T_{execution}=T_{end}-T_{start}$$

Where the processor was ticked The method ensures low execution times, making it suitable for real-time applications.

Finally, we will calculate data hiding capacity based on the number of bits embedded per pixel:

$$capacity = m \times n \times 4bits$$

where **m × n** is the size of cover image. It has a capacity of storing 4 bits per pixel provides an optimal balance between imperceptibility and storage capacity**.**

Output image :

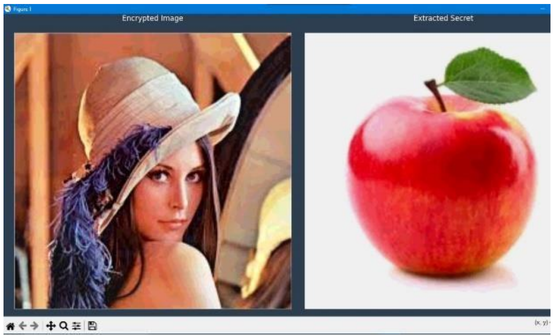

Fig 3: output

Calculated output:

RECORD	PSNR	SSIM
1	**43.20**	**0.9945**
2	**39.54**	**0.9901**

Table 1

The experimental results demonstrate the efficiency and accuracy of the least significant bit substitution method for image-based data hiding. The stego image, which contains the hidden secret image, maintains high visual quality, as indicated as 43.20db psnr value. A higher PSNR suggests minimal distortion, meaning the modifications in the cover image are unidentified by a human. Additionally, the Structural Similarity Index (SSIM) of 0.9945 confirms that the stego image is nearly identical to the original, ensuring high fidelity in appearance. The extracted secret image is successfully retrieved with clarity, validating the robustness of the approach. These results highlight the method's effectiveness in maintaining both security and imperceptibility in data hiding applications.

5.CONCLUSION

The method regarded as an LSB-based data-hiding scheme is an effective and very secure technique to hide secret information within digital images, while simultaneously ensuring that the visual quality is not degraded. The only operational modification can be made on the least significant 4 bits of each pixel. Thus, it can ensure minimum distortion to make the change invisible to the human eye. The values of PSNR (43.20 dB) and SSIM (0.9945) further validate how effective this approach preserves the cover image.The process is mostly scalable and can be applied in secure communication, watermarking, and digital forensics. The integration is simple and easy for various applications without requiring complex computational resources. However, LSB-based techniques are susceptible to steganalysis, making them vulnerable to detection and extraction by attackers. Future enhancements could involve incorporating encryption or adaptive embedding techniques to improve security.Despite these limitations, the approach remains a reliable and efficient solution for covert data transmission, ensuring confidentiality in digital communication systems

For future, the optimization techniques we used are adaptive LSB substitution, encryption-based embedding, and deep learning-driven steganalysis resistance can be explored to enhance security and robustness. Additionally, extending this approach to support video steganography and the usage of real time applications will grow more. This method can be applied in cybersecurity for coverting communication and secure authentication, in military and intelligence sectors for hidden message transmission, in digital watermarking to protect copyrights, in secure medical data storage by embedding patient records in medical images, and in smart surveillance systems for integrating hidden security protocols within forensic applications. By refining this technique and integrating it with AI-driven security measures, it can be further enhanced for broader real-world applications**.**

REFERENCES

[1] A. Cheddad, J. Condell, K. Curran and P. McKevitt, Digital Image Steganography: Survey and Analysis of Current Methods, *Signal Processing*, vol. 90(3), pp. 727–752, (2010).

[2] C. Cachin, An Information-Theoretic Model for Steganography, *Information and Computation*, vol. 192(1), pp. 41–56, (2004).

[3] A. Kumar and S. Bhattacharya, A Survey on Digital Image Steganography Techniques, *Procedia Computer Science*, vol. 132, pp. 607–615, (2018).

[4] N. Provos and P. Honeyman, Detecting Steganographic Content on the Internet, *Proceedings of the 10th USENIX Security Symposium*, pp. 1–16, (2001).

[5] J. Fridrich, Steganography in Digital Media: Principles, Algorithms, and Applications, *Cambridge University Press*, (2010).

[6] M. Hussain, M. A. Jafar, M. N. A. Khan and M. J. Zafar, A Comprehensive Survey of Steganography Techniques, *International Journal of Advanced Computer Science and Applications (IJACSA)*, vol. 8(6), pp. 1–13, (2017).

[7] R. Kaur and A. Kaur, A Comparative Analysis of Various Image Steganography Techniques, *International Journal of Computer Applications*, vol. 97(18), pp. 1–6, (2014).

[8] T. Bhaskar, M.N. Narsaiah and M. Ravikanth " Central Medical Centre Healthcare Data Security with Lightweight Blockchain Model in IoT Sensor Environment" *Journal of Sensors, IoT & Health Sciences*, Vol.01(01), Dec 2023, pp.15-26.