iJRASET
International Journal For Research in Applied Science and Engineering Technology

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 14 **Issue:** VII **Month of publication:** July 2026

DOI: https://doi.org/10.22214/ijraset.2026.84117

www.ijraset.com

Call: 08813907089 | E-mail ID: ijraset@gmail.com

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 14 Issue VII July 2026- Available at www.ijraset.com

Hybrid Transformer-Based Steganography Using Frequency Embedding and Adversarial Robustness

Aakash Bonagiri[1], N Naveen Kumar[2]
[1]*M.Tech Scholar, Dept of CSE, JNTUH UCESTH, Hyderabad, India*
[2]*Associate Professor, Dept of CSE, JNTUH UCESTH, Hyderabad, India*

Abstract: ***Steganography is the practice of hiding secret information within digital media such as images, audio, or video. It ensures confidential communication by concealing the existence of data itself, unlike encryption. Modern research focuses on improving invisibility, security, and resistance to detection using AI and deep learning. This project aims to design a secure and intelligent image steganography system using a hybrid Transformer model. It focuses on increasing data hiding capacity while maintaining image quality and reducing detectability. The system will be capable of resisting steganalysis attacks, compression, and noise distortions in real-world usage. Recent studies show that deep-learning steganography models still suffer from low robustness and poor scalability. Most approaches fail under compression or noise, and their hidden data can be detected by advanced AI models. High computational cost, limited payload capacity, and dataset dependency further affect their reliability. This project introduces a hybrid Transformer integrated with Discrete Cosine Transform (DCT) for frequency embedding. By combining spatial and frequency domains, it ensures better concealment and robustness. Adversarial training with a steganalysis discriminator will enhance security against modern detection models. The system will achieve higher PSNR and SSIM scores, proving superior imperceptibility and accuracy.***

Implementation will use Python, PyTorch, and OpenCV for model training and image processing. Datasets like COCO, BOSSBase, and ImageNet will be used for evaluation. Performance metrics such as PSNR, SSIM, MSE, and BER will measure quality and accuracy. Development and testing will be carried out in Jupyter Notebook or Google Colab environments.

Keywords: Image Steganography, Hybrid Transformer Model, Discrete Cosine Transform (DCT), Adversarial Training, Steganalysis Resistance

I. INTRODUCTION

Steganography is the process of concealing secret information within digital media in such a way that the existence of the information itself remains hidden. Derived from the Greek words *steganos* (hidden) and *graphein* (to write), steganography provides secure communication by embedding data into a cover object such as text, images, audio, or video files. Unlike cryptography, which transforms information into an unreadable form, steganography focuses on disguising the presence of the message. The effectiveness of a steganographic system is generally evaluated based on three important factors: imperceptibility, which ensures that embedding causes minimal visible or audible distortion; payload capacity, which determines the amount of data that can be hidden; and security, which measures the resistance of the hidden information against detection and extraction.

Text steganography is one of the earliest forms of information hiding and involves embedding secret messages within textual content by manipulating spaces, punctuation marks, font styles, word choices, or grammatical structures. Although these methods are simple and require minimal storage, text documents possess limited redundancy, resulting in lower payload capacity and reduced robustness. Image steganography, on the other hand, has become the most widely studied technique because digital images contain a large amount of redundant pixel information that can be modified without noticeable visual changes. Traditional methods such as Least Significant Bit (LSB) embedding alter pixel values to hide information, whereas transform-domain techniques based on Discrete Cosine Transform (DCT) and Discrete Wavelet Transform (DWT) embed data within frequency components to improve robustness and security. Image steganography finds applications in secure communication, digital watermarking, copyright protection, and privacy preservation. Audio steganography conceals secret information within digital audio signals by exploiting the limitations of the human auditory system. Common techniques include Least Significant Bit modification, phase coding, spread spectrum methods, and echo hiding. These approaches enable hidden information to be embedded while preserving audio quality and maintaining resistance against certain attacks. Audio steganography is extensively used in secure voice communication, multimedia protection, and digital rights management systems. However, compression and signal processing operations may degrade the embedded information and affect extraction accuracy. Despite these limitations, audio-based methods provide a favorable balance between payload capacity and perceptual transparency.

Video steganography extends the concept of information hiding to digital video sequences by utilizing both spatial and temporal redundancy present in consecutive frames. Secret information can be embedded in individual frames using spatial-domain methods such as LSB replacement or within frequency coefficients using transform-domain techniques based on DCT, DWT, or hybrid approaches. Since videos consist of thousands of frames accompanied by audio streams, they offer significantly higher payload capacity and improved security compared to images and audio files. Video steganography is widely employed in secure multimedia communication, copyright protection, military applications, and digital forensics. However, challenges such as video compression, frame manipulation, noise, and steganalysis attacks can affect the integrity of hidden information. With recent advancements in deep learning, Transformers, and adversarial training techniques, modern steganographic systems are evolving toward highly secure, robust, and intelligent information-hiding solutions capable of resisting sophisticated detection mechanisms.

II. LITERATURE REVIEW

A. A High-Capacity Image Steganography Based on LSB Inversion and Summing of Bit Pairs

Nowadays, computer users face a challenge in maintaining the security of their information. There are many techniques for transferring information safely over networks. Image steganography, one of these techniques, is an effective way to conceal important information. The Least Significant Bit (LSB) technique is a famous method for data hiding. This technique exchanges the LSBs of the cover image pixel with the mystery message bits. However, it was always the challenging problem of increasing the number of bits utilized for embedding that led to a decreased quality of the stego image. Hence, this research targets optimizing the stego image quality by inverting the LSBs of image pixels to reduce the changes that occur in pixels. This paper proposes a new method based on summing the pairs and inverting the LSBs of the cover image pixels. The secret data are arranged in pairs, and then the sum of every two successive pairs is computed. After that, the secret bits are embedded by inverting the pixel bits depending on that sum. Finally, for more enhancement of the results, some of the inverted bits will be inverted again. Rigorous experiments using benchmark datasets confirm that the proposed scheme yields superior results compared to the substitution LSB scheme. Also, shows the effectiveness of the proposed approach in achieving a high payload with acceptable visual quality for stego image that cannot be noticed by human eyes when compared with state-of-the-art steganography techniques.

B. A Novel LSB Steganography Technique Using Image Segmentation

Steganography is a process to hide data inside a cover file mostly used in media files like image, video, and audio files. Least significant bit (LSB) steganography is a technique where the least significant bits of pixels are used for information hiding. The purpose of using only those bits is to minimize the visual impact of the hidden data on the image file. LSB technique of steganography is one of the most popular forms of steganography available today. As a result, various steganalysis techniques are developed for this steganography technique. One of them is the visual analysis of pixels through pixel modificationsto expose hidden data in a visual manner. The proposed method achieves resistance to this attack using an image segmentation model and extracting the most texture-complex areas of an image and hiding information in these specific areas as pseudo-randomized least significant bit replacements. As the outcome of the study, an alternative approach to LSB steganography that results competitively with existing methods is provided.

C. An Effective Steganographic Technique for Hiding the Image Data Using the LSB Technique

Steganography is the art and science of writing secret messages so that neither the sender nor the intended recipient knows there is a hidden message. Data hiding is the art of hiding data for various reasons, such as keeping private data, secure, confidential data, etc. With increasing data exchange over a computer network, information security has become a significant issue. There are many methods used for data hiding, and steganography is a well-known technique. Steganography is the art of invisible contact and science. Steganography is the process through which the presence of a message can be kept secret. The objective of this paper is to hide data using the LSB (Least Significant Bit) technique into images that can be detected only by the specified user. We have developed a user-friendly GUI such that it can be used with the utmost ease. This paper is motivated to hide the message stated by the user in the dialog box given within the picture. The secret text is converted to the ciphertext to make it more stable. The sender selects the cover image, and it is used to generate the secured Stegno image, which is identical to the cover image. With the support of a private or public communication network, on the other hand, the stegno image can be saved and sent to the designated user, i.e., the recipient downloads the stegno image and can retrieve the secret text concealed in the stegno image using that same application. As for the watermarking, we have visible and invisible we have used the same LSB technique

D. An Improved Phase Coding Audio Steganography Algorithm

As AI technology continues to advance, voice cloning is becoming increasingly easy. Recently, cases of fraud involving audio forgery using AI technology have emerged, making it particularly important to covertly embed information and verify the authenticity and integrity of audio. Digital Audio Watermarking has thus become a crucial tool in this context. This study proposes an improved Phase Coding audio steganography algorithm that dynamically segments the audio signal and embeds information into the phase components of the mid-frequency range. This approach not only enhances the algorithm's resistance to steganalysis but also simplifies the computational process, ensuring the authenticity and integrity of audio both efficiently and securely.

III. PROPOSED SYSTEM

This work proposes a secure web-based multi-modal steganography platform that allows users to hide and retrieve secret information using image, audio, video, and text files. The system combines steganography and encryption techniques to provide secure communication. It also includes user authentication, activity tracking, and a layered architecture to improve security and usability.

A. Steganography Techniques

1) Image Steganography

The image steganography module hides encrypted data inside digital images such as PNG and BMP files.

Adaptive Least Significant Bit (LSB) embedding hides more data in textured regions and less in smooth regions to maintain image quality.

Random pixel selection using a user key and capacity checking improve security and ensure sufficient storage space. DCT-based embedding for JPEG images provides better robustness, and image quality is evaluated using PSNR, SSIM, MSE, and embedding capacity.

2) Audio Steganography

The audio steganography module hides secret information inside WAV audio files.

LSB encoding provides high payload capacity while preserving sound quality.

Phase coding and echo hiding techniques improve robustness and make the hidden information less noticeable.

Synchronization information is added for accurate extraction, and quality is measured using SNR and related metrics.

3) Video Steganography

The video steganography module uses both video frames and audio tracks to provide higher data hiding capacity.

Frame-selective embedding and LSB-based techniques are applied to motion-rich frames to minimize visual distortion.

DCT-based embedding and motion vector modification improve robustness against video compression and processing.

The audio track serves as an additional hiding channel, while video quality is evaluated using PSNR, SSIM, VMAF, and bitrate changes.

4) Text Steganography

The text steganography module hides information within text documents without changing their meaning.

Invisible Unicode characters and whitespace manipulation are used to encode secret data.

Synonym substitution preserves sentence meaning while embedding information.

Capital letters and punctuation patterns are used as alternative methods for hiding data.

Overall, the proposed system combines steganography and encryption techniques to provide a secure, reliable, and user-friendly platform for hiding sensitive information across multiple digital media formats.

B. System Architecture

The proposed system follows a five-layer service-oriented architecture, where each layer performs a specific task and communicates with neighboring layers through defined interfaces.

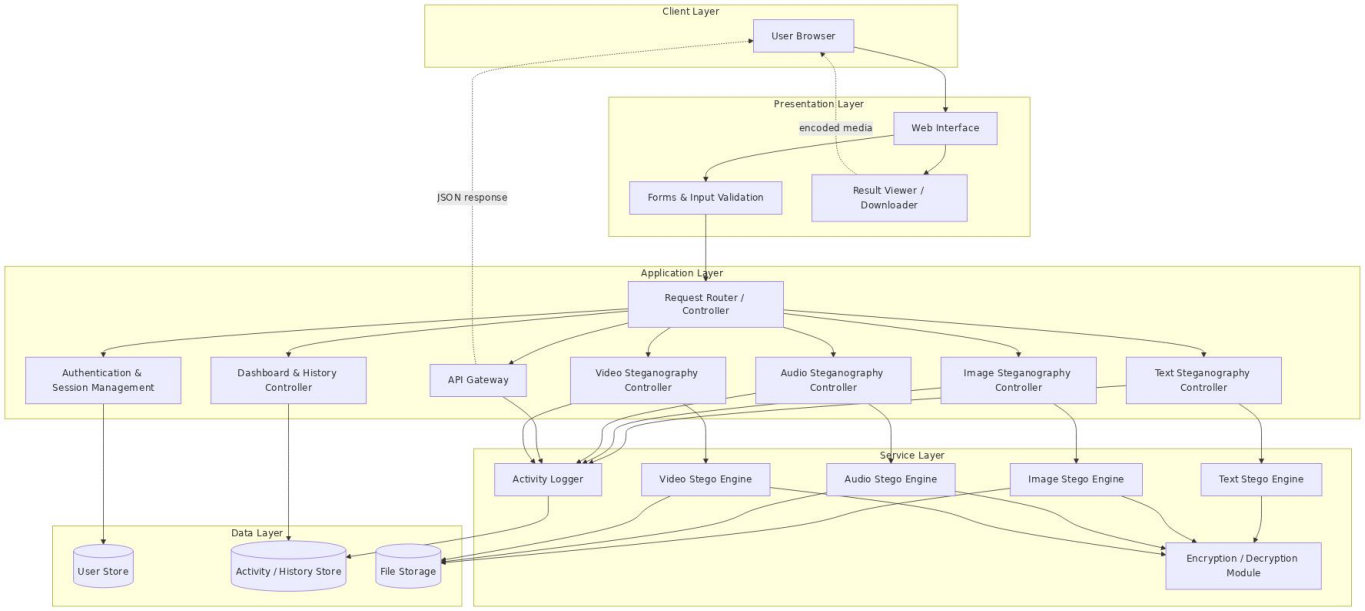


Fig 1 – System Architecture

- *Client Layer:* The client layer acts as the entry point to the system through a web browser. It sends secure HTTPS requests and receives the generated stego files for download. Users interact with the platform through this layer.
- *Presentation Layer:* The presentation layer provides the user interface for encoding and decoding operations across different media types. It validates user inputs such as file type, file size, and key strength before processing. It also previews and allows downloading of generated image, audio, video, and text outputs.
- *Application Layer:* The application layer manages requests, performs authentication, and directs operations to the appropriate steganography module. It provides REST APIs and maintains user dashboards and activity history. Separate controllers handle image, audio, video, and text processing tasks
- *Service Layer:* The service layer contains the core steganography engines responsible for embedding and extracting secret information. It also provides encryption and decryption services using AES and key generation mechanisms. In addition, an activity logger records all operations for auditing and monitoring purposes.
- *Data Layer:* The data layer stores user information, roles, and authentication details. It maintains activity logs and history records for all steganographic operations. It also stores cover files and generated stego media, including large video files, using unique identifiers.

C. *Working Flow of the Proposed System*

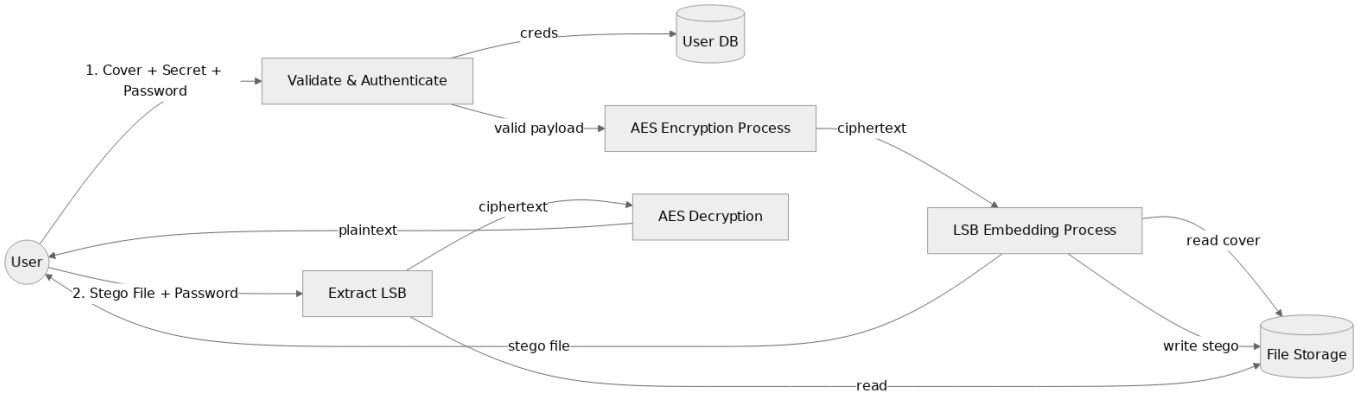


Fig 2 - Workflow of the proposed system

The Diagram describes how data moves through the proposed steganography system. The user acts as the main external entity and provides the cover file, secret message, and password during encoding, and the stego file and password during decoding. The User Database stores user information, while the File Storage module stores cover files and generated stego files.
During the encoding process, the user is first authenticated by the system. The secret message is then encrypted using the AES algorithm to provide security. The encrypted data is embedded into the cover media using the LSB technique, and the generated stego file is stored and returned to the user. The encoding flow can be represented as:
User → Authentication → AES Encryption → LSB Embedding → File Storage → User

During the decoding process, the user uploads the stego file and provides the password. The system extracts the hidden encrypted data using the LSB extraction process and decrypts it using AES to recover the original message. The recovered message is then displayed to the user. The decoding flow is represented as:
User → LSB Extraction → AES Decryption → User

The proposed system follows the principle of encrypting the secret message before embedding it into the cover media. This provides an additional layer of security and ensures that only authorized users can retrieve the hidden information. The File Storage module acts as a secure repository for storing cover and stego files.

D. System Performance

Modality	Cover Size	Payload	Encode (s)	Decode (s)
Text	5 KB	256 B	0.004	0.003
Image	512×512 PNG	8 KB	0.31	0.22
Audio	10 s WAV (1.7 MB)	4 KB	0.18	0.14
Video	10 s AVI (~14 MB)	32 KB	2.41	1.97

Table 1 – System Performance

The performance of the proposed multi-modal steganography system was evaluated by measuring the encoding and decoding times for different media types. Experimental results show that text steganography provides the fastest performance, requiring only 0.004 seconds for encoding and 0.003 seconds for decoding a 256-byte payload. Image steganography successfully embeds 8 KB of data in a 512 × 512 PNG image with encoding and decoding times of 0.31 seconds and 0.22 seconds, respectively.
For audio steganography, a 10-second WAV file of size 1.7 MB was used to hide a 4 KB payload. The encoding process required 0.18 seconds, while the decoding process took 0.14 seconds, demonstrating efficient performance with minimal delay. Video steganography, which handles larger files and higher payload capacities, required more processing time. A 10-second AVI video of approximately 14 MB was able to hide 32 KB of data with encoding and decoding times of 2.41 seconds and 1.97 seconds, respectively.
Overall, the results indicate that the proposed system achieves efficient encoding and decoding across all supported media types. Although video steganography requires more computational time due to its larger file size and complexity, the system maintains acceptable performance while providing higher payload capacity and secure data hiding.

IV. RESULTS

The proposed multi-modal steganography system was successfully implemented and tested using text, image, audio, and video carriers. Experimental results demonstrate that the system is capable of securely embedding and extracting secret information while preserving the quality of the cover media. The integration of AES encryption with steganographic techniques provided an additional layer of security, ensuring that hidden data remained protected from unauthorized access.

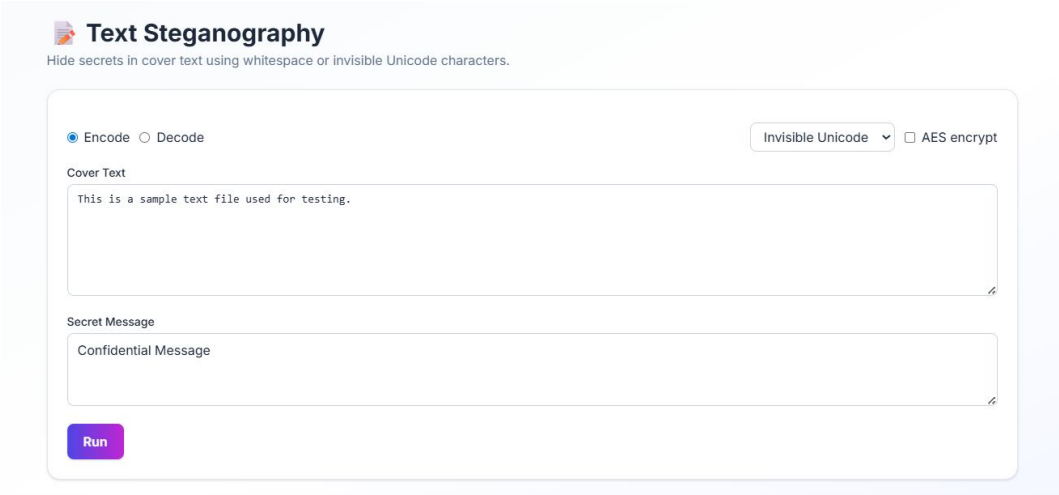


Fig 3- Text Steganography encoding mode

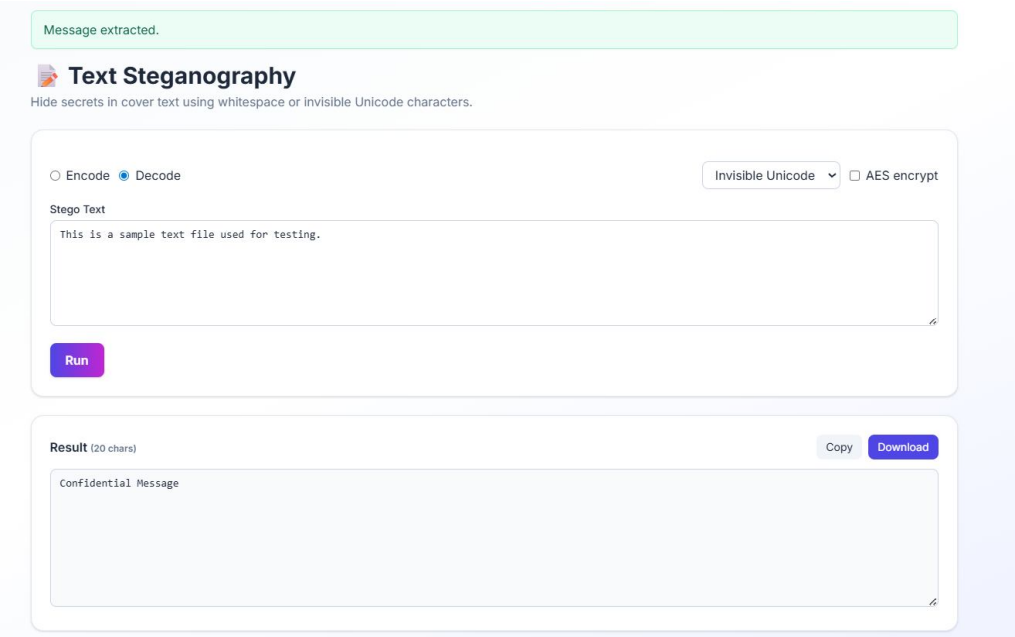


Fig 4- Text Steganography decoding mode

V. CONCLUSION

In this work, a secure multi-modal steganography platform was designed and implemented for hiding confidential information within text, image, audio, and video files. The system combines steganography techniques with AES encryption to provide an additional layer of security, ensuring that hidden data remains protected from unauthorized access. Different embedding methods were employed for each media type to achieve effective data hiding while preserving the quality of the cover media.

Experimental results demonstrated that the proposed system successfully performs both encoding and decoding operations with high accuracy and acceptable processing time. Image and audio steganography maintained good perceptual quality, while video steganography provided higher payload capacity for hiding larger amounts of information. The use of authentication, activity logging, and secure storage further enhanced the reliability and usability of the platform. Overall, the proposed system provides a secure, efficient, and flexible solution for covert communication across multiple digital media formats. The combination of steganography and cryptography improves confidentiality and makes the system suitable for applications requiring secure information exchange. Future enhancements may include the integration of deep learning-based steganography techniques, improved robustness against steganalysis attacks, and support for cloud-based and real-time communication environments.

REFERENCES

[1] A. Rehman, T. Saba, T. Mahmood, Z. Mehmood, M. Shah, and A. Anjum, "Data hiding technique in steganography for information security using number theory," Journal of Information Security and Applications, vol. 78, p. 103618, 2023. J. Breckling, Ed., *The Analysis of Directional Time Series: Applications to Wind Speed and Direction*, ser. Lecture Notes in Statistics. Berlin, Germany: Springer, 1989, vol. 61.

[2] R. Meng, Q. Cui, Z. Zhou, Z. Fu, and X. Sun, "A steganography algorithm based on CycleGAN for covert communication in the Internet of Things," IEEE Access, vol. 11, pp. 12345–12358, 2023.

[3] S. Kaur, S. Singh, M. Kaur, and H.-N. Lee, "A systematic review of computational image steganography approaches," Archives of Computational Methods in Engineering, vol. 30, no. 7, pp. 4775–4805, 2023.

[4] N. Subramanian, O. Elharrouss, S. Al-Maadeed, and A. Bouridane, "Image steganography: A review of the recent advances," IEEE Access, vol. 11, pp. 23409–23423, 2023.

[5] M. M. Hashim, A. A. Mahmood, and M. S. M. Rahim, "A comparative study among different techniques of audio steganography," Indonesian Journal of Electrical Engineering and Computer Science, vol. 33, no. 1, pp. 456–467, 2024.

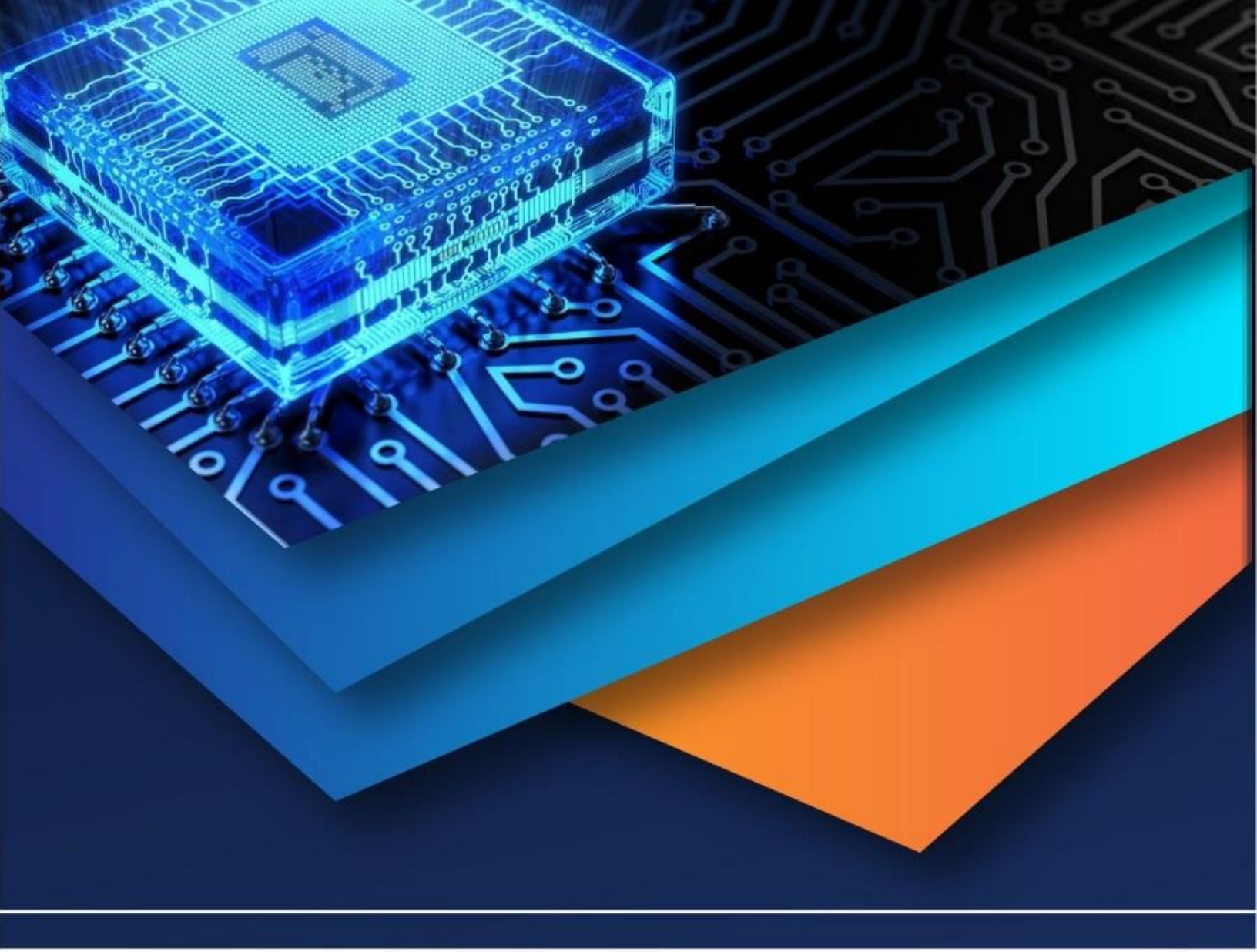
[6] P. Singh and R. Kumar, "An improved LSB image steganography technique using bit-inversion in spatial domain," Multimedia Tools and Applications, vol. 83, pp. 18211–18233, 2024.

[7] A. K. Sahu, M. Hassaballah, R. Saraswathi Rao, and G. Suresh, "Logistic-map based fragile image watermarking scheme for tamper detection and localization," Multimedia Tools and Applications, vol. 82, pp. 24069–24100, 2023. *FLEXChip Signal Processor (MC68175/D)*, Motorola, 1996.

[8] H. Kheddar, M. Hemis, Y. Himeur, D. Megías, and A. Amira, "Deep learning for steganography and steganalysis: Recent advances and future perspectives," Neurocomputing, vol. 581, p. 127495, 2024. A. Karnik, "Performance of TCP congestion control with rate feedback: TCP/ABR and rate adaptive TCP/IP," M. Eng. thesis, Indian Institute of Science, Bangalore, India, Jan. 1999.

[9] M. A. Majeed, R. Sulaiman, Z. Shukur, and M. K. Hasan, "A review on text steganography techniques," Mathematics, vol. 11, no. 4, p. 920, 2023.

[10] D. Megías, W. Mazurczyk, and M. Kuribayashi, "Data hiding and its applications: Digital watermarking and steganography," Applied Sciences, vol. 13, no. 21, p. 11924, 2023.



10.22214/IJRASET

45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY