

Enhancing Data Security: Implementing Steganography in Web Application

Salam Khalaf Abdullah¹, Ammar Mohammedali Fadhil²

¹Department of Computer Engineering, Al-Nukhba University, Baghdad, Iraq
Baghdad, Iraq

²Middle Technical University, Information and Communication Technology, Al-Zafaraniya,
Baghdad, Iraq

ARTICLE INFO

Received: 30 Dec 2024

Revised: 05 Feb 2025

Accepted: 25 Feb 2025

ABSTRACT

In the era of information technology, ensuring the security of information over the Internet is a very important and priority issue. Conventional encryption methods have become traditional and more vulnerable to attacks due to their challenging nature. Steganography is presented as a promising technique to enhance the security of information in web applications. Steganography provides the art of hiding secret data in non-confidential media in a way that the data inside cannot be perceived. In this study, we propose a new approach to integrate steganography with web applications to maintain the security of secure data transmitted over the Internet. Our research involves a method based on hiding data in image pixels by the priority method produced by the SVM classifier based on the features extracted from the image itself. The proposed method has proven its worth through experimental results that preserve the integrity and quality of the image while ensuring the transfer of a good amount of secure data. It has demonstrated the ease of implementing steganography in web applications in an imperceptible manner. This allows the user to upload and retrieve secure messages securely. The study was evaluated through quantitative and qualitative methods that relied on increasing the PSNR (85 dB) ratio and chi-square that works to find the probability of the presence of secure data in the image. This research presents a promising solution for information security against cyber threats.

Keywords: Web Applications, Data Security, Steganography, Digital Image Processing, Information Hiding, Cyber security.

INTRODUCTION

In the era of information technology and the Internet, the security of information transmitted over the Internet has become a source of concern for both individuals and organizations. Encryption is considered one of the traditional methods of securing data, but due to its explicit nature, it faces problems from hackers [1]. In order to enhance data security, especially in Internet applications that have become more popular at the present time, studies are working to find new ways to secure data. Data hiding techniques have appeared before and their techniques have been developed, especially steganography, which can be defined as hiding data inside non-secret media that are not declared confidential [2]. Steganography provides security from two parties: the first is hiding data in a way that cannot be recovered except with the presence of the hider, and the second is that hiding data is not visible to the eye [3]. Protecting Internet application data is very necessary because it is circulated among many of the public. One of the best techniques that hide data currently is steganography because it does not attract the attention of professionals and there is no challenge in advance such as encryption.

One of the important things that currently occupies researchers is the integration of data hiding technology into web applications, which has witnessed widespread use due to the spread of the Internet and its applications [4]. Through these technologies (steganography), the weaknesses that occur in data transmission over the network are addressed. By including hidden messages within images used on the web or any file, whether audio or text, on web pages, steganography technology ensures that the data is out of sight of intruders [5]. The process of hiding data, even if it is complex in terms of encryption, also makes the data hidden in applications without suspicion.

There are many web applications that provide many services on the Internet, such as e-commerce, banking, social networks, scientific research, and others, that can benefit from the data security provided by Steganography technology [6]. Applications deal with security data, often personal or financial, so they must be protected because they are basically prime targets for cyber-attacks. Steganography technology can protect these platforms and the data within them during transmission, thus achieving the ultimate goal of maintaining security in general [7].

Machine learning and artificial intelligence in general have improved the Steganography technique, making it more robust and complex for intruders. Artificial intelligence techniques can be both beneficial and harmful at the same time [8-9]. On the negative side, it is its ability to predict and provide solutions, which creates a major challenge for Steganography. Therefore, it is important to maintain the strength of the proposed algorithm to avoid any external interference. Therefore, stenography in web applications enhances and strengthens data security. By hiding important and sensitive data in unsuspecting data, stenography plays an important role in digital communications against the increasing cyber threats [10]. With the proliferation of web applications, their integration and data hiding in them maintains users' privacy and security in the digital landscape.

The main goal is to improve data security through web applications and maintain a secure environment. As well as increasing the intangibility of images in web applications and increasing the amount of information to achieve the quantity and quality of hidden data. To increase data security by using random keys and a substitution method based on machine learning in the pixel locations in the image. Increasing data security at the present time represents a major challenge to stand against cyber-attacks and maintain information security.

This manuscript is carefully designed to explore how steganography can enhance the security of web applications and the introduction discussed, highlights the objectives and challenges. Then, previous studies related to the topic are reviewed and a detailed analysis of the importance of the topic is provided. In the methodology section, the system architecture is explained in detail, the process of embedding data in the image and the principle of hiding data. The embedding and extraction mechanism from the sender and receiver sides are explained. The results section provides an analytical study and discussion of the data hiding technique and discusses the security improvements through the results. The conclusion section summarizes the implications of using steganography in web applications and important remarks about the proposed method and then explains the most important future studies in this regard.

LITERATURE REVIEW

Steganography is a term that has been used since ancient times, whether before or after the development of information technology. Many studies have shown interest in the subject of data hiding because there is no apparent challenge to detecting it, in addition to the possibility of transferring and hiding data imperceptibly [11]. The beginning of the actual development of the principle of steganography began after the Internet revolution, as data became available to everyone [12]. Data hiding has become a must for the safety and security of data. The study [13] confirmed the importance of hiding data in various media, which varies according to the size of the data they contain. A study was conducted by [14] to confirm the work of steganography in high-resolution images and the reason for choosing images is their availability on the Internet and their frequent circulation. Many studies have improved the methods of hiding data in images with a specific term known as steganography images [15]. Steganography enters into many applications, including sensitive and important ones, so the focus is on improving methods and moving away from traditional methods, especially in the field of Internet applications [16]. Artificial intelligence and its algorithms are used in the steganography technique by finding suitable places to hide the secret bits in the image pixels [17].

Machine learning techniques such as the famous classifiers SVM or KNN and others were used to find the appropriate place for the secret data in the image (pixels) and with a smart distribution system that cannot be predicted by intruders [18]. Deep learning was used in data hiding, which had a significant impact on increasing the security of data in the image or any other media [19]. Steganography was used in web applications to hide sensitive data in images contained in the web page and in an encryption method that can only be solved by the encryption key provided by the sender [20]. Many web applications have hidden data in an unusual way, which is hiding the secret message in the data transmission protocol, which is characterized by great effectiveness in terms of security, but its disadvantage is the transfer of small amounts of data [21]. Web applications were based on three pillars of media that can contain secret data or in other words can be exploited by steganography, which are images, texts and sound [22]. A study based on hiding data in images in web applications [23] by using the least significant bits, but it is considered

an expected method for hackers. The encryption method is adopted by [24] of changing the locations of random pixels, which increases the security and complexity of the data, and by using an external encryption key to choose the appropriate pixel to contain the secret data.

Web applications have been used in many studies to hide data using a URL to the application site [25]. Web applications are important and sensitive and the hidden data must be through them because they are available to everyone and do not require a specific system or environment. They depend mainly on direct browsers and can work from anywhere provided that the Internet and browser are available. Because of their wide availability, the steganography must be strong enough to achieve great imperceptibility.

PREPOSED METHODS

In this section, we proposed a data hiding technique in a web application to transfer security data to the other party without being noticed. Using the Steganography method, which hides the text in the image taken from the web application. In general, the user is interested in the public details that he sees, and even the intruder tracks things that appear to be secure, but in our topic here, the data is not suspicious because it is included in images that do not arouse suspicion.

One of the components of the web application is the web browser, and this browser consists of several elements, the most important of which are images, which are what we are looking for in this research in order to include secret data through it. When writing and designing a web page, we bring the image and include the data through it, and thus it is completely secure and does not attract attention. Then the user or anyone interested in the presentation of the web page.

The proposed algorithm aims to embed a secret message inside the image in a way that is not visible to others while maintaining the image quality. This process is called steganography, and the algorithm will be implemented in a web application environment.

The following steps consider the main issue in the proposed method:

Step1: Download the image from the site using the Java script for embedding.

Step2: Preprocessing regarding managing long secret messages with image size.

Step3: Analyze image pixels and choose candidate pixels for embedding.

Step4: Prepare secret bits for embedding into LSB pixel bits.

Step5: Reconstruct image pixels to produce a stego image.

Step6: On another side reverse the process for extracting secret messages.

The data comes from a standard dataset in the form of an image and the pre-processing begins. The secret message is selected and the appropriate length is prepared for embedding. After that, the processing is done by selecting the appropriate pixels for addition in the least significant bits (LSB). The addition process continues until the image is finished and then it is sent to the other side (receiver) to extract the data according to what is followed in the stego key generated by the sender. Figure 1 illustrate the process in detail.

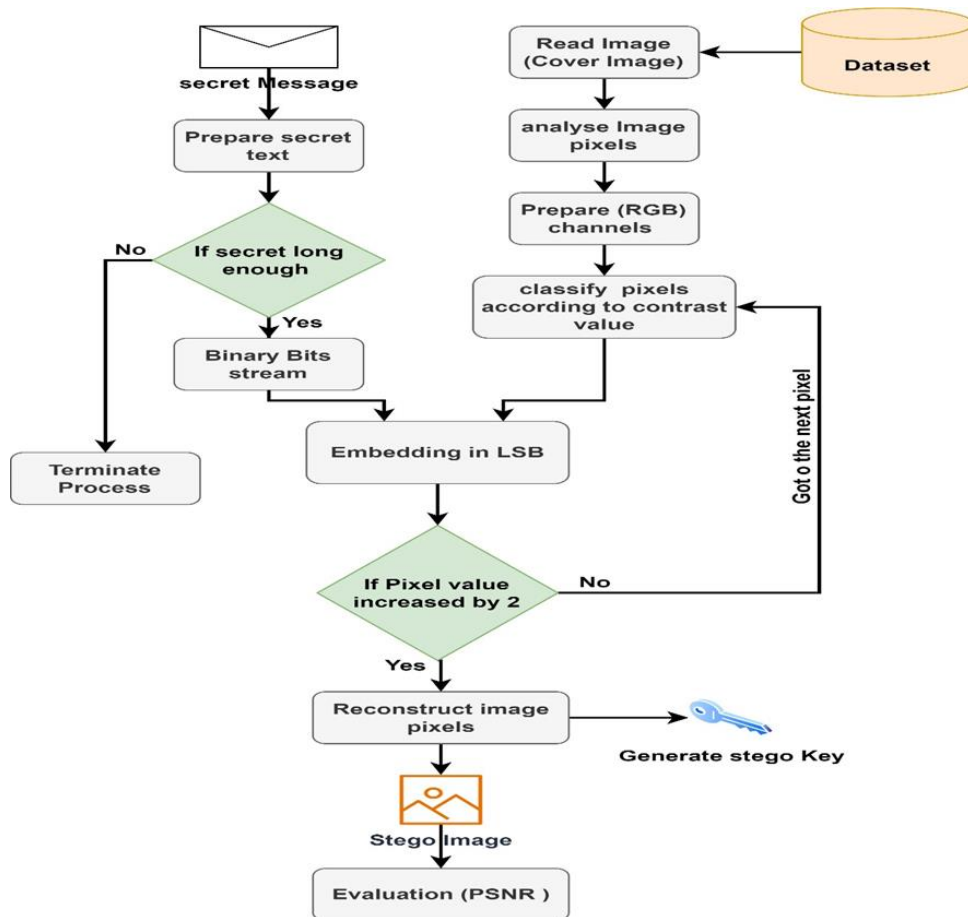


Figure 1: general framework of proposed method

The embedding process is in the form of a logical addition of the secret bits to the pixel bits in the image, but with different pixel sequences in the image. The SVM classifier classifies the pixels based on the difference and color frequency in the pixel. In other words, if the difference is large, the embedding is avoided because it will be obvious to the intruder, but if the difference is small, it is suitable for embedding. In the case of the pixel being suitable for embedding, it is checked whether the embedding exceeds the pixel value by 2 or not to avoid an increase in the sticky pixel value, which may be suspicious. Embedding process can illustrate in Figure 2.

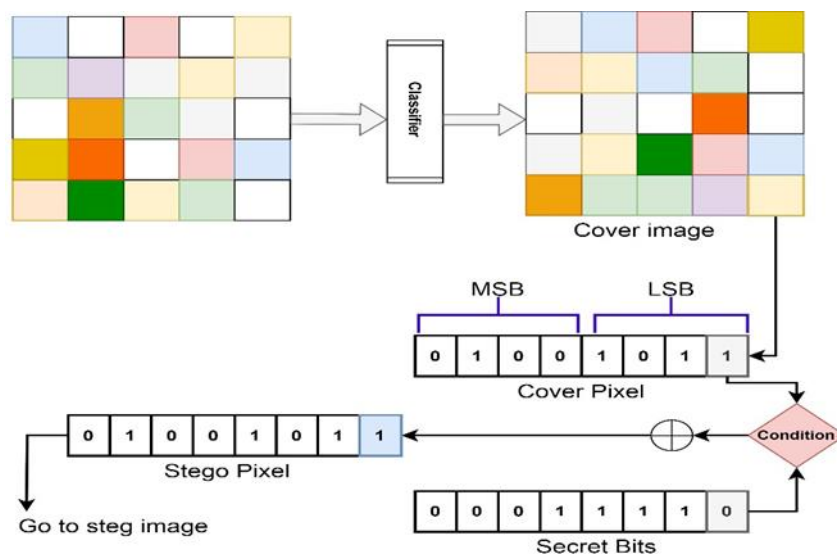


Figure 2: Embedding process within proposed model

The original image consists of a set of pixels called the cover image and each pixel is called a cover pixel. The sequence and priorities of the pixels are changed using SVM. Then the embedding process takes place which involves merging the secret bit with the last bit of the LSB to form the stego pixel which is produce the stego image.

Once the embedding process is complete, it is necessary to examine and evaluate the image to ensure that it is not imperceptible when sent. There are special criteria for evaluating the stego image, which we will discuss in the next section of the results and their analysis

RESULTS AND DISCUSSION

In stenography systems, a standard dataset is used to accurately evaluate the work with previous studies as a single basis. The dataset contains a set of images to be worked on and the results to be compared with previous studies. Color images (RGB) are used for each pixel with three color channels, so it becomes 24 bits. Each color has 8 bits and the embedding is done in each pixel of the secret message. Images that are common in previous studies will be selected for embedding.

Image processing affects its quality to some extent and may cause loss of image information. There are two types of evaluations that are used to evaluate an image. The first is to find statistical differences in the image and the second method is subjective observation through the human eye without criteria. In this paper we will present results related to Peak Signal to Noise Ratio (PSNR) which we get from statistical and objective analyses. PSNR can define by:

$$PSNR = 10 \cdot \log_{10} \left(\frac{MAX_I^2}{MSE} \right) \quad (1)$$

Then can find the expression Mean Square Error (MSE) by:

$$MSE = \frac{1}{mn} \sum_{i=0}^{m-1} \sum_{j=0}^{n-1} [I(i, j) - K(i, j)]^2 \quad (2)$$

Consider MAX maximum possible pixel value in stego image such as m, n are the dimensions of image where I, K are the original pixel and noisy pixel.

The PSNR value is negatively affected by MSE, and applying the equation to we get the result in Table 1 for Lena image from dataset.





Table 1: Imperceptibility of Lena Image (512× 512)

Payload (Bytes)	Embedding Ratio %	PSNR (dB)	
		Simple LSB	Proposed method
16383	6.25	73.534	89.522
32768	12.5	71.762	88.725
49152	18.75	70.342	87.127
65536	25	70.101	86.299

The data embedded to the image is 16384 bytes, which is about 6.25% of the image. In the normal addition to the LSB, the PSNR increases with the decrease in the amount of embedded information. In the proposed method, we found that it is much better than the normal embedding because the inclusion of the secret data is very random due to the classification of the SVM, as the selection, in embedding to being random, is carefully selected in the appropriate places by the classifier.

Table 2 summarizes the PSNR of some images taken from the standard dataset. It can be seen that the method used is feasible and achieves better results because the embedding of the secret data is after processing and not directly, and the embedding is according to a classification based on the colour contrast between the adjacent pixels (8-neighbors) and are entered as features for the classifier to be classified.

Table 2: Imperceptibility of Images from Dataset

Images	Image size	PSNR (dB)
 peppers	256×256 pixels	85.234
 Jet	227×250 pixels	83.611
 Camera man	260×280 pixels	87.481
 lake	270×290 pixels	80.928

As mentioned in Steganography, imperceptibility means that the secret message hidden inside the image cannot be observed by the human eye. In fact, there are two main factors that affect imperceptibility: the first is the amount of data included in the image and the embedding method itself. The amount of information cannot be controlled only by compressing it, but the method is the main concern of researchers' contributions. The measure of imperceptibility here is PSNR.

One of the most important advantages of Steganography is that the secret message is invisible to the public and if it is discovered by a hacker, it can only be opened using the secret key agreed upon between the sender and the recipient. The invisibility is important because it removes suspicion from the image and thus transfers information in a very secure manner. The original image is no different from the image containing secret information in different methods, and the results are identical to the human eye. After including 65782 bytes, but they differ in the amount of noise they carry, and this can only be distinguished by statistical methods previously referred to in the PSNR equation. If the amount of data included increases, the noise increases with it and is noticeable to the naked eye.

There is an inverse relationship between PSNR and the amount of data that the image can absorb. Increasing the size of the secret data leads to image distortion and thus a decrease in PSNR. Therefore, all methods aim to maintain a balance between the size of the secret data and the size of the image.

The chi-square attack works to find any hidden data in the image by finding the LSB frequency in the stego image. The chi-square graph represents the probability of the data being present in all the pixels in the image. At first, the percentage is 3% until the rest of the pixels are calculated, then the real percentage comes. Because all the letters in the English language come with an initial value close to the LSB value, then the difference begins. This test detects the percentage of differences by comparing it to the pixel frequency.

After embedding 16653 bytes in the image, different results are generated. In the usual way, LSB detects the chi-square equation. 50% of the image is considered hidden data and its maximum probability is 1.

The proposed method is difficult to detect by chi-square, as it is almost impossible to recognize the image with the naked eye. After statistical analysis, the proposed method proved its effectiveness, the probability is close to zero since the beginning of the analysis, which proves that there is no indication of secret data in the image.

In short, based on the embedding method, the security of a stenography system can be detected. By considering the probability of the data distribution or its difference from the original image of pixel values. The following equation represents the calculation of the chi-square value.

$$X^2 = \sum \frac{(\text{Observed} - \text{Expected})^2}{\text{Expected}} \quad (3)$$

CONCLUSION

With the rapid digital development, data security has become a priority in the lives of individuals and organizations. Steganography provides the ideal solution in web applications to enhance data security by embedding secret information in images within web applications. To prevent it from being detected by the naked eye, the Steganography application is integrated with the web application to avoid raising suspicions when transferring data over the network without suspicion, and hence the prevention of secret data breach by unauthorized persons.

In stenography, the information hidden in the image is in a complex way that cannot be revealed even if it is intercepted by intruders. Extracting the secret message contained in the image is almost impossible without the key. Steganography can be used on multiple types of media such as audio, video, text files, etc., which makes it flexible and adaptable. In the proposed system, the secret message is converted to ASCII code and then to a series of binary bits and then embedded in the image pixels. The image is selected based on the extracted features of the image, such as contrast and color variation, and then the pixels are classified according to the features by SVM classifier to choose the priority of the pixels to be embedded. An image is imported from the web application and the Steganography method is used to hide the secret message in it. In this way, the secret data in the web application is protected.

Two types of results were evaluated on the basis of which are qualitative and quantitative. Imperceptibility is one of the most important measures used in stenography and a satisfactory result was obtained for PSNR (87 dB). Chi-square is one of the statistical measures used and depends on measuring the probability of embedding in the pixel bits in the image.

REFERENCES

- [1] Wang, S., Zhong, H., Wang, Z., Lv, H., Jiang, J., & Pu, J. (2024). Aldehyde modified nanocellulose-based fluorescent hydrogel toward multistage data security encryption. *International Journal of Biological Macromolecules*, 256, 128359.
- [2] Rustad, S., Andono, P. N., & Shidik, G. F. (2023). Digital image steganography survey and investigation (goal, assessment, method, development, and dataset). *Signal processing*, 206, 108908.
- [3] Lakshmi Sirisha, B., & Chandra Mohan, B. (2021). Review on spatial domain image steganography techniques. *Journal of Discrete Mathematical Sciences and Cryptography*, 24(6), 1873-1883.
- [4] Ilchev, S., & Ilcheva, Z. (2014). A new approach to data hiding for web-based applications. Prof. Marin Drinov Academic Publishing House.
- [5] Azizan, N., Abdul Aziz, F. F., Kamarzaman, N. S., Abdul Jamil, N., Jaafar, J. S., & Mohd Shafiee, A. N. File Hiding Web Application (FHWA) Using Image Steganography. *European Proceedings of Multidisciplinary Sciences*.
- [6] Abed, N. K., Shahzad, A., & Mohammedali, A. (2023, October). An improve service quality of mobile banking using deep learning method for customer satisfaction. In *AIP Conference Proceedings* (Vol. 2746, No. 1). AIP Publishing.
- [7] Dhawan, S., & Gupta, R. (2021). Analysis of various data security techniques of steganography: A survey. *Information Security Journal: A Global Perspective*, 30(2), 63-87.
- [8] Sulong, G., & Mohammedali, A. (2015). RECOGNITION OF HUMAN ACTIVITIES FROM STILL IMAGE USING NOVEL CLASSIFIER. *Journal of Theoretical & Applied Information Technology*, 71(1).
- [9] Atiyha, B. T., Aljabbar, S., Ali, A., & Jaber, A. (2019). An improved cost estimation for unit commitment using back propagation algorithm. *Malaysian Journal of Fundamental and Applied Sciences*, 15(2), 243-248.
- [10] Fadhil, A. M. (2016). Bit inverting map method for improved steganography scheme. Diss. Universiti Teknologi Malaysia.
- [11] Singh, S., Singh, A. K., & Ghrera, S. P. (2017, February). A recent survey on data hiding techniques. In *2017 International Conference on I-SMAC (IoT in Social, Mobile, Analytics and Cloud)(I-SMAC)* (pp. 882-886). IEEE.
- [12] Hussain, M., Wahab, A. W. A., Idris, Y. I. B., Ho, A. T., & Jung, K. H. (2018). Image steganography in spatial domain: A survey. *Signal Processing: Image Communication*, 65, 46-66.
- [13] Shiu, H. J., Lin, B. S., Lin, B. S., Huang, P. Y., Huang, C. H., & Lei, C. L. (2018). Data hiding on social media communications using text steganography. In *Risks and Security of Internet and Systems: 12th International Conference, CRiSIS 2017, Dinard, France, September 19-21, 2017, Revised Selected Papers 12* (pp. 217-224). Springer International Publishing.

-
- [14] Rustad, S., Andono, P. N., & Shidik, G. F. (2023). Digital image steganography survey and investigation (goal, assessment, method, development, and dataset). *Signal processing*, 206, 108908.
 - [15] Xu, Y., Mou, C., Hu, Y., Xie, J., & Zhang, J. (2022). Robust invertible image steganography. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition* (pp. 7875-7884).
 - [16] Abdullah, D. M., Ameen, S. Y., Omar, N., Salih, A. A., Ahmed, D. M., Kak, S. F., ... Rashid, Z. N. (2021). Secure data transfer over internet using image steganography. *Asian Journal of Research in Computer Science*, 10(3), 33-52.
 - [17] Fadhil, A. M., Jalo, H. N., Mohammad, O. F. (2023). Improved Security of a Deep Learning-Based Steganography System with Imperceptibility Preservation. *International journal of electrical and computer engineering systems*, 14(1), 73-81.
 - [18] Sukumar, A., Subramaniaswamy, V., Vijayakumar, V., & Ravi, L. (2020). A secure multimedia steganography scheme using hybrid transform and support vector machine for cloud-based storage. *Multimedia Tools and Applications*, 79(15), 10825-10849.
 - [19] Chaumont, M. (2020). Deep learning in steganography and steganalysis. In *Digital media steganography* (pp. 321-349). Academic Press.
 - [20] Azizan, N., Abdul Aziz, F. F., Kamarzaman, N. S., Abdul Jamil, N., Jaafar, J. S., & Mohd Shafiee, A. N. File Hiding Web Application (FHWA) Using Image Steganography. *European Proceedings of Multidisciplinary Sciences*.
 - [21] Gurunath, R., & Samanta, D. (2022). A novel approach for semantic web application in online education based on steganography. *International Journal of Web-Based Learning and Teaching Technologies (IJWLTT)*, 17(4), 1-13.
 - [22] bin Mohd Nai, M. K. A., Zulkipli, N. H. N., & Aziz, S. R. A. (2021). Web-Application For Securing Message Using LSB Algorithm Steganography And Hybrid Encryption. *i-JaMCSIIX Universiti Teknologi MARA Cawangan Melaka Kampus Jasin 77300 Merlimau, Melaka Tel: 062645000*, 35.
 - [23] Othman, N. A., Osman, M. N., Sedek, K. A., & Shariffudin, M. N. H. (2023). Image Steganography Using Web Application. *Journal of Computing Research and Innovation*, 8(2), 1-11.
 - [24] Hoffman, A. (2024). *Web application security*. " O'Reilly Media, Inc."
 - [25] Khalid, M. N., Rasheed, K., & Abid, M. M. (2020). Web vulnerability finder (WVF): automated black-box web vulnerability scanner. *Int J Inf Technol Comput Sci*, 12(4), 38-46.