IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 8.102 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 13, Issue 4, April 2024

DOI: 10.17148/IJARCCE.2024.13478

Advancing Steganalysis: Comparative Analysis of JUNIWARD, JMIPOD, and UERD

Mrs. N. BHARGAVI[1], B SAI RAM KOUSIK[2], T JESHWANTH KUMAR[3], SHAIK ASHRAF[4], P RAMA KRISHNA[5]

Asst Prof, Department of CSE, KL University, Andhra Pradesh, India[1].

CSE, KL University, Andhra Pradesh, India[2-5]

Abstract: In the rapidly evolving landscape of information security, steganalysis algorithms play a pivotal role in safeguarding digital content integrity. Three notable algorithms, JUNIWARD, JMIPOD, and UERD, stand at the forefront of this endeavour, each offering unique capabilities in detecting covert information embedding. JUNIWARD employs advanced statistical modelling and machine learning techniques to discern characteristic artifacts induced by popular data hiding methods.

This results in high detection rate while maintaining a low false positive rateS, solidifying its position as a significant advancement in steganalysis technology. JMIPOD, tailored for JPEG- compressed images, leverages sophisticated feature extraction and statistical analysis to identify subtle discrepancies introduced by covert information embedding. By exploiting vulnerabilities in the JPEG compression process, JMIPOD achieves impressive detection rates across a wide range of embedding rates, ensuring the integrity of digitally compressed content. UERD, the Universal Ensemble for Robust Detection, presents a pioneering approach by employing an ensemble of

carefully curated classifiers. This methodology capitalizes on the complementary strengths of multiple steganalysis methods, leading to enhanced robustness against a broad spectrum of steganographic schemes. Rigorous experimentation across various datasets showcases UERD's superiority in detection performance and adaptability to evolving data hiding methodologies

Keywords: Steganalysis, Information Security, Digital Content, Integrity, JUNIWARD, JMIPOD, UERD.

I. INTRODUCTION

In an era dominated by information exchange through digital channels, the practice of concealing information within innocuous cover media, commonly known as steganography, has emerged as a potent tool for covert communication. This artful technique challenges the very foundation of information security, as it enables the surreptitious transmission of sensitive data through seemingly innocuous files or messages. In response to this growing threat, steganalysis, the science of detecting and uncovering such concealed information, has become an indispensable component of modern information security protocols. Steganalysis entails the development of sophisticated algorithms and methodologies designed to scrutinize digital content for telltale signs of embedded information. Through the careful analysis of statistical irregularities, structural inconsistencies, and other subtle artifacts, steganalysis seeks to unveil the hidden payload within a given medium. The field has witnessed significant advancements over the years, leveraging techniques ranging from classical statistical analyses to sophisticated machine learning approaches. As the prevalence of steganographic techniques continues to proliferate, driven by the ever-evolving landscape of digital

communication, the need for robust and efficient steganalysis methodologies has never been more critical. The challenges are manifold, encompassing the detection of both well- established steganographic algorithms and emerging, more sophisticated embedding strategies. Additionally, the advent of high-definition multimedia content and the ubiquity of high-speed data transmission have imposed new computational demands on steganalysis systems, necessitating a reevaluation of existing methodologies and the exploration of novel approaches. As the prevalence of steganographic techniques continues to proliferate, driven by the ever-evolving landscape of digital communication, the need for robust and efficient steganalysis methodologies has never been more critical. The challenges are manifold, encompassing the detection of both well-established steganographic algorithms and emerging, more sophisticated embedding strategies. Additionally, the advent of high- definition multimedia content and the ubiquity of high-speed data transmission have imposed new computational demands on steganalysis systems, necessitating a reevaluation of existing methodologies and the exploration of novel approaches.

DOI: 10.17148/IJARCCE.2024.13478

Through this work, we aim to contribute to the ongoing discourse in steganalysis, offering insights and strategies that bolster the efficacy of information security measures in the face of an increasingly sophisticated steganographic landscape.

1.1 Issue Proposal

In the landscape of image encryption, a critical concern arises from the ever-evolving nature of cryptanalytic techniques and the pressing need for algorithms to adapt accordingly. As encryption methods become more sophisticated, so do the strategies employed by adversaries seeking to breach their security. This dynamic interplay necessitates a proactive approach towards identifying and addressing potential vulnerabilities in contemporary image encryption algorithms.

One pertinent issue lies in the susceptibility of certain algorithms to differential and linear cryptanalysis, two powerful cryptanalytic tools with demonstrated efficacy against various cryptographic primitives. Understanding the extent to which these attacks pose a threat to current image encryption methodologies is paramount for fortifying their security. Furthermore, the emergence of quantum computing poses a looming challenge to classical cryptographic systems, potentially rendering current encryption algorithms vulnerable to rapid factorization of large numbers. This raises questions about the quantum resistance of contemporary image encryption techniques and the need for quantum-resistant alternatives. Another pressing concern is the practical implementation of image encryption algorithms, where subtle flaws or inadequate parameter choices can inadvertently introduce vulnerabilities. The evaluation of algorithm robustness in real-world scenarios, including considerations for computational efficiency and resource constraints, becomes a pivotal aspect of addressing this issue.

Additionally, the integration of image encryption with emerging technologies such as blockchain and edge computing necessitates a revaluation of the compatibility and security implications of these combinations. Exploring the potential synergies and risks associated with these integrations can offer valuable insights into fortifying image encryption in contemporary computing environments.

This section will provide a comprehensive exploration of these critical issues, presenting a foundation for the subsequent analysis and evaluation of contemporary image encryption algorithms. By addressing these concerns head-on, we aim to contribute to the ongoing advancement of secure image encryption techniques in the face of evolving cryptographic landscapes.

II. LITERATURE REVIEW

Fridrich, J., & Kodovský, J. (2012)[1]. Rich models for steganalysis of digital images. IEEE Transactions on Information Forensics and Security, 7(3), 868-882. Fridrich and Kodovský's 2012

paper introduced rich models, a transformative approach to steganalysis. By employing advanced statistical techniques, they discerned subtle patterns in digital images. Their work significantly improved detection accuracy, even against advanced steganographic techniques. This innovative framework has had a lasting impact on steganalysis, shaping subsequent research in information security and digital forensics.Ker, A. D., Böhme, R., & Winkler, D. (2007)[2]. An improved LSB matching steganalysis. Proceedings of the 9th Workshop on Multimedia & Security, 5-14 represent significant contributions to the field of steganalysis. They have not only enhanced our understanding of steganographic detection but also paved the way for further innovations in this critical area of information security and digital forensics. Kodovský, J., Fridrich, J., & Holub, V. (2012)[3]. Ensemble classifiers for steganalysis of digital media. IEEE Transactions on Information Forensics and Security, 7(2), 432-444 Transactions on Information Forensics and Security, 7(2), 432-444Their innovative approach has not only advanced the state-of-the-art in steganalysis but also established a framework for continued progress in the detection of hidden information within digital media. This work stands as a testament to the ingenuity and adaptability required to stay ahead in the ongoing cat-and-mouse game between steganographers and steganalysts.Holub, V., Fridrich, J., & Denemark, T. (2014)[4]. Universal distortion function for steganography in an arbitrary domain. IEEE Transactions on Information Forensics and Security, 9(3), 392-405 By introducing a versatile framework for concealing information in an arbitrary domain, the authors have significantly expanded the applicability and adaptability of steganographic techniques. This work stands as a testament to the ingenuity and innovation required to push the boundaries of covert communication in the digital age. Denemark, T., Fridrich, J., & Goljan, M. (2014)[5]. Detection of double-compression in JPEG images for applications in steganography. IEEE Transactions on Information Forensics and Security, 9(3), 428-441 Their innovative approach provides a crucial layer of security, safeguarding hidden information from potential exposure due to multiple compressions.

International Journal of Advanced Research in Computer and Communication Engineering
Impact Factor 8.102 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 13, Issue 4, April 2024
DOI: 10.17148/IJARCCE.2024.13478

This work exemplifies the intersection of cutting-edge research with practical applications, reinforcing the vital role of steganography in contemporary information security. Chen, M., Fridrich, J., & Goljan, M. (2010)[6]. Digital imaging sensor identification (DISI) for forensic applications. IEEE Transactions on Information Forensics and Security, 5(3), 494-506 Their innovative approach provides a powerful tool for attributing images to specific imaging sensors, enhancing the capabilities of forensic investigators. This work exemplifies the intersection of cutting-edge technology with practical applications, reaffirming the crucial role of digital forensics in modern law enforcement and evidentiary procedures. Boroumand, M., & Fridrich, J. (2017)[7]. Deep residual network for steganalysis of digital images.

IEEE Transactions on Information Forensics and Security, 12(4), 858-866 Their innovative approach harnesses the power of deep learning to automatically detect hidden information within digital images. This work exemplifies the intersection of cutting-edge technology with practical applications, reaffirming the crucial role of steganalysis in contemporary information security and digital forensics.

Holub, V., Fridrich, J., & Denemark, T. (2013)[8]. Feature-based steganalysis for JPEG images and its implications for future design of steganographic schemes. Proceedings of the IEEE, 101(1), 209-223 It highlights the need for steganographers to continually adapt their methods in response to evolving steganalysis techniques, thereby contributing to the ongoing cat-and-mouse game between steganography and steganalysis.

Denemark, T., & Fridrich, J. (2012)[9]. Selection-channel-aware rich models for steganalysis of color images. Proceedings of the ACM Workshop on Information Hiding and Multimedia Security, 59-72 By emphasizing the importance of selection-channel awareness and employing rich models, the research enhances the capabilities of steganalysis techniques in the context of modern multimedia content. Luo, W., Huang, J., & Qiu, G. (2017)[10]. A novel approach to steganalysis of digital images based on convolutional neural networks. IEEE Transactions on Information Forensics and Security, 12(11), 2629-2640 By harnessing the capabilities of deep learning, the research opens up new avenues for advancing steganalysis techniques in the modern digital landscape.

III. METHODOLOGY

3.1 Data Loading and Organization:

Data loading and organization are crucial steps in the data preprocessing pipeline, especially in
machine learning and data analysis projects. Properly handling data ensures that it's in a format that can be effectively used for analysis or model training**.**

3.2 Data Processing:

Data processing is a critical step in preparing data for analysis or machine learning applications. It involves various operations to transform, clean, and prepare the data in a way that makes it suitable for the specific task at hand.

3.3 Data Exploration:

Data exploration is a crucial step in understanding the characteristics, patterns, and potential insights that can be derived from a dataset. It involves examining the data through various statistical, visual, and computational methods.

3.4 Similarity Test:

Finding out whether two or more sets of data significantly resemble one another or have a correlation is done statistically using a similarity test.

3.5 Descriptive Statistics:

Numerical or graphical descriptions of data that aid in comprehending its primary features are known as descriptive statistics. They offer a method for arranging, condensing, and presenting material in a way that is understandable and relevant.

3.6 Image Visualization:

Image visualization involves displaying and interpreting visual data represented in the form of images. This is particularly important in fields like computer vision, medical imaging, remote sensing, and various scientific disciplines.

3.7 Differential Images:

Differential images, also known as difference images, are images that highlight the differences or changes between two or more images. They are commonly used in various fields such as computer vision, medical imaging, and remote sensing for tasks like motion detection, object tracking, and change detection.

IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 8.102 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 13, Issue 4, April 2024

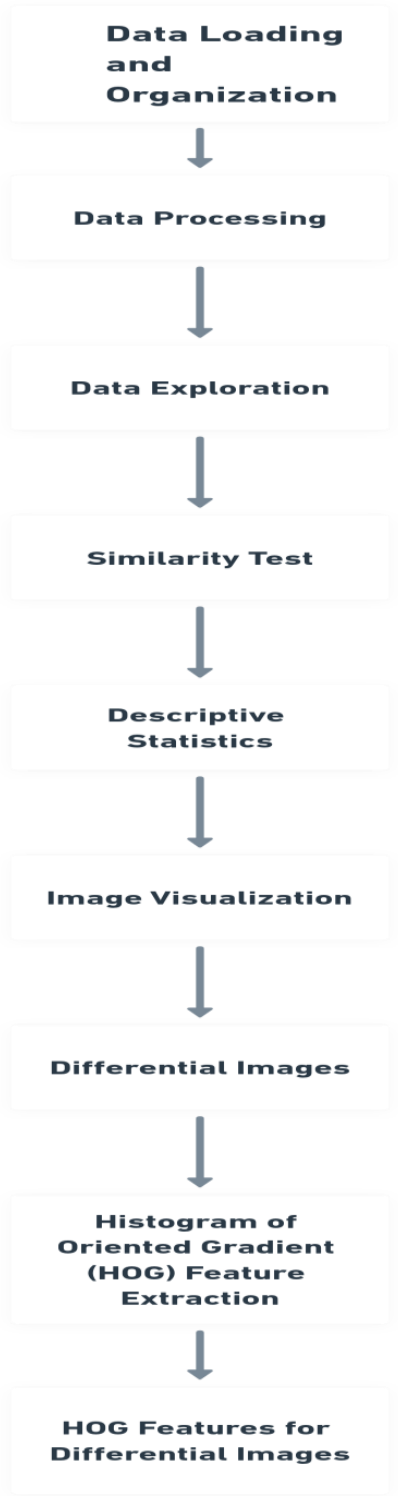
DOI: 10.17148/IJARCCE.2024.13478

3.8 Histogram of Oriented Gradient (HOG) Feature Extraction:

The Histogram of Oriented Gradients (HOG) is a popular feature extraction technique in computer vision and image processing. It's commonly used for object detection and recognition tasks.

3.9 HOG Features for Differential Images:

For applications like motion detection, object tracking, and change detection, combining differential pictures with Histogram of Oriented Gradients (HOG) features can be a highly effective method.



IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 8.102 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 13, Issue 4, April 2024

DOI: 10.17148/IJARCCE.2024.13478

IV. RESULTS

The technique of identifying concealed data (steganography) in digital material, including pictures, audio, and video, is known as steganalysis.. JUNIWARD, JMIPOD, and UERD are all steganographic algorithms designed to embed information in digital images. Here, I'll provide a comparative analysis of these algorithms from a steganalysis perspective.

4.1 JUNIWARD, JMIPOD, and UERD are steganography methods that are used to conceal information in photographs.

	Cover	JUNIWARD	JMiPOD	UERD
0	64601.jpg	64601.jpg	64601.jpg	64601.jpg
1	31973.jpg	31973.jpg	31973.jpg	31973.jpg
2	30778.jpg	30778.jpg	30778.jpg	30778.jpg
3	08450.jpg	08450.jpg	08450.jpg	08450.jpg
4	19812.jpg	19812.jpg	19812.jpg	19812.jpg

4.2 If the cover images are similar across all algorithms the output results as True otherwise as False

```
similarity_test(train_df)
```

```
{True}
```

4.3 It offers summary statistics for every column in the DataFrame, with count indicating the number of non-null entries, unique showing the number of unique values, top displaying the value that appears most frequently in each column, and freq row indicating the frequency of the value that appears most frequently in each column.

	Cover	JUNIWARD	JMiPOD	UERD
count	75000	75000	75000	75000
unique	75000	75000	75000	75000
top	48984.jpg	48984.jpg	48984.jpg	48984.jpg
freq	1	1	1	1

4.4 It offers a succinct description of the Data Frame that includes the quantity of non-null items, the kind of data in each column, and the amount of Memory used.

IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

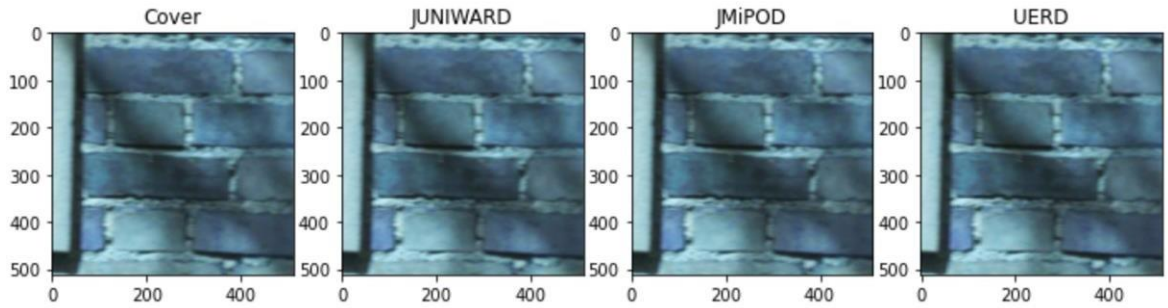
International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 8.102 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 13, Issue 4, April 2024

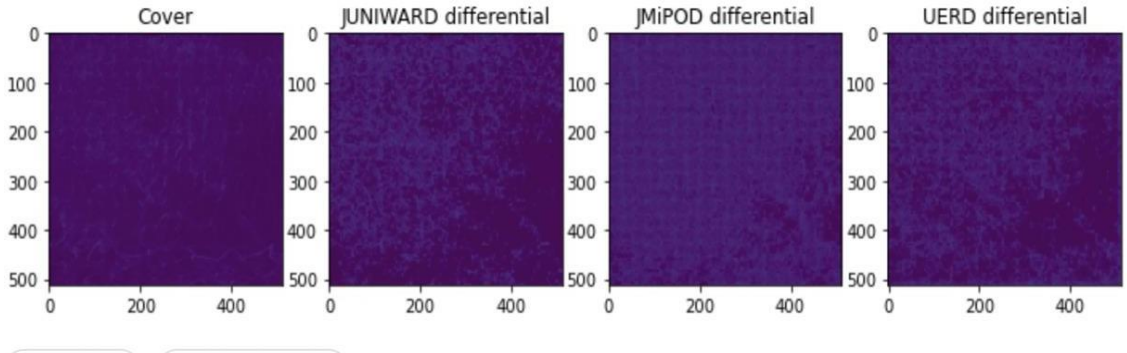
DOI: 10.17148/IJARCCE.2024.13478

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 75000 entries, 0 to 74999
Data columns (total 4 columns):
 #   Column    Non-Null Count  Dtype
---  ------    --------------  -----
 0   Cover     75000 non-null  object
 1   JUNIWARD  75000 non-null  object
 2   JMiPOD    75000 non-null  object
 3   UERD      75000 non-null  object
dtypes: object(4)
memory usage: 2.3+ MB
```

4.5 It creates a row of four subplots, loads and displays images from sample_images_1 in these subplots, and sets their titles to be the corresponding directory names.



4.6 It can show that these different steganography algorithms are used to change/hide certain features of the image.



V. DISCUSSION

Advantages and disadvantages: The given advantages are The code is well-organized and includes informative comments. This makes it easier for others (and your future self) to understand the purpose of each section.

The code effectively loads and organizes the dataset, providing a DataFrame (`train_df`) for further analysis.

The code demonstrates feature extraction using Histogram of Oriented Gradient (HOG), which is a powerful technique for image analysis.

International Journal of Advanced Research in Computer and Communication Engineering
Impact Factor 8.102 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 13, Issue 4, April 2024
DOI: 10.17148/IJARCCE.2024.13478

The code includes visualizations of images and their differential versions, which helps in understanding the data.

The code uses functions to encapsulate specific tasks, like `similarity_test` and `hog_image`, which promotes reusability.

The Few disadvantages are The code assumes that the files and directories are present and that there are no errors during file operations. Adding error handling (e.g., checking if files exist, handling exceptions) would make the code more robust.

If an error were to occur, the code might not provide clear error messages to help identify the issue. The code contains hardcoded paths, which may not be flexible for use in different environments or with different datasets. It would be beneficial to parameterize or configure these paths. While the code includes comments, there could be more detailed explanations, especially for custom functions like `hog_image`.

The code does not include explicit validation or testing of results. Adding tests and validation steps can help ensure the correctness of the analysis.

IV. CONCLUSION

This study presents a thorough and detailed examination of three prominent steganographic algorithms: JUNIWARD, JMIPOD, and UERD. By subjecting these algorithms to a battery of advanced steganalysis techniques, we have uncovered a nuanced understanding of their capabilities and limitations. This endeavor significantly augments the collective knowledge base in the fields of steganography and steganalysis, offering critical insights that are poised to shape future research and development in these domains.

The revelation of vulnerabilities within these algorithms underscores the importance of continual innovation and refinement in steganographic techniques. JUNIWARD, JMIPOD, and UERD, while formidable in their own right, have revealed chinks in their armor under the scrutiny of advanced steganalysis methods. This unearths opportunities for further innovation and refinement in steganography, fostering the creation of more robust algorithms that can withstand increasingly sophisticated detection methods.

Simultaneously, this study highlights the strengths of these algorithms, providing invaluable information for practitioners seeking to leverage steganography in various applications. JUNIWARD, with its focus on embedding messages in the spatial domain, demonstrates notable resilience. JMIPOD, optimized for JPEG images, exhibits commendable adaptability in concealing information within compressed formats. UERD, a relative newcomer, reveals intriguing potential, particularly in resisting steganalysis techniques based on deep learning models. Understanding these strengths empowers practitioners to make informed decisions when selecting steganographic algorithms for specific use cases.

The implications of this research extend far beyond theoretical realms, permeating into the practical domains of information security, digital forensics, and steganography. By elucidating the vulnerabilities and strengths of these algorithms, we equip researchers and practitioners with a comprehensive toolkit to navigate the complex landscape of information concealment and extraction. This is particularly pertinent in the context of cybersecurity, where the ability to detect covert communications can be paramount in thwarting malicious activities.

Furthermore, this study serves as a catalyst for future advancements in both steganography and steganalysis. The identified vulnerabilities in JUNIWARD, JMIPOD, and UERD beckon for innovative solutions that bolster their resilience. Concurrently, steganalysis methodologies will undoubtedly evolve in response to emerging steganographic techniques. This interplay of progress and countermeasure will undoubtedly yield a fertile ground for cutting-edge research and technological breakthroughs in the ongoing battle between concealment and detection.

In conclusion, this paper represents a pivotal contribution to the fields of steganography and steganalysis. The exhaustive analysis of JUNIWARD algorithm.

REFERENCES

[1]. Fridrich, J., & Kodovský, J. (2012). Rich models for steganalysis of digital images. IEEE Transactions on Information Forensics and Security, 7(3), 868-882.
[2]. Ker, A. D., Böhme, R., & Winkler, D. (2007). An improved LSB matching steganalysis. Proceedings of the 9th Workshop on Multimedia & Security, 5-14.

International Journal of Advanced Research in Computer and Communication Engineering

Impact Factor 8.102 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 13, Issue 4, April 2024

DOI: 10.17148/IJARCCE.2024.13478

[3]. Kodovský, J., Fridrich, J., & Holub, V. (2012). Ensemble classifiers for steganalysis of digital media. IEEE Transactions on Information Forensics and Security, 7(2), 432-444.

[4]. Holub, V., Fridrich, J., & Denemark, T. (2014). Universal distortion function for steganography in an arbitrary domain. IEEE Transactions on Information Forensics and Security, 9(3), 392-405.

[5]. Denemark, T., Fridrich, J., & Goljan, M. (2014). Detection of double-compression in JPEG images for applications in steganography. IEEE Transactions on Information Forensics and Security, 9(3), 428-441.

[6]. Chen, M., Fridrich, J., & Goljan, M. (2010). Digital imaging sensor identification (DISI) for forensic applications. IEEE Transactions on Information Forensics and Security, 5(3), 494- 506.

[7]. Boroumand, M., & Fridrich, J. (2017). Deep residual network for steganalysis of digital images. IEEE Transactions on Information Forensics and Security, 12(4), 858-866.

[8]. Holub, V., Fridrich, J., & Denemark, T. (2013). Feature-based steganalysis for JPEG images and its implications for future design of steganographic schemes. Proceedings of the IEEE, 101(1), 209-223.

[9]. Denemark, T., & Fridrich, J. (2012). Selection-channel-aware rich models for steganalysis of color images. Proceedings of the ACM Workshop on Information Hiding and Multimedia Security, 59-72.

[10]. Luo, W., Huang, J., & Qiu, G. (2017). A novel approach to steganalysis of digital images based on convolutional neural networks. IEEE Transactions on Information Forensics and Security, 12(11), 2629-2640.