DEVELOPMENT OF DATA HIDING METHODS FOR CONTENT PROTECTION IN MEDICAL DOMAIN

A Thesis Submitted
in Partial Fulfillment of the Requirements for the Degree of

DOCTOR OF PHILOSOPHY
by

SONAL GANDHI
(2K21/PHDCO/02)

Under the supervision of

Dr. Rajeev Kumar

(Supervisor)

Delhi Technological University Delhi

Department of Computer Science and Engineering
DELHI TECHNOLOGICAL UNIVERSITY
(Formerly Delhi College of Engineering)
Bawana Road, Delhi 110042

DECEMBER, 2025

CANDIDATE DECLARATION

I, Sonal Gandhi (2K21/PHDCO/02), hereby declare that the thesis titled **DEVELOPMENT OF DATA HIDING METHODS FOR CONTENT PROTECTION IN MEDICAL DOMAIN**, submitted to **Delhi Technological University, Delhi**, in partial fulfillment of the requirements for the award of the degree of Doctor of Philosophy in the Department of Computer Science and Engineering, is my original work.

This research work has been completed under the supervision of **Dr. Rajeev Kumar** (Supervisor), Department of Computer Science and Engineering, Delhi Technological University, Delhi, India.

The explanations presented in this thesis are based on my understanding and comprehension of the original texts. I have not submitted this work to any other institution for the award of any degree, diploma, associate-ship, fellowship, or other title or honor.

Sonal Gandhi
2K21/PHDCO/02
Computer Science and Engineering
Delhi Technological University
(Formerly Delhi College of Engineering)
Delhi-110042, India
Place: Delhi
Date : December 12, 2025

CERTIFICATE

Certified that **Sonal Gandhi (2K21/PHDCO/02)** has carried out her research work presented in this thesis titled **DEVELOPMENT OF DATA HIDING METHODS FOR CONTENT PROTECTION IN MEDICAL DOMAIN** under my supervision. The research work was undertaken for the award of the degree of **Doctor of Philosophy** from Delhi Technological University, New Delhi, India.

The thesis embodies the results of original research and studies conducted by the student herself. The contents of the thesis do not form the basis for the award of any other degree to the candidate or to anybody else from this University or any other University/Institution.

Dr. Rajeev Kumar
Assistant Professor
Computer Science and Engineering
Delhi Technological University
Delhi, India
Place: Delhi
Date: December 12, 2025

ACKNOWLEDGEMENT

Vakratunda Mahakaya, Surya Koti Sama Prabha |
Nirvighnaṃ Kuru Me Deva, Sarva Karyeṣhu Sarvada ||

First and above all, I thank God Almighty for giving me the strength, wisdom, and persistence needed throughout this academic journey. The faith in a higher power has been my anchor, offering both comfort and inspiration. This belief has not only sustained me through the challenging moments but also heightened the joy of my successes, reminding me that every step forward is part of a larger journey.

Most importantly, my heartfelt appreciation goes to my supervisor, **Dr. Rajeev Kumar**. His exceptional mentorship, profound expertise, and steadfast belief in my potential have been the cornerstone of my research journey. His guidance has been more than just academic advice; it has been a source of constant motivation, encouraging me to explore uncharted territories and embrace challenges with resilience. His detailed feedback and constructive critiques have been invaluable, pushing me to refine my work and strive for excellence. His perspectives have not only enhanced the quality of my research but have also broadened my own understanding and approach to problem-solving.

My thanks also go to the members of the Student Comprehensive Report Committee and Departmental Research Committee for their valuable advice and thought-provoking suggestions. Their contributions have been crucial in making my research stronger and more focused.

I am thankful for my colleagues and friends at the Department of Computer Science, Delhi Technological University. Their friendship, insightful discussions, and readiness to share your knowledge have been a great help throughout this journey. Their support has been a pillar for me during both good and tough times.

I owe my deepest gratitude to my family, whose love, patience, and unwavering support have been the foundation of my journey. To my parents, thank you for instilling in me the values of hard work, humility, and perseverance. Your constant prayers and encouragement have been my guiding light. A special note of heartfelt appreciation to my spouse, who stood by me through every challenge and triumph with immense understanding and care. My profound love to my

daughter, **Mahira**, whose innocent laughter and presence have been a source of boundless joy, strength, and motivation. You have all been my pillars, and this accomplishment is as much yours as it is mine.

I am eternally grateful to the divine force that has guided me with strength and wisdom on this academic path. In moments of doubt and success, my faith in a greater purpose has offered comfort and motivation, helping me to tackle the complex challenges of this research.

This thesis stands as a testament to the collective efforts and support of these remarkable individuals. Their contributions have shaped not only my academic growth but also my personal development, making this journey truly enriching and unforgettable.

Place: Delhi Sonal Gandhi

Date: December 12, 2025

ABSTRACT

Reversible Data Hiding (RDH) has emerged as a vital research area within information security, driven by the growing demand for secure communication that preserves data confidentiality while ensuring perfect recovery of the original cover image. Recognizing its significance across sensitive domains such as medical imaging, cloud storage, and digital forensics, this thesis presents an extensive investigation into contemporary RDH techniques. Building upon insights derived from a comprehensive review of state-of-the-art methods, the work identifies key research gaps and proposes novel methodologies aimed at enhancing embedding efficiency, prediction accuracy, and visual fidelity in reversible data hiding.

The thesis begins with an in-depth exploration of RDH by exploring major techniques and advancements within the field in the last decade and so. It systematically analyzes the progression of RDH methodologies, highlighting key contributions, emerging trends, and research frontiers identified through an extensive survey of scholarly works indexed in reputable databases such as the Web of Science (WoS). This review not only consolidates the existing body of knowledge but also identifies prevailing challenges, unresolved research issues, and potential avenues for future investigation in the domain of RDH.

In the second part of the research, a high-capacity RDH method integrating contrast enhancement and brightness preservation is proposed for medical images. The method divides the image into Region of Interest (ROI) and Non-Region of Interest (NROI), applying region-specific embedding strategies that align with their distinct characteristics. A novel preprocessing technique is further introduced for the ROI, efficiently handling areas with low Pixel Concentration Ratio (PCR) to generate additional vacant bins for embedding. This approach significantly enhances embedding capacity while improving contrast and visual quality without compromising the inherent brightness of the medical image.

The third contribution introduces an advanced RDH method for color images employing a Convolutional Neural Network Convolutional neural network (CNN)–based predictor. To address the limited dependency range in conventional RDH methods, a Self-Attention CNN (SA-CNN) predictor is designed to capture both local and global spatial dependencies. A novel error adjustment mechanism further leverages the inter-channel correlation among RGB components,

resulting in superior reconstruction quality and increased embedding performance.

The fourth part of the work presents a Two-Stage Interpolation-Based Reversible Data Hiding Framework that integrates attention-driven prediction to enhance embedding efficiency. To overcome the limitations of conventional interpolation techniques, the proposed approach combines bicubic interpolation with a highly accurate deep-learning predictor. A novel multi-head attention–based U-Net model, termed UMANet, is introduced to capture a broader spatial context, yielding improved interpolation accuracy and embedding performance.

The fifth contribution focuses on Reversible Data Hiding in Encrypted Images (RDHEI), adding an additional layer of security over traditional RDH. A novel SCAM-Net predictor is proposed, equipped with a multiscale extraction module for capturing both fine and coarse feature details. The extracted features are refined using a Convolutional Block Attention Module (CBAM) that leverages channel and spatial attention mechanisms. This architecture achieves highly precise image prediction and outperforms existing state-of-the-art methods, leading to superior embedding performance and robust data concealment in encrypted environments.

Collectively, the research contributes a comprehensive suite of RDH methodologies that advance the state of the art in prediction accuracy, embedding efficiency, and visual fidelity, reinforcing the role of intelligent predictive modeling in secure image-based communication and storage.

LIST OF ABBREVIATIONS

AMBTC Absolute Moment Block Truncation Coding

AMP Adaptive mean predictor

ACERDH Automatic contrast enhancement method with reversible data hiding

ATD Adaptive Threshold Detector

BIP Bilinear interpolation

bpp Bits per pixel

CA Channel Attention

CBAM Convolutional Block Attention Module

CNN Convolutional neural network

DE Difference Expansion

DH Data Hiding

DP Difference predictor

EC Embedding Capacity

EHRs Electronic health records

ER Embedding Rate

GAP Gradient Adaptive Predictor

HE Histogram equalization

HS Histogram Shifting

LSB Least Significant Bit

LSBs Least Significant Bits

MED Median Edge Detector

MHA multi-head attention

MSE Mean Squared Error

MSB Most Significant Bit

MSBs Most Significant Bits

NMI Neighbor Mean Interpolation

NROI Non-Region of Interest NROI

NNI Nearest Neighbor Interpolation

PCR Pixel concentration ratio

PE Prediction Error

PEE Prediction Error Expansion

PEs Prediction Errors

PEH Prediction error histogram

PSNR Peak Signal-to-Noise Ratio

PEO Prediction Error Ordering

PVO Pixel Value Ordering

RDH Reversible Data Hiding

ROI Region of Interest

RCE Relative contrast error

RDHCE RDH with contrast enhancement

RDHEI RDH for encrypted images

RMBE Relative mean brightness error

RRBE Reserving Room Before Encryption

RRDH Robust RDH

RP Rhombus predictor

SCAM-Net Spatial-Channel Attention Multi-Scale Network

SOTA State-Of-The-Art

SA-CNN Self attention based CNN

SSIM Structural Similarity Index

TCIA The Cancer Imaging Archive

UMANet Multi-head attention based U-Net

VQ Vector Quantization

VRAE Vacating Room After Encryption

WoS Web of Science

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

CHAPTER 1

INTRODUCTION

Over the past few decades, increased social media usage, advancements in digital technology, and the popularity of cloud applications have led to an exponential increase in the transmission of valuable information and multimedia content over networks. This has increased the requirement of safeguarding sensitive information from intruders, especially in the field of medical and military applications, where privacy and security are paramount. Cryptography and Data Hiding (DH) [5, 6] have evolved as two primary ways to achieve information security; where cryptography converts the information from one form to another form that is unrecognizable by unintended users, and DH solves the issue by hiding the sensitive information into a cover media, keeping the intruders unaware of the information. However, traditional data-hiding methods are irreversible in nature, as these techniques introduce permanent changes to the cover media, and recovering the cover media from the marked image is not feasible.

To overcome the aforementioned challenge, RDH [7] has emerged as a promising technique that enables cover media to be recovered along with the hidden data [8].This dual capability, image reversibility as well as data concealment, enables the RDH to become the optimum choice to address the critical needs of various applications where, along with information security, image preservation is also crucial.

1.1 Necessity of RDH in Medical Images

The healthcare sector has witnessed a remarkable digital transformation over the last decade, fueled by the widespread adoption of Electronic health records (EHRs), telemedicine, and advanced medical imaging technologies. Medical images such as X-rays, CT scans, MRIs, and ultrasounds play a vital role in diagnosis, treatment planning, and patient monitoring. According to a 2023 report by Statista, the global medical imaging market is projected to reach over USD 45 billion by 2030, underscoring the humongous growth in the production, transmission, and storage of digital medical data.

Medical images often contain not only visual data but also embedded metadata, such as patient identity, diagnosis details, and hospital information. With the rise of the worldwide Internet and cloud-based healthcare solutions, medical images are now routinely transmitted over public networks for diagnosis, consultation, and remote analysis. However, this digitization has introduced serious security and privacy challenges. The medical domain faces increasing threats such as data tampering, unauthorized access, identity theft, and ransomware attacks. In 2023, the average cost of a data breach hit USD 4.45 million. One of the recent cyberattacks at UnitedHealth Group revealed that 192.7 million people in the US were affected by this data breach. The information that was hacked includes health insurance details, patient diagnosis, social security numbers, and treatment details. Another report in 2024 by IBM's "Cost of a Data Breach" study revealed that the healthcare industry suffers the highest average data breach cost, approximately USD 10.93 million per incident, emphasizing the vulnerability and sensitivity of medical information.

Any unauthorized modification or leakage of this data could lead to diagnostic errors, patient privacy violations, and even legal consequences. Ensuring data security while preserving image quality is therefore a critical concern. RDH offers a practical solution by enabling confidential information to be embedded within medical images, with the assurance that both the embedded data and the original image can be fully recovered [9].

1.2 RDH Architecture

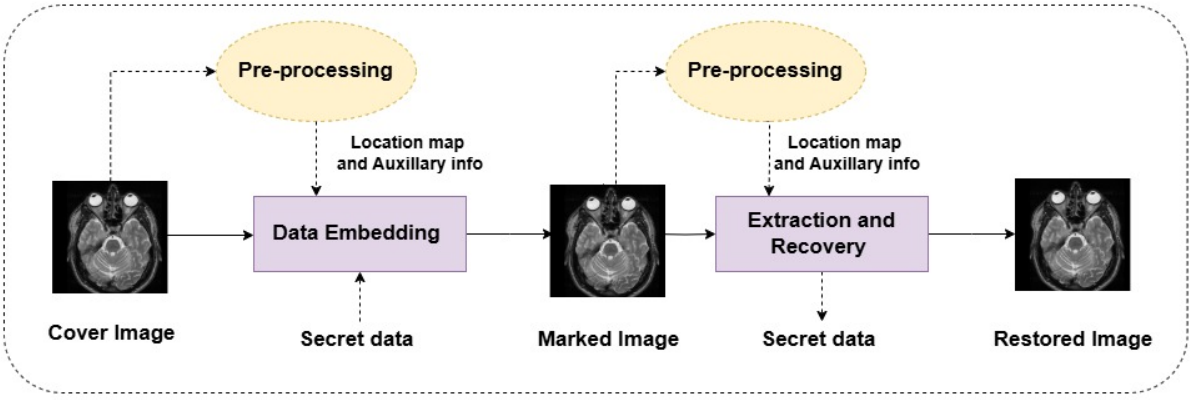


Fig. 1.1: General Framework of RDH

The general framework of RDH [10] is illustrated in Fig. 1.1. In RDH, the cover media, which may be an image, video, or audio file, acts as the carrier for the hidden secret information. The process begins with pre-processing the cover media to generate a location map and auxiliary information required for accurate data extraction. These elements, together with the secret data, are embedded into the cover media in a manner that allows them to be separated during the extraction phase, as shown in the framework. The cover image containing the embedded data is

referred to as the marked image or stego image [9,11]. Using auxiliary information, the receiver can reverse the embedding operation performed by the sender to recover both the original cover media and the concealed secret data.

1.3 Classification of RDH Techniques

Over the past two decades, RDH techniques have been extensively researched, leading to numerous approaches aimed at achieving lossless recovery of both the secret message and the cover media. These approaches can be broadly categorized into four main domains: frequency domain, spatial domain, compressed domain, and Robust RDH (RRDH) techniques. A detailed classification of these domains is illustrated in Fig.1.2. Spatial-domain RDH techniques are preferred for applications requiring high embedding capacity and strict reversibility, such as Medical and forensic areas. Frequency-domain RDH techniques are preferred for applications that require robust data embedding during transmission and processing, such as watermarking. Compression-domain RDH is ideal for storage- and bandwidth-constrained systems, such as cloud and streaming services.

In the **Frequency domain**, RDH techniques utilize well-known transform techniques to first transform the cover image into the frequency domain [12–19]. The frequency coefficients obtained after the transformation step are prudently selected to conceal secret data within them while ensuring the reversibility of the cover image. Based on the technique used for transformation, the RDH techniques can be sub-categorized into multiple categories, such as Discrete Cosine Transform [12, 13], Discrete Wavelet Transform [14, 15], Integer Wavelet Transform [16, 17], and other transforms [18, 19]. Frequency domain RDH techniques are often considered robust and play a crucial role where content security is of utmost importance. However, the transformation step introduces additional complexity and transmission overhead. Also, the Embedding Capacity (EC) of frequency domain RDH techniques is generally limited.

In the **Compression domain**, RDH techniques involve the use of various compression [20–25] methods to generate compressed codes from the cover image, thereby reducing its size. The secret data is embedded within these compressed codes in a manner that minimally affects the original content. Depending on the compression technique applied, RDH methods are classified into lossless and lossy compression techniques. Lempel-Ziv-Welch [20], one of the most widely explored lossless compression algorithms, substitutes recurring data sequences with compact codes to effectively minimize redundancy and achieve efficient compression without data loss. Within the lossy compression RDH techniques, Vector Quantization (VQ) [21,22] and Absolute Moment Block Truncation Coding (AMBTC) [24, 25] are the two most widely used compression techniques explored by the researchers. VQ operates by segmenting each image into blocks and utilizing a pre-designed codebook to index each block, thus compressing the entire image. Thereafter, indexes are utilized to conceal data within them. AMBTC is another popular lossy compression technique that first divides the image into blocks, calculates

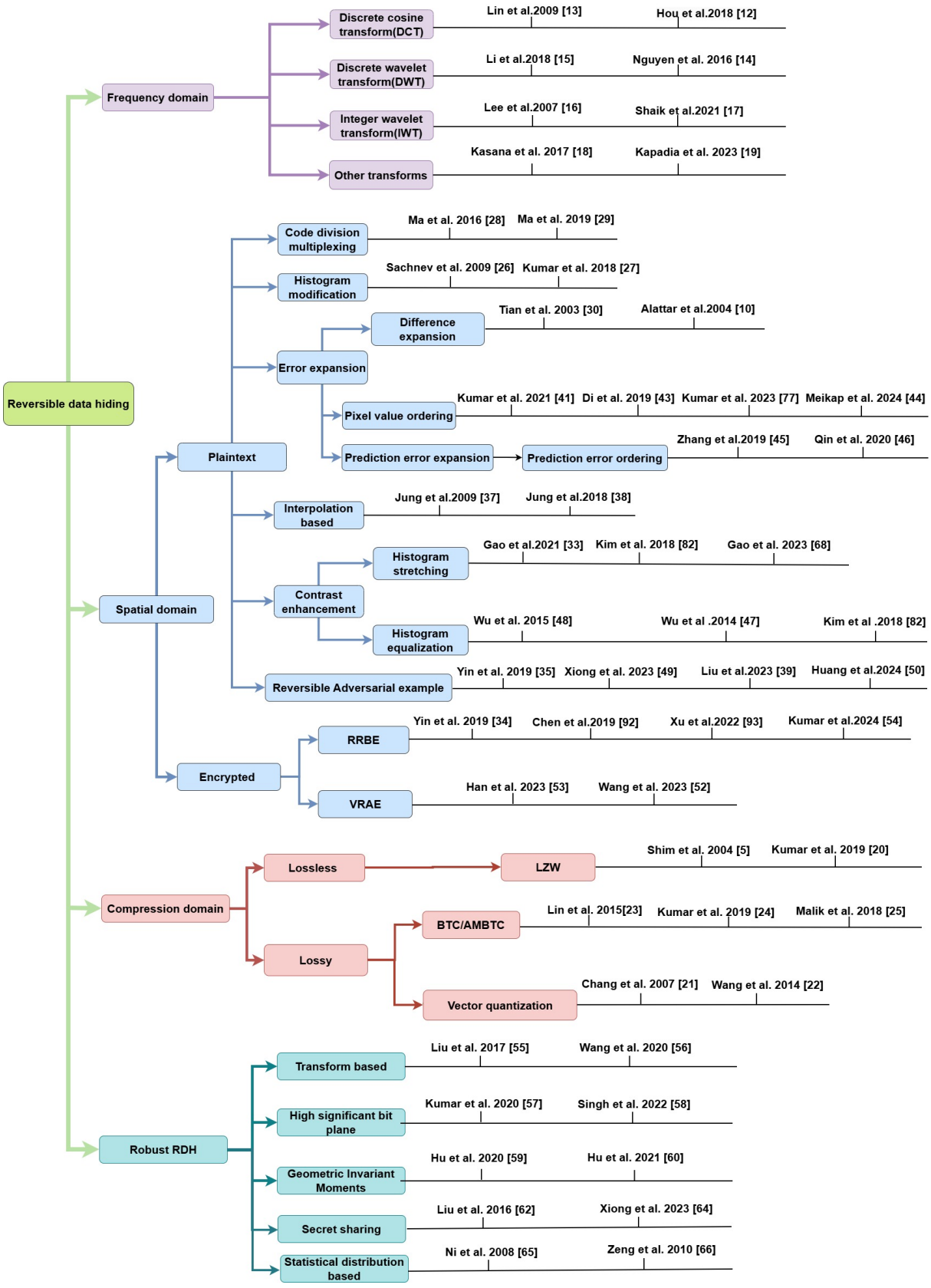


Fig. 1.2: Classification of RDH methods

their statistical moments, and then quantizes the blocks into a reduced number of levels based

on the statistical moments, providing effective compression. Lossy compression techniques introduce some distortion in the images; however, provide enhanced data-hiding capacity while maintaining acceptable image quality.

The spatial domain RDH techniques focus on directly modifying the pixel intensities of the cover image. This direct approach makes them relatively simple and user-friendly. Spatial domain techniques [26, 27] are generally divided into two main categories: plain-text RDH [10, 26–34] and RDHEI [35, 36]. The plain-text RDH techniques are designed to maximize the EC while maintaining minimal distortion in the marked images. Low distortion in the marked image is crucial in RDH, as the aim is to ensure that the embedding changes made to the cover image are imperceptible to unauthorized users. However, a trade-off often exists between EC and distortion. Plain-text RDH techniques can be further categorized based on their design into several methods: code division multiplexing [28, 29], histogram modification [26, 27], error expansion [10, 30–32], interpolation [37, 38], contrast enhancement [33] and reversible adversarial example [34, 39].

In **code division multiplexing** based RDH techniques [28, 29], similar to telecommunication, the secret data is represented by spread sequences that are orthogonal to each other and are generated by Walsh Hadamard matrix. As these sequences are orthogonal to each other, the data embedded in the cover image does not interfere with each other, providing adequate EC and low distortion.

In histogram modification based RDH techniques, the histogram of the pixels is plotted based on their intensities. Then, different strategies are employed to modify the histogram to conceal the secret data in the respective pixels [8]. In general, peak bins, acting as expansion bins, are effectively selected to hide the data. Depending on the direction of the modification, other bins are shifted to accommodate the shift of expansion bins. These techniques generally require auxiliary information to be embedded along with the secret data so as to reverse the Histogram Shifting (HS) operation during data extraction.

Error Expansion is a high-fidelity RDH approach that utilizes redundancies and correlations within cover media to conceal information. It is broadly categorized into three main techniques: Difference Expansion (DE) [30], Pixel Value Ordering (PVO) [40], and Prediction Error Expansion (PEE) [26]. DE techniques exploit the correlation between adjacent pixels by calculating their difference and expanding it for embedding. PVO techniques [41–44] involve ordering pixels by their intensity values before embedding data, ensuring lossless recovery. PEE techniques use a predictor to leverage correlations among multiple neighboring pixels. The effectiveness of PEE in RDH heavily depends on the quality of the predictor: higher prediction accuracy leads to better EC and improved image quality. Thus, accurate predictions enhance embedding efficiency by ensuring that the data is concealed with minimal distortion.

In recent years, various predictors have been developed to enhance the performance of PEE-RDH. Some notable examples include the Difference predictor (DP) [30], Median Edge Detector (MED) [31], Gradient Adaptive Predictor (GAP) [32], Rhombus predictor (RP) [26] and

Comprehensive RP [40], all aimed at improving prediction accuracy. Despite these advancements, traditional predictors often rely on a limited number of context pixels and may struggle with capturing non-linear relationships, which can limit their effectiveness. To address these issues, recent developments have integrated RDH with deep learning networks, leveraging larger context pixels and complex non-linear relationships to enhance prediction performance. To further boost PEE-RDH performance, a notable advancement is the Prediction Error Ordering (PEO) method [45, 46], which draws inspiration from PVO techniques. This approach enhances PEE performance by sorting Prediction Errors (PEs) by magnitude before embedding, thereby prioritizing those errors that minimize distortions in the final embedded image.

Interpolation-based RDH [37, 38] is an important technique in digital image processing, particularly relevant in the healthcare domain, where it enables the upsampling of the cover image while embedding secret information, all while preserving the ability to recover the original image. In RDH, various interpolation techniques are utilized to up-sample the image, preserving image quality while creating space for data insertion. Common methods include Nearest Neighbor Interpolation (NNI) [37], Neighbor Mean Interpolation (NMI) [38], Bilinear Interpolation, and Bicubic Interpolation. Each method offers different trade-offs in terms of computational complexity, processing time, and image quality, providing flexible options for RDH applications.

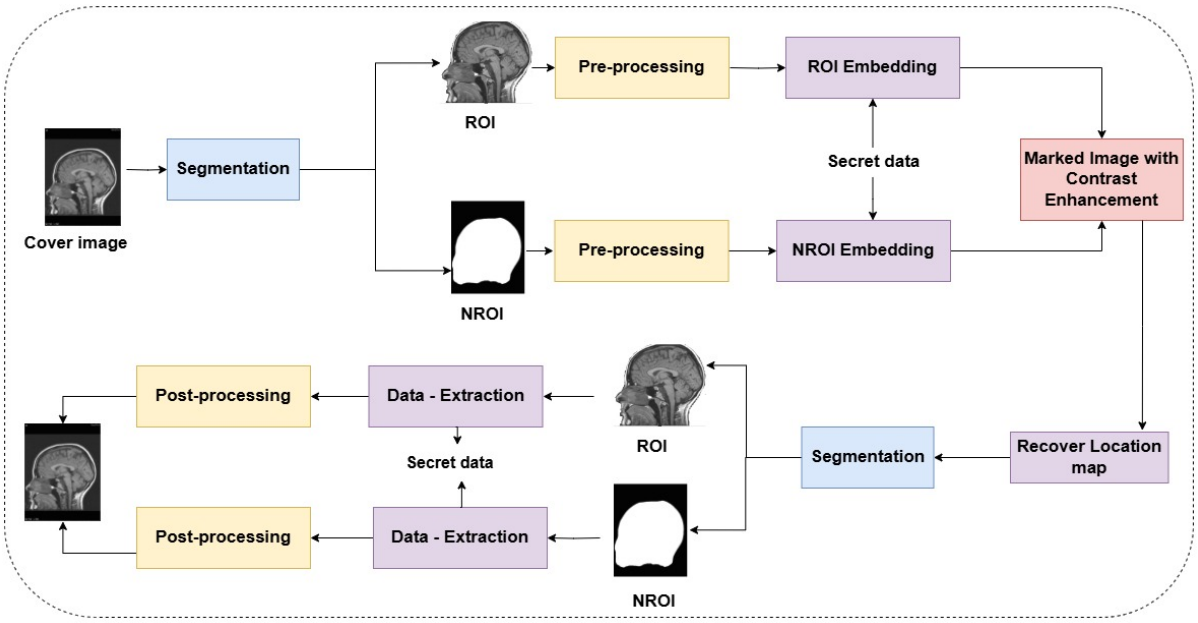
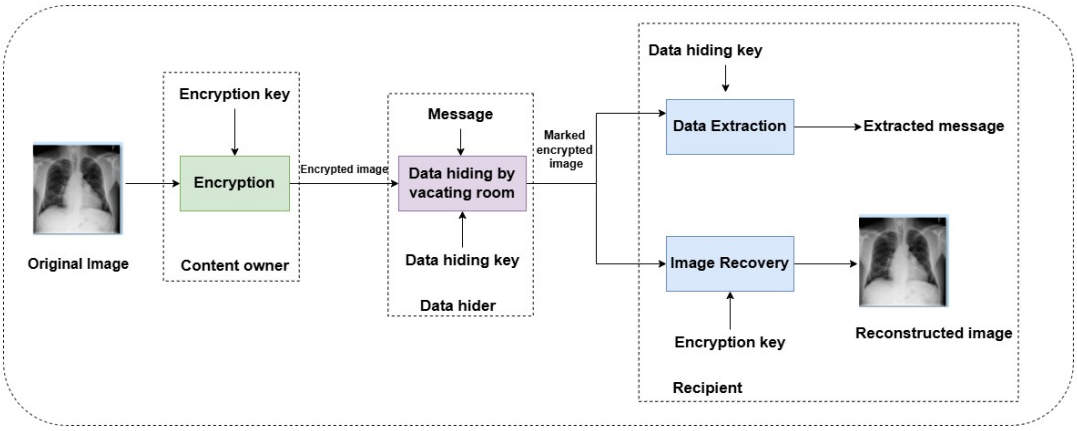


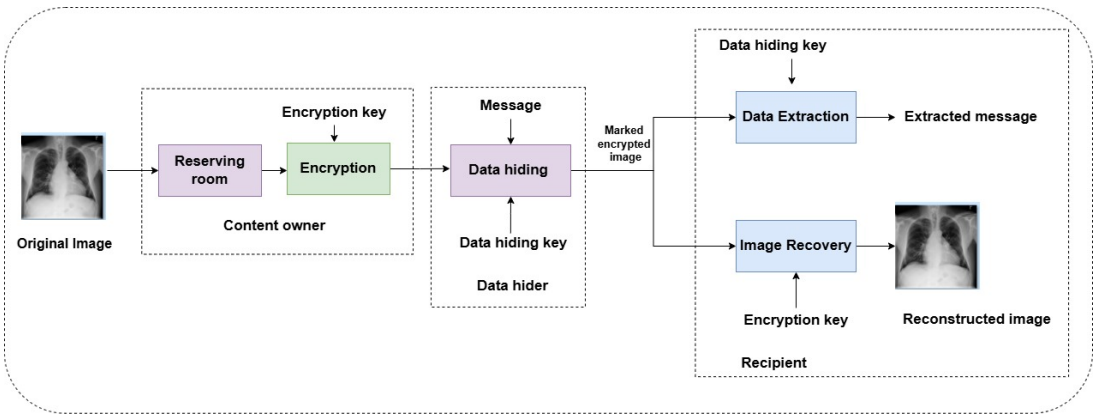
Fig. 1.3: Example Framework of RDHCE in medical images

RDHCE [47–49]techniques aim to enhance the contrast of the marked images while embedding data, which is particularly valuable in scenarios where the original cover images suffer from poor contrast due to limitations in sensing systems or environmental conditions. This technique is especially beneficial in medical imaging, where improved contrast can aid diagnostic analysis. RDHCE methods primarily employ two techniques to achieve contrast enhancement:

Histogram equalization (HE) [47] and histogram stretching [33]. HE adjusts pixel values to distribute them across the entire range of gray scales, thus enhancing global contrast but potentially introducing distortions. Histogram stretching, on the other hand, focuses on modifying the histogram distribution to cover the full gray-scale range and splits large bins to embed data, providing a more localized contrast enhancement with reduced distortion. Fig. 1.3 illustrates a basic framework for applying RDHCE techniques in medical images. Typically, medical images are divided into two regions: Region of Interest (ROI) and Non-Region of Interest NROI (NROI). The ROI contains critical information necessary for diagnosis, while the NROI comprises background information not crucial for diagnostic purposes. Processing these regions differently allows for efficient embedding with high capacity and improved contrast. Despite their benefits, RDHCE techniques often face challenges in achieving high EC and are vulnerable to steganalysis attacks due to the inherent nature of HS.



(a) VRAE



(b) RRBE

Fig. 1.4: Two different frameworks of RDHEI methods (a) VRAE and (b) RRBE

An emerging application of RDH involves **adversarial examples** [34, 50], which are particularly relevant in the context of privacy concerns raised by technologies such as gender identification and facial recognition. RDH techniques are used to embed adversarial perturbations into images, making them appear unchanged to the human eye while confusing AI classifica-

tion models. This approach helps protect personal data from unauthorized AI scrutiny while ensuring that the original image can still be accurately recovered by authorized users. Recent studies [34, 39, 50, 51] demonstrate how RDH can be effectively used to create reversible adversarial images that safeguard privacy in AI-driven environments.

RDHEI techniques offer enhanced security by encrypting the cover image before embedding data. As depicted in Fig. 1.4, RDHEI techniques fall into two primary categories: VRAE (Vacating Room After Encryption) and RRBE. VRAE techniques, illustrated in Fig. 1.4a [52, 53], involve encrypting the image first and then creating space within the encrypted image to hide the secret data. Conversely, RRBE techniques, shown in Fig. 1.4b [35, 36, 54], reserve space for embedding by exploring pixel correlations before encryption. These techniques can be designed as joint or separable, allowing for flexible security configurations based on who should access the data and image. RDHEI techniques not only provide robust security for both the cover media and the secret data but also offer high EC, making them effective for secure and efficient DH.

The **RRDH** techniques aim to provide resilience against distortions or attacks, both intentional and unintentional, such as compression, filtering, or noise addition. These techniques can be broadly categorized into five types: transform-based RRDH [55, 56], (HSB)-based RRDH [57,58], geometric invariant moments-based RRDH [59–61], secret sharing-based RRDH [62–64], and statistical distribution-based RRDH [65, 66].

Transform-based techniques leverage transform coefficients, such as those derived from the Krawtchouk [55] or Haar wavelet [56] transforms, for embedding secret data. These coefficients are less susceptible to environmental changes or compression, ensuring robustness. HSB-RRDH techniques [57] embed data in higher significant bits of pixels, which are less prone to alterations compared to lower significant bits, further enhancing resilience. Geometric invariant RRDH approaches achieve robustness against geometric deformations by utilizing invariant moments, such as Zernike moments [61], for embedding. Secret sharing-based RRDH techniques enhance robustness by dividing content into multiple shares (n), distributed across different locations. With a threshold ($t < n$), the original content can be fully reconstructed from just t shares, ensuring lossless recovery even if some shares are lost or damaged [64]. Lastly, statistical distribution-based RRDH techniques explore statistical properties of the image, such as histogram and PEs, and embed the data in a way that the content remains robust against compression and other processing operations on the image.

1.4 Applications of RDH

RDH is significant across various fields due to its capability to securely embed and later recover data without altering the original content. Key application areas include watermarking [10], medical imaging [49], cloud storage [52, 54], adversarial examples [39], and covert communication [67].

- **Copyright Protection**: One of the primary uses of RDH in watermarking is copyright protection. By embedding watermarks into digital media, such as images, videos, or documents, creators can assert ownership and protect their intellectual property from unauthorized use. RDH ensures that these watermarks can be extracted and the original media fully restored, maintaining the integrity of the content. Multiple industries are focusing on implementing copyright protection, for example, Digimarc Corporation in its patent has studied embedding of auxiliary data in multimedia content.

- **Medical Imaging**: With the rise in digital medical imaging, securing patient data has become crucial. RDH allows for the embedding of confidential patient information within medical images [68] while preserving their diagnostic integrity. This ensures that sensitive data remains protected without compromising the quality of the images. Some of the famous hospitals are currently supporting studies for RDH to embed patient metadata into CT and MRI images, during telemedicine communication. This will help to achieve dual purpose, keep the patient's confidentiality and avoid the patient mismatches during the online consultations.

- **Cloud Storage**: In the digital age, the demand for cloud storage has significantly increased, driven by the need for widespread access and real-time data retrieval.RDH is particularly useful in this context, as it allows essential management data—such as file names and user details—to be embedded into encrypted content without increasing storage costs. This approach ensures that both the original data and the embedded management information can be fully recovered, preserving data integrity and optimizing storage efficiency. RDHEI techniques, an advancement of RDH, are specifically designed for encrypted data, providing a robust solution for secure and efficient cloud storage management. Cloud services providers, such as Aliyun have conducted several case studies to investigate the performance of RDH to ensure integrity and authenticity of the images and content on the cloud storage. The study observed that the RDH preserves the embedded information as well as the cover image even after using recompression.

- **Adversarial Examples**: In the digital age, online services often use artificial intelligence for tasks like facial recognition, which can raise privacy concerns. RDH offers a solution by embedding adversarial perturbations into images, making them appear unchanged while misleading unauthorized artificial intelligence systems. RDH ensures that authorized users can still recover the original image, thus balancing privacy protection with usability.

- **Covert Communication**: RDH is widely used in covert communication, particularly in military applications. By embedding secret data within cover media, RDH makes the hidden information virtually undetectable through visual inspection. Additionally, RDH

can authenticate cover media to ensure it has not been tampered with, further enhancing the security of confidential communications.

1.5 Performance Evaluation Metrics

Evaluation metrics play a crucial role in assessing the effectiveness of a method and validating its performance against other techniques. In the context of RDH, the evaluation is primarily based on two key parameters: **EC** and **visual quality**. Each of these parameters is measured using distinct quantitative metrics, as described below:

1.5.1 Embedding Capacity

EC refers to the maximum amount of secret data that can be embedded into a cover image without violating the reversibility constraint. It is typically expressed in bpp and is calculated using the following formula:

$$\text{Embedding Capacity (bpp)} = \frac{\text{Total number of embedded bits}}{\text{Total number of pixels in the cover image}} \tag{1.1}$$

1.5.2 Visual Quality

Visual quality evaluates the level of distortion introduced in the cover image due to the embedding and extraction processes. An effective RDH technique should minimize this distortion, ensuring that the visual or diagnostic utility of the image is preserved. Common metrics used to assess visual quality include:

1.5.2.1 Peak Signal-to-Noise Ratio (PSNR): PSNR is widely used to quantify the distortion between the original and the marked image. For an 8-bit grayscale image, it is computed using the following equation:

$$\text{PSNR} = 10 \log_{10} \left(\frac{255^2}{\text{MSE}} \right) \tag{1.2}$$

where Mean Squared Error (MSE) is the Mean Squared Error between the original and the marked images. PSNR is measured in decibels (dB), and higher values indicate better image quality with less distortion.

1.5.2.2 (SSIM): The Structural Similarity Index (SSIM) metric is used to measure the similarity between the cover image and the stego image. The range of SSIM is [0,1]. For better performance, its value should be closer to 1. It can be calculated as follows:

$$SSIM = \left(\frac{2\mu_I \mu_{I'} + C_1}{\mu_I^2 + \mu_{I'}^2 + C_1} \right) \left(\frac{2\sigma_{II'} + C_2}{\sigma_I^2 + \sigma_{I'}^2 + C_2} \right) \tag{1.3}$$

where μ_I is the average of cover image I, μ'_I is the average of marked image I',$\sigma_{II'}$ is the covariance of cover image and marked image, σ_I^2 and σ_I^2 are the variances of cover image and marked image, respectively, C_1 and C_2 are the regularization constants.

1.5.2.3 Relative Contrast Error(RCE): Relative contrast error (RCE) is a metric used to evaluate contrast effectiveness. If the value of RCE is less than 0.5, it means there is no enhancement. If the value is more than 0.5, it means the contrast of an image is significantly improved in comparison to the original image. The RCE indicates the degree of contrast enhancement.

$$RCE = \frac{\sigma_{\text{new}} - \sigma_{\text{original}}}{255} + 0.5 \tag{1.4}$$

where σ_{new} and σ_{original} represent the standard deviations of the enhanced/resultant and original images, respectively.

1.6 Benchmark Datasets

The evolution of RDH highlights the need of a comprehensive and standardized test environment to evaluate the effectiveness and robustness of the various proposed techniques. To accomplish this need, USC-SIPI [69], UCID [70], BOSSBase [71], and BOWS-2 [72] have emerged as some of the popular databases often utilized by researchers. These databases, with huge amount of images, possessing diverse characteristics along with widely used standard sizes, offer an exhaustive testbed for examining the performance of the proposed methods, while ensuring the methods are universal, robust, and widely accepted.

1.6.1 BOSSBase

One of the most popular databases is BOSSBase (Break Our Steganographic System), which is specifically used for evaluation in most of the steganography and steganalysis research. The database consists of a whopping set of 10000 grayscale images, each having a standard size of 512×512 pixels. These images are captured using seven different digital cameras, covering different scenarios, leading to a diverse dataset of images with varying characteristics. This diversity and the absence of compression make BOSSBase an ideal database for evaluating both performance and robustness of RDH methods.

1.6.2 BOWS-2

The BOWS-2 database, best known for its real-life images, is a version-2 collection of 10,000 images, captured by professional photographers with different level of complexity and details, ensuring a diverse collection best suited for research. It is one of the most widely used databases used in research related to digital watermarking and image processing. The images

size 512×512 grayscale images, available in the JPEG format. The dataset is a wide collection of landscapes, portraits, architecture, and still-life scenes, helping researchers test the robustness of RDH algorithms for most of the real-life applications.

1.6.3 UCID

The UCID dataset is another popular dataset, organized into several categories, such as animals, nature, buildings, and textures, enabling the researchers to evaluate the RDH algorithms into desired application areas. The dataset consists of 1338 images, with two set of resolutions, 512×384 to 3264×2448 pixels. In addition to providing diverse characteristic and varied resolution images, the UCID dataset also provides metadata that help the researchers in designing and examining context-aware algorithms, resulting into development of robust and high performing algorithms.

1.6.4 SIPI

USC-SIPI is a comprehensive database of digital images provided by the University of Southern California – Signal and Image Processing Institute (USC-SIPI), known for its diverse range of standard test images used in the field of steganography, compression algorithms, and other digital image processing. It is a database of grayscale and color images categorized into multiple cateogies such as classical, miscellaneous, aerial, textures and sequence. $Baboon$, $Boat$ are categorized as highly textured images, providing researchers with common ground to test their algorithms on complex images. Images such as $Lena$, $Pepper$, and Man are part of classical image set and are most widely used in image processing literature. Sequence category images are popular in research related to motion sensing and aerial images are widely used in research related to remote sensing, and feature extraction. Thus USC-SIPI provides a rich dataset for evaluating and benchmarking performance of RDH, watermarking, and cryptography related algorithms.

The collective use of BOSSBase, UCID, BOWS-2, and USC-SIPI databases ensure comprehensive evaluation of RDH algorithms on diverse characteristics images. The wide popularity of these images provides common ground for benchmarking various algorithms against each other. From the widespread use of these databases, it is evident that these databases play crucial roles in the development of RDH while ensuring their robustness.

1.7 Research Gaps and Problem Statement

The primary objective of the RDH techniques is to achieve a higher EC and low distortion in the image with high security of data. However, there exists a trade-off between the EC and image quality. Larger the data embedded into cover image, there will be high distortion. Over

the past few years, various RDH techniques have been proposed by researchers to improve the performance in terms of EC, image quality and other image properties (such as contrast and brightness). Based on the literature review, following research gaps have been identified:

- It has been observed that in RDHCE methods, the pre-processing of the cover images has not been performed considering the properties of Medical Images.
- In case of high-fidelity RDH methods, the EC has been limited as the image correlation has not been efficiently used.
- Most of the RDH methods are unable to hide significant amount of data in encrypted images as these methods make use of redundancy among neighboring pixels which limits the scope of RDH in privacy-preserving applications.

Problem statement: Considering the research gaps, the major challenge being faced by security researchers is the optimal management of classical trade-off between EC and image quality with newly added dimensions such as image enhancement and security.

1.8 Objectives

The objectives to be achieved in this work are summarized as follows:

- **Objective 1**: To study the existing reversible data hiding techniques with contrast enhancement and develop a novel RDHCE for enhancing the EC.
- **Objective 2**: To develop a new high-fidelity reversible data hiding scheme(s) with optimal utilization of spatial correlation in order to achieve enlarged EC.
- **Objective 3**: To propose a new Prediction Error (PE)-based data hiding scheme for privacy preservation while ensuring lossless restoration of the cover media at the receiving end.

1.9 Contributions of Thesis

The research presented in this thesis makes noteworthy contributions to the field of Reversible Data Hiding RDH by addressing several key gaps and challenges identified in existing literature. Through extensive qualitative evaluation and experimental analysis, this work offers a comprehensive assessment of current RDH techniques, emphasizing their strengths, limitations, and potential areas for enhancement. Such an evaluation provides a worthwhile reference point for researchers, presenting deeper insights into the significance and applicability of various approaches. In addition, the thesis includes a quantitative analysis that explores publication trends,

technology landscape, influential contributors, and collaborative patterns within the RDH domain. The integration of both qualitative and quantitative perspectives offers a multidimensional understanding of the field, helping researchers to explore promising directions for future work.

One of the primary contributions of the thesis is the development of the RDHCE method that focuses on improving the contrast of the stego image obtained during embedding, while preserving the image brightness. The method offers unprecedented improvements in EC, image contrast and brightness, marking a significant advancement in RDH research. In some of the other key contributions of the thesis, the prediction performance of the predictor has been improved by designing deep learning based predictors that involve utilizing a Novel Self-attention-based CNN predictor, Multi-head attention based U-Net (UMANet) Predictor, and a SCAM-Net architecture. By utilizing these novel predictors, the thesis effectively bridges several critical research gaps, notably the ineffective utilization of spatial correlations. By presenting these innovative contributions that mitigate the key challenges, the research contributes to the enhancement and progression of RDH methodologies.

Overall, the thesis introduces multiple RDH techniques involving, RDHCE, RDH for color images, interpolation-based RDH, and RDHEI, typically involving different methodologies. Despite contributing to different methodologies, they collectively contribute to a unified research objective: the development of secure, reversible, and robust data-hiding method, specifically tailored for content protection in medical images. RDHCE focuses on improving contrast of grayscale medical images; another research work extends data hiding in multi-channel color image environment, interpolation-based RDH helps in efficient data hiding in medical images where resolution of original images is low, and finally, RDHEI schemes add additional layer of security making cloud-bases telemedicine networks more secure. Collectively, these techniques constitute a comprehensive framework that support visual enhancement, secure storage and transmission of medical content.

The proposed RDH methods have significant implications for practical applications such as Healthcare and the military. By improving the security, contrast, EC, and robustness of the RDH techniques, this research fosters the advancement of more robust systems for protecting sensitive data. Along with proposing significant advancements, this thesis also highlights the future directions and potential areas for further investigation within RDH. Overall, these contributions mark a significant advancement in the era of RDH, introducing novel techniques that diminish the tradeoff between EC and image quality.

1.10 Thesis Organization

- **Chapter 1. Introduction:** Serving as a foundation, the introductory chapter establishes the groundwork of the research conducted in the thesis. It begins with highlighting the need of security and protection in the era of digital evolution within the healthcare industry. The chapter discusses DH as possible solution addressing the growing need of privacy

and security in various applications. The chapter next discusses conceptual framework of RDH along with its dual capability of content protection and storage saving through lossless recovery of original cover media post data extraction. Subsequently, the chapter delves into the motivation for adopting the RDH as prominent solution for safeguarding the patient data in the healthcare industry, The chapter then briefly touches different application areas of RDH, such as coyright protection, adversarial examples, and covert communication, showcasing its widespread applicability and adaptability. By offering a comprehensive introduction, this chapter lays the foundations stone for forthcoming sections, underscoring the context and value of the research work.

- **Chapter 2. Literature Review:** Literature review is a crucial and integrated part of research highlighting the research area's evolution, research gaps, and unsolved mysteries. This chapter offers a comprehensive overview of RDH by exploring major techniques and advancements within the field. The chapter includes a detailed analysis of existing research trends, key contributions, and emerging hot topics, derived from an extensive review of literature from popular sources such as WoS. Through the extensive literature study, we aimed to provide insights into prevalent research areas, significant achievements, and future directions in RDH.

- **Chapter 3. A high-capacity reversible data hiding with contrast enhancement and brightness preservation for medical images:** Chapter 3 introduces a novel and innovative approach to RDHCE with the aim of improving contrast of medical image while embedding data into it. Addressing the literature gap of state-of-the-art techniques, this chapter presents a method that leverages unique properties of medical images to design effective pre-processing technique. The method segments the medical image into ROI and NROI, and then utilize different embedding technique for the two regions. The proposed method introduces a novel pre-processing technique for ROI region optimally utilizing empty bins. The method further processes the region with low Pixel concentration ratio (PCR) to create additional empty bins required for embedding. The experimental results showcase that the balanced and optimal approach proposed by the method not only results into contrast enhancement but also preserves brightness of the marked images.

- **Chapter 4. Reversible Data Hiding for Color Images using a Novel Self-attention-based CNN predictor and Error Adjustment:** This chapter presents a novel RDH method for color images utilizing a CNN-based predictor. The proposed method introduces a novel self-attention based CNN (SA-CNN) predictor, which enhances the multi-receptive features of CNN to effectively capture long-range dependencies among pixels. Self-attention mechanism allows the model to attend to different spatial locations across the entire input image when making predictions for a particular pixel, improving the prediction accuracy significantly, achieving a large peak of zero PE. The proposed RDH

method is specifically tailored for color images leveraging inter-channel correlations of the color image to improvise the prediction of color channels.

- **Chapter 5. UMANeT: A Two-Stage Interpolation-Based Reversible Data Hiding Framework with Attention-Enhanced Prediction:** This chapter presents a novel Two-Stage Interpolation-Based RDH Technique that integrates bicubic interpolation with a deep learning-based predictor. This two-stage framework improves the quality of the interpolated cover image, leading to more effective DH. The proposed method introduces a novel UMANeT: A Multi-Head Attention-Enhanced Predictor for the second stage of the interpolation. UMANeT is a novel U-Net-inspired architecture that leverages multi-head attention to capture a wider spatial context, resulting in more accurate pixel prediction. Experimental results highlighted in the chapter showcase that the proposed method achieves a higher EC compared to existing state-of-the-art methods, while also preserving high visual quality in both cover and stego images.

- **Chapter 6. SCAM-Net: Spatial-Channel Attention Multi-Scale Network for Reversible Data Hiding in Encrypted Images :** Underscoring the need for stronger protective measures, this chapter introduces a novel RDHEI scheme that utilizes SCAM-Net predictor for generating highly accurate predicted images. The SCAM-Net predictor features a multiscale extraction module that captures both fine and broad details effectively. Additionally, an integrated CBAM module enhances feature refinement by leveraging spatial and channel attention mechanisms. Experimental results reveal that this predictor outperforms existing state-of-the-art alternatives. For the majority of USC-SIPI test images, the proposed RDHEI method achieves markedly improved embedding performance.

- **Chapter 7. Conclusion, Future Scope and Societal Applications:** The chapter discusses the findings and outcome of the research and study performed in RDH. In summary, the proposed research seeks to address the traditional tradeoff between EC and image quality through a comprehensive investigation of prediction techniques. By integrating deep learning based techniques, the study aims to enhance the prediction accuracy, hence resulting in improved embedding performance. The future work will focus on further enhancing the prediction performance by predicting PEs and complexity using deep learning techniques. The proposed RDH methods have societal relevance in secure medical image sharing, forensic authentication, and digital content protection, ensuring privacy and trust in sensitive data communication.

1.11 Research Objective Mapping with Publications

Research objective mapping Table 1.1 intents to demonstrate the mapping between research objectives and publications, highlighting the contribution of this thesis in addressing the research

objectives.

Table 1.1: Research Objectives and Corresponding Publications

Research Objectives	**Publication(s)**
RO1. To study the existing reversible data hiding techniques with contrast enhancement and develop a novel RDHCE for enhancing the embedding capacity.	✓ Sonal Gandhi and Rajeev Kumar. "Survey of reversible data hiding: Statistics, current trends, and future outlook." *Computer Standards and Interfaces, Elsevier*. IF: 4.1. March 25, 2025. **(Published)** ✓ Sonal Gandhi and Rajeev Kumar. "A high-capacity reversible data hiding with contrast enhancement and brightness preservation for medical images." *Multimedia Tools and Applications, Springer*. IF: 3.0. April 02, 2024 **(Published)**
RO2. To develop a new high-fidelity reversible data hiding scheme(s) with optimal utilization of spatial correlation in order to achieve enlarged embedding capacity.	✓ Sonal Gandhi and Rajeev Kumar. "High Fidelity Reversible Data Hiding For Color Images using CNN Predictor and Reference Error." *15th International Conference on Computing, Communication and Networking Technologies (ICCCNT)*, 2024, IIT Mandi, IEEE, SCOPUS. November 4, 2024. **(Published)** ✓ Sonal Gandhi and Rajeev Kumar. "UMANeT: A Two-Stage Interpolation-Based Reversible Data Hiding Framework with Attention-Enhanced Prediction." *Journal of Information Security and Applications*, Elsevier, IF: 3.8. April 2, 2025 **(Published)** ✓Sonal Gandhi and Rajeev Kumar. " Reversible Data Hiding for Color Images using a Novel Self-attention based CNN predictor and Error Adjustment." Submitted to *Signal, Image and Video Processing*, Springer, IF: 2.0. December 2, 2024. **(Minor Revisions Submitted)**

Table 1.1 (Continued): Research Objective Mapping with Publications

Research Objectives	Publication(s)
RO3. To propose a new prediction error-based data hiding scheme for privacy preservation while ensuring lossless restoration of the cover media at the receiving end.	✓ Sonal Gandhi and Rajeev Kumar. "SCAM-Net: Spatial-Channel Attention Multi-Scale Network for Reversible Data Hiding in Encrypted Images." *3rd International Conference on Women Researchers in Electronics and Computing (WREC)*, MIET Jalandhar, Springer, SCOPUS. **(Accepted & Presented)** April 10, 2025.

CHAPTER 2

LITERATURE REVIEW

In the era of increasing digital media storage and transmission over networks, data hiding (more specifically RDH) has evolved as a prominent area of research, mitigating information security risk. To study the evolution of research, highlight its achievements over the years, and provide future prospects, this chapter presents an extensive review of RDH utilizing the dataset extracted from one of the most popular and exhaustive databases, Web of Science. The study aims to first perform quantitative analysis that includes trend analysis, citation analysis, prominent authors and organizations, and geographical coverage, along with qualitative analysis focusing on key research areas and future prospects within RDH. The study further provides a structured view of sub-technologies within RDH, along with the key contributors and their proposed techniques that have led to the evolution of RDH over the years. Next, we provide a comprehensive review of some of the prominent works in each of the sub-technologies of RDH. Finally, several key research directions, identified based on current research trends and early-stage problems and motivations, are discussed. Overall, the proposed study provides valuable insights into the evolution, key milestones, current state, and future prospects of RDH, serving as a guide for the research community

The methodology utilized to perform the survey analysis of RDH research is organized into three key parts: **Data Sources and Retrieval**, detailing the use of the Web of Science database and search strategy; **Data Collection and Pre-Processing**, which covers the refinement and validation of the dataset; and **Analysis Tools and Visualization**, explaining the software tools used, including VOSViewer, CiteSpace, and Rawgraphs, for in-depth analysis and visualization.

A. Data Sources and Retrieval

In this analysis, we have employed WoS, the most popular database comprising a vast range of high-impact research articles published in prominent worldwide journals and conferences, providing comprehensive information about scientific research. The database includes several key

indices, such as Science Citation Index Expanded (SCI-EXPANDED), Social Sciences Citation Index (SSCI), Conference Proceedings Citation Index-Science (CPCI-S), and Conference Proceedings Citation Index-Social Science and Humanities (CPCI-SSH). All of these indices are highly relevant to our analysis of RDH. On May 7, 2025, we performed a specific search query that encompassed the period from 2004 to 2025. The search string chosen to retrieve the extensive dataset is: (reversib* OR lossless) NEAR/3 (hiding OR steganography OR watermark* OR embed*). Upon conducting this search, a whopping total of 2462 research documents were obtained.

B. Data collection and pre-processing

For an efficient and accurate analysis, it is important to sanitize and validate the data before processing. Hence, the pre-processing step was performed by filtering and refining the data to eliminate documents that are irrelevant or unreliable, which includes the removal of 5 retracted articles. Along with the statistical refinement, manual analysis was also performed to ensure all the documents retrieved were in the domain of RDH. Post performing screening and manual analysis steps, a dataset comprising 2370 documents was obtained, ensuring that the proposed study is conducted only using relevant publications.

C. Analysis Tools and Visualization

To obtain in-depth insights from the retrieved data, a set of cultivated tools such as VOSViewer, Rawgraphs, and CiteSpace, along with basic tools, such as MS Excel, were utilized. All the visuals were meticulously elected after performing a substantial number of trials. The details of the primary tools used in the proposed study are as follows:

- Citespace [73] is a powerful automated tool that has the capability to analyse the raw data and produce interactive visual charts. Along with the publication trend analysis, it also helps to analyse the co-citation network.
- VOSViewer [74] is another powerful visualisation tool that generates a variety of visuals for co-authorship networks, citation analysis, keyword analysis, and co-occurrence networks.
- Rawgraphs [75] is an open-source data visualisation tool to simplify the visual coherence of complex data.

We collected a comprehensive dataset of articles related to RDH using this approach, reflecting the development and patterns in this field during the specified period. These findings are detailed in the following sections of this study.

2.1 Trend Analysis

In this section, we analyze the publication structure from various perspectives, such as annual publication trends, year-on-year citation progression, leading research domains, productive countries, and top-notch journals. The analysis is further detailed by highlighting key breakthroughs and influential papers in the field.

2.1.1 Publication Trends and Distribution Patterns

To present the evolution of research in RDH, we have presented the yearly publication count along with percentage distribution across the years 2004–2025 in Fig.2.1. The research in RDH commenced in 2004 [76] with some of the early publications but the growth remained stagnant until 2009. A major breakthrough occurred in 2009 with the work of Sachnev et al. [26], who introduced an efficient RDH technique utilizing rhombus prediction and HStechniques. Post 2009, the RDH saw a gradual increase in interest and research output. 2011 witnessed research advancement in another dimension within the RDH, that enhanced the security by introducing RDH in encrypted images [40, 77]. This new dimension catalyzed two-fold research in RDH leading to a steady rise in publication counts In 2014, Wu et al. [47] introduced another aspect within RDH, RDHCE [78, 79] that focuses on improving the contrast of the image while performing embedding. A significant increase in publications is evident from 2015 onwards, with the number of publications exceeding 100 for the first time. This period highlights the growing interest and rapid development within the RDH field. One of the most important breakthroughs occurred in 2021 with the integration of RDH with emerging technology, deep learning. The use of multi-receptive field capability of CNN [80, 81] significantly increased prediction accuracy, improving embedding performance at large. With 236 publications, the peak of RDH research activity occurred in 2022, indicating increasing interest and widespread recognition of its importance. The year 2025 is still in progress and seems to be on track with the trend observed in 2024. Overall, the graph reflects the evolution of RDH research, punctuated by key breakthroughs and innovations that have shaped the field over the past two decades. Fig. 2.2 depicts the citation trend in RDH from 2004 to 2024. Each data label represents the number of papers cited in a particular year. A substantial rise in citations was witnessed across the span of 20 years (2004-2024), reaching the top of its trajectory in 2022. The total number of citations retrieved by the refined dataset (2013 documents) is 50665. So the average number of citations per document is 25. The consistent upward trend in citations represents the increasing interest of researchers in RDH and proves that the earlier research is being found useful in subsequent studies.

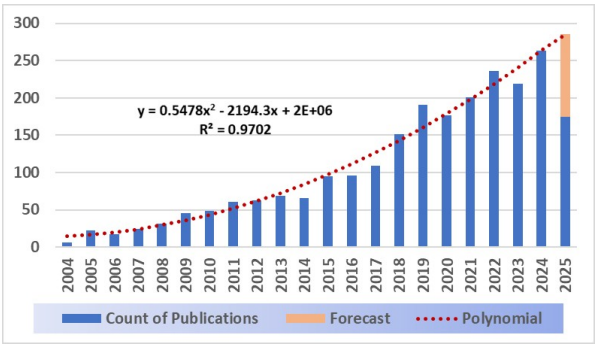


Fig. 2.1: Year wise publication count

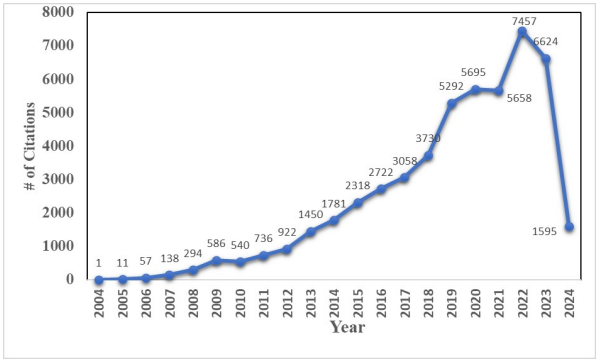


Fig. 2.2: Year-wise total citations received

2.1.2 Geographical Trends

The structure of the graph in Fig.2.3 illustrates the collaborative relationships among the leading 24 countries that contribute to research on RDH through co-authored articles. The size of each

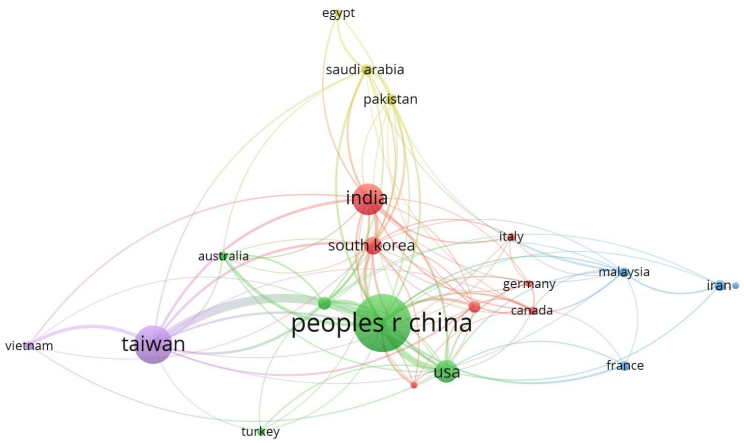


Fig. 2.3: Collaboration network of top countries

node corresponds to the volume of papers produced by each country, with China, the United States, and Canada being the most prominent contributors. The edges represent connections between co-authors, whereas the colors of the nodes reflect different groups of researchers. Based on collaborations, clustering has been performed using the VOSviewer tool. Several groups can be partially explained by the geographical distribution and patterns of research collaboration. Cluster one (green) comprises the People's Republic of China, the United States, Canada, Germany, and Italy, indicating robust collaborative connections among these prominent academic centers. Cluster two, represented by the color red, consists of India and South Korea, demonstrating substantial coordination between both nations. Cluster three, demonstrated through the color purple, consists of Taiwan and Vietnam, indicating the presence of regional cooperation in East Asia. Cluster four, represented by the color light green, includes Australia and Turkey, indicating significant connections between two physically remote nations. Cluster five, reflected

in the color light blue, includes Iran and Malaysia, emphasizing relationships and cooperation within these geographical areas. Cluster six, referred to by the color yellow, comprises Pakistan, Saudi Arabia, and Egypt, suggesting a notable level of scientific collaboration among these nations. Within this group are six distinct clusters, which are partly driven by geographical proximity, although some clusters demonstrate collaboration across continents. Significantly, the United States plays a vital role as a vital bridge connecting China and European nations in this network. The graph highlights the geographical trends and collaboration networks, offering significant insights into the global research collaborations in this specific field.

2.1.3 Co-occurrence Analysis

This section explores the co-occurrences of keywords within published RDH research articles. For the comprehensive analysis, a large number of keywords were collected from WoS and subsequently, the analysis was performed using fractional counting functionality within VosViewer, with a minimum occurrence threshold applied.

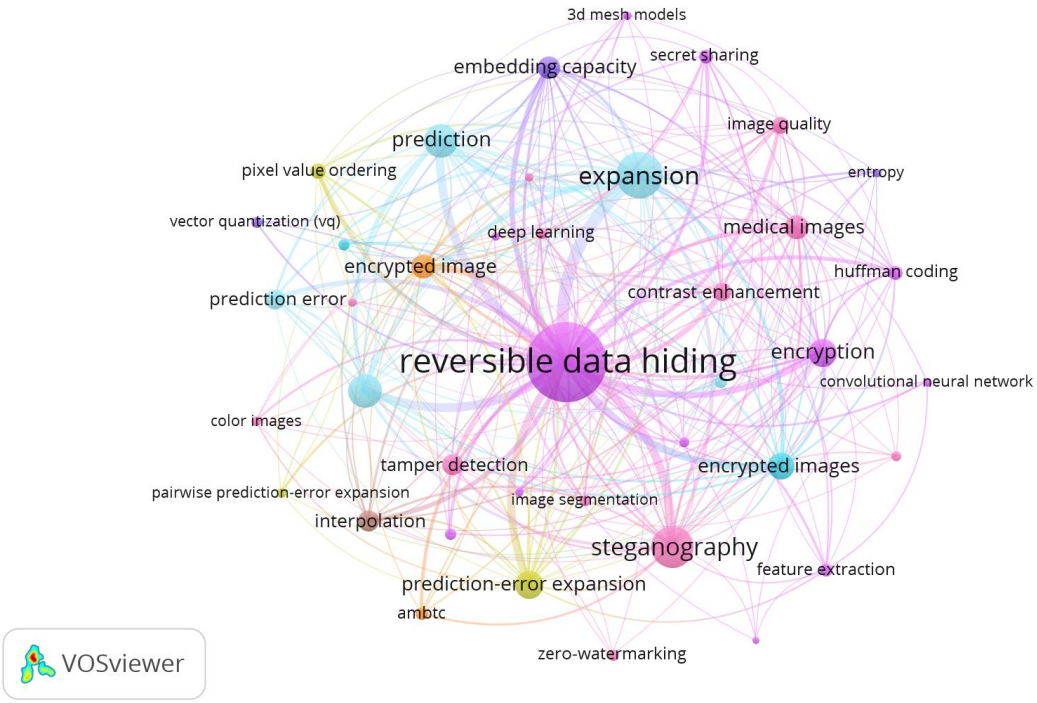


Fig. 2.4: Co-occurrence network of keywords

Fig. 2.4 represents the keyword cloud where the size of a keyword node represents the frequency of its occurrence. "Reversible data hiding" emerges as the dominant keyword, forming the largest node in the network. This is closely followed by "watermarking" and, "steganography", indicating the core focus areas of the field. Keywords like "image segmentation", "encrypted images", "medical images", and "color images" suggest a strong focus on the use of the image as cover media in majority of the RDH methods. "Prediction", "encryption", and

"expansion" form a distinct cluster, highlighting key technical approaches. Keywords like "expansion", " contrast enhancement", and "encryption" suggest the popularity of these techniques contributing to the evolving research focus within RDH. The high density of links between keywords demonstrates that the concepts within RDH are tightly integrated.

Based on the keyword clustering and their occurrence, we have identified hot topics and the recent trends in research. These are discussed in detail in subsequent sections.

2.2 Detailed Review of Prominent RDH Methods

This sub-section highlights some of the key areas identified from the keyword cluster, that represent the primary area of research within RDH evolved during the years.

2.2.1 Evolution of Contrast Enhancement:

RDHCE remains a mention-worthy area of research, specifically in the field of medical images where improving the contrast of the images helps in better diagnosis. Table 2.1 highlights some of the prominent RDHCE techniques along with their comparison. The concept was initially introduced by Wu et al. [47], who utilized histogram equalization to perform embedding while enhancing the contrast of the marked image simultaneously. Building on this foundation, Gao et al. [33] (2021) utilized an Adaptive Threshold Detector (ATD) method for accurate ROI segmentation and proposed a contrast stretching technique to improve both contrast and embedding capacity. Gao et al. [68] utilized a deep learning model for ROI segmentation along with a novel multi-group stretching method. These developments showcase the ongoing evolution and importance of contrast enhancement techniques in RDH, particularly in medical imaging applications.

Table 2.1: Prominent contributions within contrast enhancement

Author	Year	Methodology	Advantage	Limitation
Wu et al. [47]	2014	Histogram equalization	Known to be the first RDHCE method. Utilizes the histogram bin expansion to achieve contrast enhancement.	Pre-processing is performed to avoid overflow or underflow of pixel values but artificial distortions are introduced, especially when a large number of histogram bins are expanded for data hiding.

Table 2.1: Prominent contributions within contrast enhancement (continued)

Author	Year	Methodology	Advantage	Limitation
Wu et al. [48]	2015	Adaptive histogram equalization	Adaptively selects the minimum pixel interval on both sides of the histogram and then performs overlapping with the neighboring pixels.	Reduces the distortion, but does not preserve the order of pixel values impacting visual quality. Increase in complexity due to adaptive interval selection.
Wu et al. [49]	2018	Histogram equalization	Leveraging an order-preserving histogram bin expansion approach to improve the visual quality by retaining the order of pixel values.	Limited EC due to the limitation of adjacent bin merging. Over-enhancement of the image at high EC.
Kim et al. [82]	2018	Histogram equalization	Adaptively select bins for embedding based on the original brightness of the image. It preserves the brightness and limits the over-enhancement effect.	Computing the mean brightness in each iteration in this method results in extra calculation for the embedding system. EC is limited.
Yang et al. [83]	2018	Histogram stretching	ROI based high-capacity reversible data hiding that utilizes contrast stretching for embedding and contrast enhancement. Improves contrast even at a low embedding rate.	Limited EC.
Gao et al. [33]	2021	Histogram stretching	Intelligently exploits available empty bins in the histogram, and performs histogram stretching to increase the EC while providing stable contrast enhancement for a range of embedding rates.	EC and contrast enhancement is limited when the number of empty bins in the image is small.

Table 2.1: Prominent contributions within contrast enhancement (continued)

Author	Year	Methodology	Advantage	Limitation
Shi et al. [84]	2022	Histogram Stretching & Brightness Preservation	Pre-processes the ROI of the image utilizing histogram stretching and controls the direction of embedding with brightness calculation at each step.	Overall EC is limited, especially in images with few empty bins.
David et al. [85]	2022	Histogram equalization	Checks the pixel concentration on both half sides of the histogram. If the pixel concentration is less in the right half (128-255), then all the pixels on the right side (128-255) are overlapped on the left side pixels (0-127) or vice versa, thus creating 128 empty bins available for embedding on one side and vice versa.	Generates significantly high distortion.
Chen et al. [86]	2016	PVO, Adaptive HS	Utilizes three techniques: adaptive histogram shifting to improve the contrast, limiting the pixel modification to reduce distortion, and using PVO to increase the EC.	Computational complexity increased
Gao et al. [68]	2023	UNET3+ segmentation model, Multigroup Histogram stretching	Improved RDHCE method based on deep learning-based segmentation and utilizes multi-group stretching to improve contrast and EC.	Cannot be applied directly to color medical images.

2.2.2 Evolution of RDH in Encrypted Domain

With its capability of adding an additional layer of security, RDHEI has emerged as a crucial area of interest among the research community, combining the challenges of data hiding and

Table 2.2: Prominent contributions on RDHEI

Author	Year	Encoding Type	Decoding Type	Methodology	Summary	Limitation
Puteaux et al. [87]	2018	RRBE	Separable	MSB prediction-based	First PE-based RDHEI method utilising MED. Embeds one bit per embeddable pixel.	Limited EC.
Yin et al. [88]	2018	VRAE	Joint	Parametric binary tree labeling (PBTL)	Provides separability and exploits local spatial correlation using small blocks.	Does not leverage spatial correlation fully
Yin et al. [35]	2019	RRBE	Separable	Multi-MSB prediction and Huffman coding	Utilizes overall spatial correlation. Significantly improves EC.	Doesn't perform well in cases where PE is small.
Wu et al. [89]	2019	RRBE	Separable	Improved PBTL	Exploits spatial correlation of the entire image, which enables it to achieve better EC.	Embedding performance is limited.
Mohammad et al. [90]	2020	RRBE	Separable	Local difference predictor	Utilises block label prediction which enables EC of up to 2 bits per pixel	Fails to fully utilize larger block correlations.
Yu et al. [91]	2021	RRBE	Separable	Hierarchical bitplane labeling and Arithmetic encoding	Provides higher EC with hierarchical labels	Does not fully exploit small region-image correlation
Chen et al. [92]	2019	RRBE	Separable	Block-based MSB bit-plane rearrangement and uses Run length encoding	Achieves high EC.	Computational complexity increases by combining multiple techniques.
Xu et al. [93]	2022	RRBE	Separable	Hierarchical block variable length coding	Adaptive block sizing enables efficient compression of bit-planes with varying '1' bit distributions	Computationally intensive due to multiple levels of partitioning and threshold checks
Wang et al. [52]	2023	VRAE	Separable	PVO, histogram shifting	Achieves high EC.	Complex implementation due to scrambling-based encryption

encrypted image processing. Table 2.2 highlights some of the prominent RDHEI techniques along with their comparison. In one of the recent researches, a groundbreaking contribution by Ping et al. [94] proposed a method involving a two-stage RDHEI approach using asymmetric CNN-based and adaptive mean predictors. In another contribution, Zhang et al. [95] eliminated the need for a sign bit-plane and concentrated the zeros within high-bit planes by utilising asymmetric coding. Further, the method enhanced the overall embedding capacity by utilising bit-plane compression to compress redundant data in low-bit planes.

2.2.3 Evolution of PEE

Prediction plays an important role within the RDH framework, as prediction accuracy directly correlates to embedding performance. Among the various predictors proposed by the researchers, PEE has evolved as a key technique that utilizes correlation among pixels to perform prediction. The primary goal of these techniques is to select pixels that are highly correlated and derive computational models to accurately predict the target pixels. Table 2.3 highlights some of the notable methods representing the evolution of PEE techniques throughout the years. MED [96] is one of the earliest and computationally efficient predictors in this category. MED utilizes 3 neighbour pixels and considers horizontal and vertical edges to predict target pixels. However, due to its static nature, it does not perform well in complex regions. One of the other predictors, GAP [97], proposed the calculation of vertical and horizontal gradients to detect edges and smooth regions of an image. GAP provides an adaptive approach but has higher computational complexity in comparison to MED. One of the most popular and widely used predictors is the Rhombus predictor, which was proposed by Sachnev et al. [26] in 2009. The Rhombus predictor, also known as full-context predictor, takes into account the four pixels located at left, right, top, and bottom of the target pixel location to predict it. It provides highly accurate prediction in smooth areas; however, has limited accuracy in non-smooth regions. The traditional methods lack in devising non-linear relationships between pixels, hence having limited accuracy. To resolve this issue, Hu et. al. proposed a CNN-based predictor that, with its multi-receptive field capability, is able to capture non-linear relationships and predict complex patterns of the image. However, the predictor accuracy is dependent on the type and size of the training dataset. The latest work by Chang et al. [98] introduces a novel reversible data hiding method for color images based on adaptive mapping selection that leverages high inter-channel correlation by generating a three-dimensional (3D) prediction-error histogram (PEH) and adjusting 3D mappings in an ordered iterative manner to optimize embedding performance. Another eye-catching paper by Kumar et al. [9] discusses PEE-RDH and provides a thorough examination of multiple predictors. The paper also presented a new ensemble learning-based XGBoost predictor for RDH, which utilizes the rhombus context for prediction, aiming to enhance prediction accuracy in data-hiding processes.

Table 2.3: Prominent contributions within prediction error expansion

Author	Year	Methodology	Context of pixel	Advantage	Limitation	Complexity
Hong et al. [96]	2009	Median Edge Detection (MED)	3	Simple and computationally efficient.	Limited to vertical and horizontal edges; does not perform well in highly complex regions due to non-adaptiveness.	$O(n)$
Conetter et al. [97]	2010	Gradient Adjusted Predictor (GAP)	7	Highly adaptable; recognizes various edge types.	Computationally intensive.	$O(n)$
Avramovic et al. [99]	2011	Gradient Edge Detection (GED)	5	Combines simplicity of MED with efficiency of GAP; simpler than GAP, more effective than MED.	Requires threshold adjustment for gradient.	$O(n)$
Sachnev et al. [26]	2009	Rhombus Mean (Full Context)	4	Good prediction accuracy in smooth regions.	Limited accuracy; may not capture complex patterns. Performance may decrease in non-smooth regions.	$O(n)$
Dragoi et al. [100]	2012	Dragoi and Coltuc (D&C)	4	Effective in capturing directional information; improves performance in some non-smooth regions.	May not always outperform Rhombus Mean.	$O(1)$

Table 2.3: Prominent contributions within prediction error expansion (continued)

Author	Year	Methodology	Context of pixel	Advantage	Limitation	Complexity
Ma et al. [101]	2019	Least Square (LS)	Adaptive	Uses a linear combination of prediction coefficients.	Computationally intensive.	$O(n^3)$
Hwang et al. [102]	2016	Least Absolute Shrinkage and Selection Operator (LASSO)	Adaptive	Reduces overfitting and provides feature selection; performs both variable selection and regularization.	Computationally intensive.	$O(np)$, where p is the number of features.
Wang et al. [103]	2021	Ridge Regression	10	Solves the overfitting problem of LS method; addresses multicollinearity and improves prediction accuracy.	May not perform well with many irrelevant features.	$O(np^2 + p^3)$
Hu et al. [80]	2021	CNN Predictor (CNNP)	Adaptive	Can capture complex patterns and non-linear relationships; potentially more accurate.	Requires significant training data; computationally intensive.	$O(n)$
Kumar et al. [9]	2023	XGBoost	11	Often outperforms other methods; handles complex datasets effectively.	Can be computationally intensive and may overfit if not properly tuned.	$O(n \cdot d \cdot t)$, where d is max depth and t is number of trees.

2.2.4 Evolution of Robust RDH

RDH has gained widespread popularity due to its unique ability to recover the original cover image while extracting hidden secret data. However, traditional RDH techniques often lack robustness against attacks or image processing operations such as compression, geometric deformation, and noise addition. These vulnerabilities hinder their ability to reliably extract embedded data, posing significant challenges in applications like watermarking, where accurate watermark retrieval is crucial for copyright protection. To address these limitations, researchers have developed RRDH techniques that focus on enhancing the resilience of embedding processes against such distortions. The primary goal of RRDH is to ensure the lossless recovery of embedded data, even under adverse conditions. Table 2.4 summarizes notable advancements in RRDH, highlighting their robustness techniques, performance characteristics, and limitations.

Table 2.4: Prominent contributions within Robust RDH

Author	Year	Category	Robustness Technique	Performance Characteristics	Limitation
Ni et al. [65]	2008	Statistical Distribution based	Difference between pixel pairs of a block is used as a statistical parameter	Robust against salt-n-pepper noise. Robust to JPEG compression. Avg. PSNR $>$ 38 dB. EC: 512–1024 bits.	Very limited embedding capacity; cannot handle geometric deformation attacks effectively.
Zeng et al. [66]	2010	Statistical Distribution based	Arithmetic difference-based block-level embedding; robustness achieved via threshold gap maintenance.	Robust against JPEG compression. Avg. PSNR $\sim$38 dB. EC: 2048 bits (double of Ni et al.).	Moderate robustness against geometric attacks.
Zong et al. [104]	2015	Statistical Distribution based	Gaussian low-pass filter preprocessing; safe band between embedding and non-embedding pixels.	Robust against cropping and bending; improved compression robustness.	Ineffective against geometric deformation due to feature sensitivity.
Coltuc et al. [105]	2007	Statistical Distribution based	Two-stage embedding with compensation data embedding for robust extraction.	Improved robustness to JPEG compression; high imperceptibility and reversibility.	Weak robustness due to reuse of embedding regions; geometric instability.

Table 2.4: Prominent contributions within Robust RDH (continued)

Author	Year	Category	Robustness Technique	Performance Characteristics	Limitation
Liu et al. [55]	2017	Transform-based	Fractional Krawtchouk transform; fractional orders as secret keys enhance robustness.	High robustness and imperceptibility. Avg. PSNR ∼38.35 dB.	Limited effectiveness against geometric deformations.
Wang et al. [106]	2017	HSB-based	Significant-bit-difference expansion.	Robust against noise and compression; EC: 50K bits.	Lacks robustness against geometric attacks.
Wang et al. [56]	2020	Transform-based	Haar wavelet transform decomposing image into low and high frequency components.	Low-frequency embedding achieves robustness; Avg. PSNR ∼39.5 dB.	Wavelet-based features distort under geometric transformations.
Hu et al. [59]	2020	Geometric Invariant Moments	Two-stage embedding: low-order Zernike moments + distortion compensation.	Robust against geometric deformation; Avg. PSNR ∼39.2 dB.	Sensitive to numerical instability in Zernike moments.
Hu et al. [60]	2021	Geometric Invariant Moments	Polar harmonic transform with quantization watermarking.	Robust to JPEG compression (QF 10) and geometric deformations. Avg. PSNR ∼38.55 dB.	Sensitive to small variations, leading to inaccuracies.
Tang et al. [61]	2023	Geometric Invariant Moments	Two-stage optimization and rounded error compensation in selected Pseudo-Zernike moments (PZMs).	Outperforms other moment-based methods; robust to multiple distortions. Avg. PSNR ∼40 dB, EC: 34K bits.	Orthogonal moments prone to numerical instability in noisy data.
Ma et al. [107]	2024	Geometric Invariant Moments	Polar Harmonic Fourier Moments (PHFMs) with Quantization Index Modulation.	High robustness to rotation, scaling, and compression. Avg. PSNR: 40.4 dB.	Effective only within tested geometric distortion range.

Xiong et al. [64]	2023	Secret-sharing based	Chinese Remainder Theorem-based secret sharing.	Lightweight encryption and high visual quality; robust multi-party embedding.	Multi-layer PEE increases computational complexity.

Early approaches like Ni et al. [65] utilized pixel-pair differences for HS, offering robustness against JPEG compression but limited embedding capacity. Zeng et al. [66] introduced arithmetic difference-based block embedding, achieving a significant increase in capacity while maintaining robustness. Zong et al. [104] improved resilience against cropping and random bending attacks by employing Gaussian low-pass filtering and safe embedding bands.

Two-stage embedding methods also gained traction. Coltuc et al. [105] ensured robust data extraction with compensation embedding, and Wang et al. [106] used HSBs for robustness against compression and random noise. Transform-based approaches, such as Wang et al. [56], leveraged Haar wavelet transforms to enhance robustness but faced challenges with geometric transformations. Geometric invariant moment-based techniques addressed geometric instability. Hu et al. [59] utilized Zernike moments for embedding, achieving robustness against rotation and scaling, albeit with numerical instability. Tang et al. [61] advanced this approach using adaptive normalization with Pseudo-Zernike moments, offering strong resistance against attacks and achieving high image quality. Ma et al. [107] further enhanced robustness using Polar Harmonic Fourier Moments and quantization index modulation, demonstrating superior performance against geometric deformations with a PSNR of 40.4 dB.

In a distinct approach, Xiong et al. [64] introduced a secret-sharing-based RDH technique leveraging the Chinese Remainder Theorem. This method achieved lightweight encryption, high visual quality, and robust multi-party embedding. Despite its robustness, the multi-layer PEE increased computational complexity, which may limit scalability. These advancements illustrate the progression of RRDH techniques, with each addressing specific challenges and contributing to the broader goal of achieving robust RDH.

2.3 Summary

In summary, the domain of RDH has experienced significant growth and innovation, largely driven by the contributions of RDHCE, RDHEI, and prediction methods. These methods have notably enriched the content security and visual quality of the embedded images. Despite significant progress in enhancing EC, PSNR, reversibility, and prediction performance, achieving an optimal balance among these factors remains an open challenge. While some methods demonstrate strong performance in one dimension but at the expense of another, this underscores the need for further research and innovation. For example, a number of deep learning

based studies have emerged focusing on prediction performance enhancement leveraging the multi-receptive field capabilities of deep neural networks. However, there remains a huge scope to explore other deep learning techniques, such as attention mechanisms, transformers, etc., that have proven themselves to be more efficient in other applications.

CHAPTER 3

A HIGH-CAPACITY REVERSIBLE DATA HIDING WITH CONTRAST ENHANCEMENT AND BRIGHTNESS PRESERVATION FOR MEDICAL IMAGES

The healthcare industry has witnessed an increase in the use of cloud storage, resulting in a significant demand for safeguarding medical records from potential attackers. In response to this challenge, the RDH ensures the concealment of private and confidential data with minimal loss to the cover image while offering opportunities for image enhancement. In some scenarios, particularly with inadequate illumination and substandard capability imaging systems, the captured images suffer from low contrast, hindering their interpretability in medical diagnosis. However, the traditional RDH techniques primarily focus on achieving high PSNR and embedding rate while ignoring the visual enhancement. This Chapter introduces a new RDH method with contrast enhancement for medical images. The method aims to provide high EC while improving the contrast and also preserving brightness. To achieve this, the proposed method initially segments cover images into regions of interest ROI and NROI and employs different embedding strategies based on the characteristics of each region, thereby enhancing embedding performance. A novel pre-processing technique is introduced for pre-processing ROI pixels, which capitalizes on the unique properties of medical images, reorganizing and creating empty bins to provide an enlarged EC with less distortion.

3.1 Introduction

As the healthcare industry progresses towards digitization, cloud storage has become a popular and convenient choice to store and analyze medical data. Electronic Patient Records [108, 109] contain sensitive information, making it crucial to protect them from potential threats. However, due to the semi-trusted nature of cloud applications, medical data breaches occur occasionally. To address this issue, RDH in medical images has gained popularity. RDH allows

for embedding confidential information within a cover media, which can be later extracted as needed, all while also ensuring the lossless restoration of the cover. [110–112].

However, recent studies have shown that maximizing PSNR alone does not ensure satisfactory visual quality, especially for low-contrast images. Improving the visual quality of low-contrast images is a challenge for RDH methods. Enhancing image visibility is a fundamental objective, and one effective means to achieve this is through contrast enhancement. The assessment of image contrast enhancement is typically based on the RCE, which measures the extent of contrast improvement between the original and enhanced images, with values ranging from 0 to 1. In applications such as medical diagnosis [113], enhancing image quality through contrast enhancement is crucial for providing a more detailed and visually comprehensible content representation. As a result, researchers have started exploring techniques combining RDH with contrast enhancement, often called RDHCE. The goal of RDHCE is to enhance visual quality while performing data embedding.

Medical images consist of an ROI containing valuable information and an NROI less relevant to the observer. Therefore, RDHCE techniques are designed to use histogram shifting to enhance ROI contrast, while NROI can be used for extensive data embedding. Repeated histogram shifting redistributes the higher histogram bins toward their outer bins, effectively achieving the effects of histogram equalization. The outcome expands the dynamic range of the image histogram, resulting in enhanced image contrast and an overall improvement in visual quality. Some popular contrast enhancement methods coupled with RDH for medical images have been proposed in [1, 114–122]. This chapter proposes a novel high-capacity RDHCE method for medical images. The main contribution of the proposed method can be summarized using the following points:

- The proposed method initially partitions the image into ROI and NROI regions, taking into account their relevance to the medical practitioners. It then employs distinct embedding strategies for hiding the secret data based on their sensitivity to modifications. This approach enhances EC without compromising image quality for medical practitioners.

- Furthermore, the proposed method addresses the shortcomings of existing embedding methods through a novel pre-processing technique that fully leverages the characteristics of medical images. The selection of bins for shifting and overlapping is performed carefully based on pixel concentration ratio (PCR), effectively reducing overall distortion.

- As a result, the proposed method achieves a superior embedding rate. It produces significantly improved stego-images with higher PSNR, SSIM, and contrast, all while preserving brightness for all the test images.

3.2 Literature Review

This section discusses the SOTA RDHCE methods and is divided into two subsections. The first subsection (3.2.1) provides a brief overview of the SOTA RDHCE methods and their strengths and limitations. These methods assist in highlighting the core principle utilized in the proposed technique. In the second subsection (3.2.2), we conduct a detailed review of Gao et al.*'s* scheme. We provide an in-depth analysis of its operation and examine its shortcomings and advantages.

3.2.1 Related Work

Most of the RDH techniques aim to achieve high EC while reducing structural distortions in the resulting stego images. However, due to certain limitations of the image-capturing systems, such as degraded imaging sensors, inadequate illumination, and inappropriate system settings, the images possess low contrast. This is not desirable in medical applications, as the images contain critical information to be analyzed by practitioners. To address this concern, in 2014, [114] introduced the first RDH method that enhances the contrast of stego-images while embedding secret data. This method achieves histogram equalization by performing histogram shifting, resulting in contrast enhancement and data embedding simultaneously. However, the method can lead to pixel intensity mismatches after a large number of repetitions in the embedding process. This occurs because the approach involves a pre-processing step that disrupts pixel values, leading to intensity-mismatched artifacts.

To address the issue present in [114], a new RDHCE method was introduced in 2015 by Wu et al. [115]. The method adaptively selects the minimum pixel interval on both sides of the histogram and performs overlapping with neighboring pixels. While this approach results in low distortion, it does not preserve the order of pixel values. Additionally, the adaptive interval selection also increases the computational complexity. In 2018, Wu et al. [116] discussed a new RDH method based on order-preserving histogram bin expansion (OPHBE) to enhance visual quality by retaining the order of pixel values. However, with the increase in capacity, the method over-enhances the image and introduces distortion.

In 2015, Kim et al. [117] introduced the first Automatic contrast enhancement method with reversible data hiding (ACERDH). This method effectively eliminates the intensity mismatch artifacts caused by discrepancies in intensity levels. It achieves increased contrast while ensuring that overflow/underflow does not occur during the embedding process, eliminating the need for pre-processing. However, the ACERDH method sometimes results in over-enhancement of contrast due to the lack of consideration of the original image's mean brightness. To overcome the aforementioned drawback of [117], Kim et al. [118] discussed a new RDHCE method in 2018. This method adaptively selects the histogram bins based on the cover image's original brightness, enhancing contrast while preserving image brightness. However, this modification necessitates additional calculation at each embedding iteration, which increases the processing time. Nevertheless, the method successfully preserves image brightness while avoiding over-

enhancement, resulting in a less susceptible image. However, one of the drawbacks of this method is its relatively weaker contrast enhancement.

In 2019, Wu et al. [123] proposed a novel RDH scheme based on 2D histogram modification. The method discusses a new pre-processing approach to avoid overflow/underflow conditions and is able to achieve higher contrast in comparison to 1D histogram modification techniques. However, the overall embedding capacity of the scheme is limited. In 2020, Chen et al. [124] introduced a novel pre-processing technique to calculate an upper bound for embedding. Using the calculated upper bounds, the method achieved high embedding capacity with minimal distortion in image quality. However, the method only focuses on preserving structural characteristics and does not focus on enhancing visual characteristics such as contrast. In 2021, Mansouri et al. [125] proposed an RDHCE technique that includes a bidirectional shifting of the histogram to embed the secret data. Along with achieving high contrast and brightness preservation, the method achieves low complexity by avoiding the additional calculations required to calculate the shifting direction. However, the method has limited embedding capacity and also lacks security analysis.

In another pre-processing technique proposed by Nunez et al. [120], the method assesses the pixel concentration on both halves of the histogram. If the pixel concentration is lower on the right half (pixel values 128-255), then all the pixels on the right side are overlapped over the left side pixels (pixel values 0-127). This action creates 128 empty bins available for embedding on one side and vice versa. However, it's worth noting that this method introduces a significantly high overall distortion level.

In 2016, Liu et al. [121] noted that medical practitioners often focus more on the foreground part of medical images, known as the ROI, rather than the background, referred to as the NROI, as the NROI typically lacks essential information. In response to this observation, Yang et al. [122] introduced an ROI-based RDH method that utilizes contrast stretching for contrast enhancement. This approach strategically distributes empty bins to create space for embedding while preventing overflow and underflow, thus reducing the size of side information. While this method can enhance contrast even at lower embedding rates, its overall embedding capacity is limited.

In 2022, Shi et al. [126] proposed a hybrid approach that combined the concepts used in [118] and [1]. The method pre-processes the ROI of the image using contrast stretching and controls the embedding direction with brightness calculation at each step. The resulting stego-image possesses improved brightness preservation and standard deviation. However, the overall embedding capacity remains limited, especially in images with few empty bins.

3.2.2 Review of Gao et al. [1]

In 2021, Gao et al. [1] introduced a novel RDH method with automatic contrast enhancement for medical images, often abbreviated as "RDHACEM" in this work. RDHACEM divides the

original medical image into two distinct regions: the foreground region and the background region. The foreground region, referred to as the ROI, typically contains crucial information necessary for medical diagnosis. On the other hand, the background region, known as the NROI, is generally monochromatic and lacks any vital diagnostic information. This division is implemented to apply different RDH techniques based on the region's sensitivity. More specifically, since the NROI region lacks valuable information, it can accommodate a more significant number of bits for data embedding without adversely affecting medical diagnosis. However, extra attention is required for the sensitive ROI region. To address this, Gao et al.*'s* RDHACEM initially calculates the PCR. PCR represents the ratio of pixels with values between [0, 127] or [128, 255] to the total number of pixels on both sides of the histogram. Subsequently, the method performs histogram shifting in the region with a low PCR, using Eq. (3.1), to ensure the production of high-quality resultant images. In medical images, pixels primarily fall within the ranges of [0, 127] or [128, 255], and expanding the entire histogram of the images often leads to a decrease in visual quality.

$$R_s(p,q) = \begin{cases} R(p,q) + T_s & \text{if } PCR[0,127] > r, R(p,q) \in [128,255] \\ R(p,q) - T_s & \text{if } PCR[128,255] > r, R(p,q) \in [0,127] \\ R(p,q)Y & \text{otherwise} \end{cases} \tag{3.1}$$

where $R_s(p,q)$ represents the pixel value obtained after shifting the corresponding $R(p,q)$ from the original image. T_s denotes the distance over which a pixel can be shifted. Shifting is employed to create additional space within the region with high PCR, facilitating efficient histogram stretching. Subsequently, stretching is performed using Eq. (3.2) to redistribute empty bins to locations where more embedding is feasible.

$$R_c(p,q) = round\left[(L_x - L_n) * \frac{R_s(p,q) - R\prime_n}{R\prime_x - R\prime_n} + L_n\right] \tag{3.2}$$

where $R_c(p,q)$ represents the pixels after the stretching operations, R'_x and R'_n denote the maximum and minimum pixel values in the ROI, excluding the shifted pixels, while L_x and L_n are the upper and lower boundaries of the stretched pixels. Following the aforementioned automatic stretching, embedding occurs within the ROI, resulting in improved visual perception compared to the erstwhile SOTA methods. Additionally, embedding occurs in non-ROI (NROI) regions to incorporate the secret data, expanding the EC. In this process, the least significant bits (LSBs) of pixels are replaced with the secret data bits, resulting in increased EC with minimal distortion. However, the method's performance diminishes when there are very few or zero empty bins, resulting in low EC and no contrast enhancement.

After reviewing the existing methods of RDHCE [123–129], it becomes apparent that these methods lack an optimal solution that can simultaneously achieve contrast enhancement with a high embedding rate while minimizing image artifacts. These methods often do not fully

leverage the unique characteristics of medical images, especially when the number of available empty bins is limited. Additionally, many of these methods focus on enhancing only one aspect, such as contrast, brightness, or embedding capacity. Therefore, there is a growing need for a universal approach capable of effectively enhancing all types of medical images. Such an approach should not only improve contrast and embedding capacity but also preserve image brightness.

To overcome the aforementioned limitations of the approach presented by [1] and other SOTA methods, this chapter introduces a novel RDH for medical images tailored to their unique characteristics. The proposed method strives to achieve a high EC while simultaneously enhancing contrast and preserving image brightness. Further details regarding the proposed method are elaborated in the following section.

3.3 Proposed Method

The proposed method for embedding data in medical images comprises several stages, as illustrated in Fig. 3.1. The first stage is partitioning the original image into the ROI and NROI

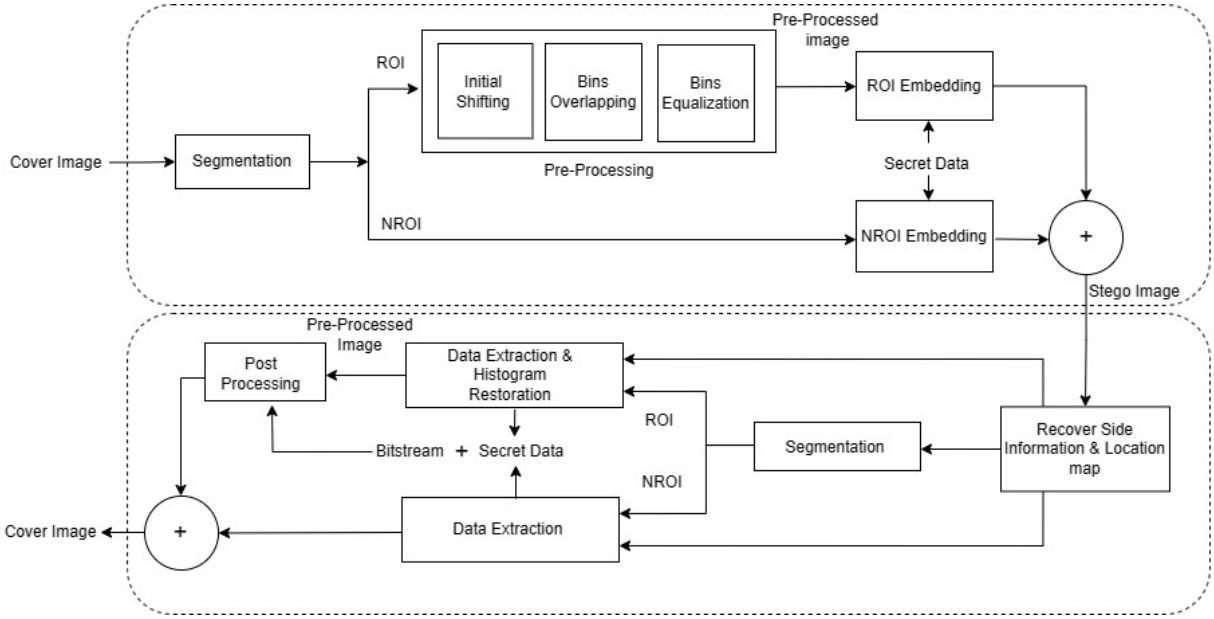
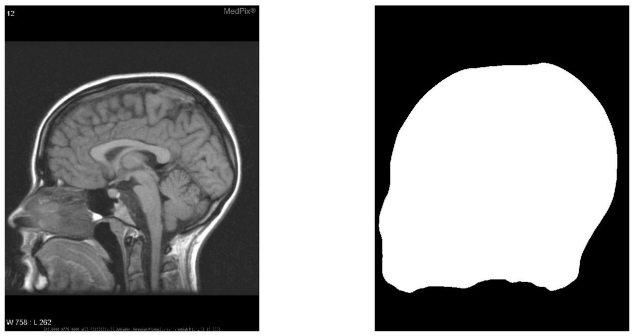


Fig. 3.1: Framework of the proposed method.

regions. Then, pre-processing is performed to make sure there are enough empty bins for embedding without the risk of underflow or overflow. The third stage is to embed data in both the Region of Interest (ROI) and Non-Region of Interest (NROI). Finally, an almost inverse process to the embedding is used to recover the original image from the stego-image. Each of the aforementioned stages is detailed in separate subsections as follows:

3.3.1 ROI and NROI Segmentation

As a medical image usually comprises two regions, ROI and NROI, different data-hiding approaches with varying levels of tolerance can be used to embed the secret data to achieve optimal performance. However, before this, accurate image segmentation into the ROI and NROI is required. To accomplish this, two well-known segmentation approaches, namely Adaptive Threshold Detector (ATD) and Otsu Thresholding Method (OTM) introduced by [130] and [131], have been utilized in SOTA methods. Among the two, ATD has gained more popularity within the research community due to its precise segmentation of images, which has been validated by elaborative experimentation done in [126], [130]. Notably, [130] designed ATD to enhance image classification accuracy by minimizing within-class variance. In contrast, OTM performs classification using within-class variances and combines these variances across classes into a single overall variance for thresholding. Pai et al. emphasized that their standard deviations cannot be directly compared when the mean difference between two classes is substantial.



(a) Original image **(b)** ROI

Fig. 3.2: Segmentation of ROI and NROI

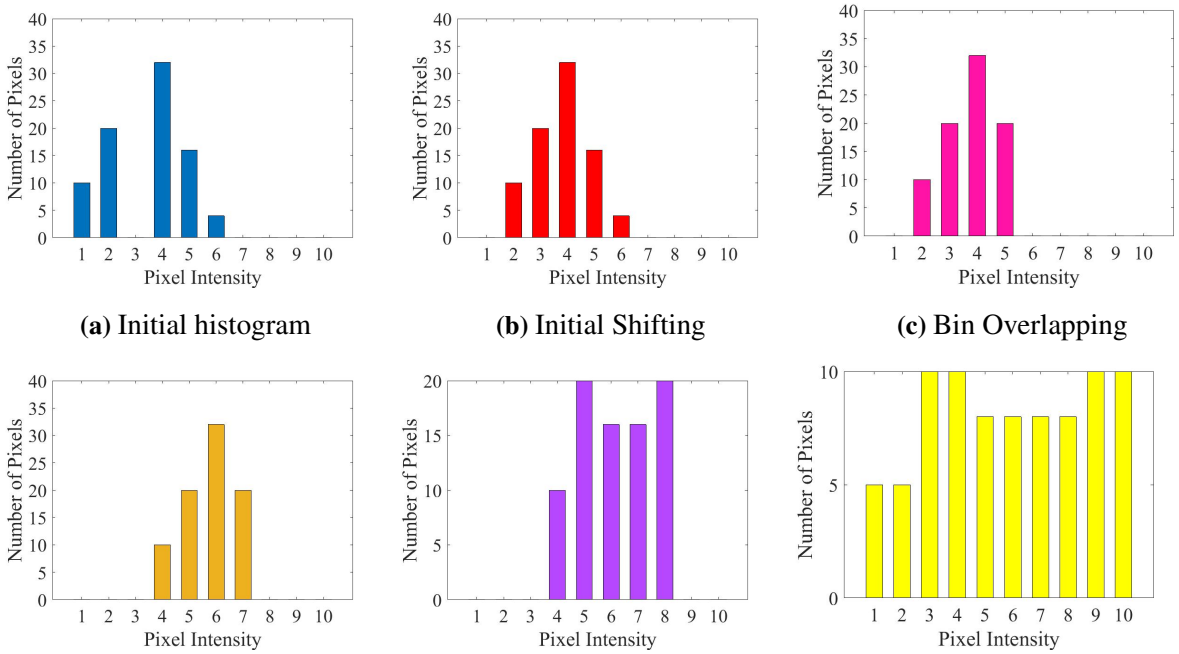
Building upon ATD's demonstrated adaptability, the proposed work also utilizes it for image segmentation. Once the optimal threshold (T) is determined, the threshold image is calculated using Eq. (3.3).

$$g(x,y) = \begin{cases} 255 & \text{if } f(x,y) > T \\ 0 & \text{if } f(x,y) < T \end{cases} \tag{3.3}$$

where $f(x,y)$ denotes the grey-level pixel intensity at position x, y, and $g(x,y)$ denotes the updated grey-level pixel intensity. Pixels with values lower than the T are marked as 0, while others are marked as 255. To visualize the resultant threshold image after applying ATD to the original image, both images are presented in Fig. 3.2. In this context, all pixels in the original image corresponding to the value 255 in the threshold image are referred to as the ROI region. In contrast, the pixels corresponding to the value 0 are referred to as the NROI region for further reference in this chapter.

3.3.2 Pre-processing

It has been observed that RDHCE methods often introduce some artificial distortion while performing embedding in an image, which can sometimes result in overflow and/or underflow issues. In SOTA methods, different pre-processing approaches have been proposed to create the required number of empty bins that help in avoiding overflow/underflow during embedding. However, these approaches do not consider image characteristics and tend to introduce artificial artifacts in the images. To address these concerns, this chapter introduces a novel pre-processing approach for the ROI segment based on the image properties to adaptively create and position the empty bins in the image histogram to ensure minimal distortion. The proposed pre-processing stage not only enhances the quality of the resultant image but also improves the EC and image brightness.



(a) Initial histogram **(b)** Initial Shifting **(c)** Bin Overlapping

(d) Bins Equalization **(e)** Embed 32 bits, S=1 **(f)** Embed 114 bits, S=6

Fig. 3.3: Example histograms generated during embedding by the proposed method

The proposed pre-processing algorithm is presented in Algorithm 1, consisting of three main steps: Initial Shifting, Overlapping, and bin equalization. To facilitate understanding, we have provided a visual illustration of these steps using an example in Fig. 3.3. For simplicity, the example employs a pixel value range of [1, 10] instead of the standard range [0, 255] for digital grayscale images.

In **Step 1**, given an ROI image (R) and the number of iterations (S), the algorithm first calculates the pixel frequency distribution through histogram analysis. The proposed pre-processing algorithm computes the pixel frequency distribution via histogram analysis, as shown in Fig. 3.3a, enabling the calculation of the total number of empty bins ($Ebins_{[0,255]}$), empty bins on the left side $Ebins_{[0,127]}$ and empty bins on the right side $Ebins_{[128,255]}$. It can be clearly ob-

Algorithm 1 Proposed Pre-processing

Input: ROI image (R), Number of iterations (S)
Output: Pre-Processed Image (R_P)
Step-1: Calculate the frequency distribution of image R using histogram analysis.
Step-2: Initial Shifting- Shift the existing empty bins on both sides of the histogram to extreme left and right, respectively.

- Count the number of empty bins $Ebins_{[0,127]}$, $Ebins_{[128,255]}$ and $Ebins_{[0,255]}$.
- Determine E_P, which denotes the position of the empty bin P_{th}, where $E_P \in [0, 255]$ and $P \in [1, Ebins_{[0,255]}]$.
- For all the empty bins $Ebins_{[0,255]}$, repeat below shifting steps:

if ($Ebins_{[0,127]} > 0$) & ($E_P <= 127$) **then**
  if $R(i,j) < E_P$ **then**
    $R_P(i,j) = R(i,j) + 1$;
  end if
end if
if ($Ebins_{[128,255]} > 0$) & ($E_P >= 128$) **then**
  if $R(i,j) > E_P$ **then**
    $R_P(i,j) = R(i,j) - 1$;
  end if
end if

R_P describes the modified pixels after performing initial shift.
Step 3: Bin Overlapping-
if $S > Ebins_{[0,255]}$ **then**
  perform overlapping for (S - $Ebins_{[0,255]}$)
  if $PCR_{[0,127]} > PCR_{[128,255]}$ **then**
    Perform overlapping of (S-$Ebins_{[0,255]}$) on right
  else if $PCR_{[0,127]} < PCR_{[128,255]}$ **then**
    Perform overlapping of (S-$Ebins_{[0,255]}$) on left
  end if
end if
Total empty bins after overlapping = S
Step 4: Bins Equalization-
if $Ebins_{[0,127]} > Ebins_{[128,255]}$ **then**
  $toshift = (\frac{S}{2} - Ebins_{[128,255]})$
else if $Ebins_{[0,127]} < Ebins_{[128,255]}$ **then**
  $toshift = -(\frac{S}{2} - Ebins_{[0,127]})$
else if $Ebins_{[128,255]} = Ebins_{[0,127]}$ **then**
  $toshift = 0$
end if

served that there are a total of 5 empty bins at the position E_P (E_1, E_2, .. E_5 = 3, 7, 8, 9, and 10) in Fig. 3.3a

In **Step 2**, Initial shifting is performed to move the existing empty bins to the far left and far right. Initial shifting enlarges the contiguous space for embedding and facilitates the efficient design of the embedding algorithm. As illustrated in Fig. 3.3b, there is one empty bin ($E_P = 3$) in the left half of the histogram, which is shifted to the far left. There are four empty bins already located on the far right, and they do not require further shifting.

In **Step 3**, Bin overlapping is performed. However, it is important to note that bin overlapping is executed only if the total number of empty bins present in the histogram is less than the required number of iterations (S). The proposed method adaptively designs the overlapping process by evaluating the PCR on both sides of the histogram and selecting empty bins for overlapping on the side with a lower PCR. Furthermore, the overlapping is limited to $S - Ebins_{[0,255]}$ bins, ensuring minimal distortion. The number of iterations (S) is set to 6 in the example provided. As seen in Fig. 3.3a, the PCR for the left side of the histogram, covering bins from 1 to 5, is 0.95, which is higher than the PCR for the right side (bins from 6 to 10), which is 0.05. Consequently, the method performs overlapping of $S(6) - Ebins_{[0,255]}(5)$=1 bin in the region with a low PCR, which is the right side. Fig. 3.3c illustrates the histogram obtained after overlapping bin 6 over bin 5. Following this overlapping step, the number of empty bins matches the required number of iterations (S).

In **Step 4**, Bin equalization (Final Shifting) is executed to ensure that the empty bins obtained after Step 3 are evenly distributed on both sides of the histogram. This equalization process enhances contrast and preserves brightness during the secret data embedding. Fig. 3.3d displays the histogram of the pre-processed image after bin equalization, revealing an equal number of bins (3) on both sides of the histogram. The embedding method for the ROI region is discussed in the following sub-section.

3.3.3 Embedding in ROI

The embedding in the ROI region is done to enhance the contrast of the ROI region while hiding some bits of the secret data. To achieve this, we first identify the peak bin (h(p_{max})) representing the most frequent pixel value (p_{max}) in the image R_P. Subsequently, the embedding is performed using the following Eq. (3.4).

$$R_E(i,j) = \begin{cases} R_P(i,j) + D_s * b_i & \text{if } R_P(i,j) = p_{max} \\ R_P(i,j) - 1 & \text{if } R_P(i,j) < p_{max}, D_s = -1 \\ R_P(i,j) + 1 & \text{if } R_P(i,j) > p_{max}, D_s = 1 \\ R_P(i,j) & \text{otherwise} \end{cases} \tag{3.4}$$

Where $b_i \in [0,1]$ represents a bit of the secret data with a length of len, $R_E(i,j)$ denotes a

pixel in the stego-ROI image R_E, and $D_s = [-1, 1]$ defines the direction of shift. The direction depends on the number of available empty bins on either side. D_s can be alternated to shift the histogram on both sides.

$$D_s = \begin{cases} -1 & \text{if } Ebins_{[0,127]} > 0 \\ 1 & \text{if } Ebins_{[128,255]} > 0 \end{cases} \tag{3.5}$$

As mentioned in Eq. (3.4), the embedding step is repeated for S iterations. Figures from 3.3e to 3.3f demonstrate the steps taken in embedding using the proposed method. At the first iteration, S=1, the peak bin (p_{max})= 6 is identified, which contains 32 pixels. With empty bins available on both sides, the shifting direction D_s can be chosen as either of the sides. In the illustrated example, for the first iteration, D_s=1 is selected, and data is embedded according to Eq. 3.4. Let b_i=1010..10 be the 32 bits, with an equal number of 1s and 0s, to be embedded in bin (p_{max}). As D_s=1, all the pixels with values greater than 6 are shifted in the right direction by 1, and the pixels to its left remain unchanged. Pixels with value 6 are modified based on the bits to be stored. If b_i=1, the pixel is updated to value 7; otherwise, if b_i=0, the pixel remains unchanged with value 6. The modified histogram after iteration 1 is illustrated in Fig. 3.3e. As S=6, embedding is performed for 6 iterations, and Fig. 3.3f shows the histogram obtained after the final iteration. As indicated in the illustrated example, the proposed method allows for embedding 114 bits, whereas RDHACEM [1] can only embed 88 bits. This shows that the proposed method has a higher EC in the ROI. Experimental results are presented in section 3.4 to validate this.

3.3.4 Embedding in NROI

As discussed above, the NROI region is usually less important to the observer, as it generally contains no useful information. Hence, it provides an opportunity to embed a large amount of data in this region. As per the embedding requirement, we can embed up to 3 bits per pixel in this region. The proposed method simply replaces the required number of LSBs with the payload bits. The required number of embedding bits per pixel is calculated by the total number of NROI pixels and the total number of bits to be embedded.

3.3.5 Details of Auxiliary Information

To completely recover the secret data and the cover image from the stego-image R_E, it is important to record the side information along with the secret data. The side information mainly includes:

- Location map denoting ROI and NROI segmentation [1]
- Length of secret data for ROI and NROI, respectively
- Peak bin pixel at each iteration (p_{max})

- Direction of shift at each iteration (D_s)
- Number of bits embedded into each NROI pixel (N_B)
- Number of iterations (S)

To reduce the size of the side information, it is compressed by Arithmetic coding, and then the length of the condensed side information, which is denoted by L_{SI}, is calculated. The parameter N_B is embedded into LSBs of 2 pixels, and the parameter L_{SI} is embedded into LSBs of the next 18 pixels reserved for side information. The rest of the side information (i.e., Location map, secret data length, last iteration D_s, last iteration p_{max}, iterations S) is embedded next. Further, during each iteration of ROI embedding, the previous iteration's p_{max} and D_s are embedded along with the secret data. It is to be noted that we also need to record the additional auxiliary information generated during the pre-processing stage, which includes the following

- Total number of empty bins ($Ebins_{[0,255)]}$) available in cover image
- Positions of all empty bins (E_P)
- Bin value of overlapped bin n, if overlapping is performed.
- Direction of overlapping (O_D). $O_D = -1$, if (n+1) bins are overlapped onto n. $O_D = 1$, if (n-1) bins are overlapped onto n.
- Bit-stream (B_S): A bit-stream is generated by performing a raster scan on the image to record the order of pixels of two overlapping bins. Record bit 0 for all the pixels having value n-1/n+1 and value 1 for all the pixels having value n.
- Final Shifting value ($toshift$) of the pre-processing and its direction

To clearly understand the requirement of recording the above-mentioned information, let us elaborate on this by using an example as depicted in Fig. 3.3. As we note that the number of empty bins available (5) is lower than the number of required iterations (6), we perform overlapping of 1 bin. Further, as the PCR is lower on the right side of the histogram, overlapping on the right side is performed. Consequently, bin 6 overlaps with bin 5. We record the bin value 5 and a bit stream of 20 bits, with 1s at the place of pixel 5 and 0s at the place of pixel 6. This is repeated for all the bins on which overlapping is performed. The additional side information is recorded by concatenating all the recorded information. If there are a large number of overlapping bins, the side information size will be significant and is compressed using Arithmetic coding.

3.3.6 Extraction and Recovery

As discussed above, the RDH technique relies on its ability to ensure lossless extraction of the secret message and recovery of the original image. For this, the proposed method embeds all the required side information inside the image itself. At the receiving end, the receiver, first of all, reads the LSBs of the first 2 pixels to get the value of N_B and then reads the LSBs of the next 18 pixels to get the length of side information (L_{SI}). Afterward, all the side information, which includes location map, secret data length, D_s and p_{max} obtained in the last iteration, and the number of iterations (S), are extracted. Following this, the extraction of secret data and recovery of the original cover image is performed using the following steps:

Step 1: The stego-image is partitioned into regions of interest (ROI) and non-regions of interest (NROI) with the help of the location map.

Step 2: To obtain the hidden secret data from the NROI region, the N_B least significant bits (LSBs) of the pixels in the NROI region are extracted.

Step 3: The secret data is extracted from the ROI region, and the pixels are recovered to their original value using the following Eqs. (3.6) and (3.7), respectively.

$$b_i = \begin{cases} 0 & \text{if } R_E(i,j) = p_{max} \\ 1 & \text{if } R_E(i,j) = p_{max} + D_s \end{cases} \tag{3.6}$$

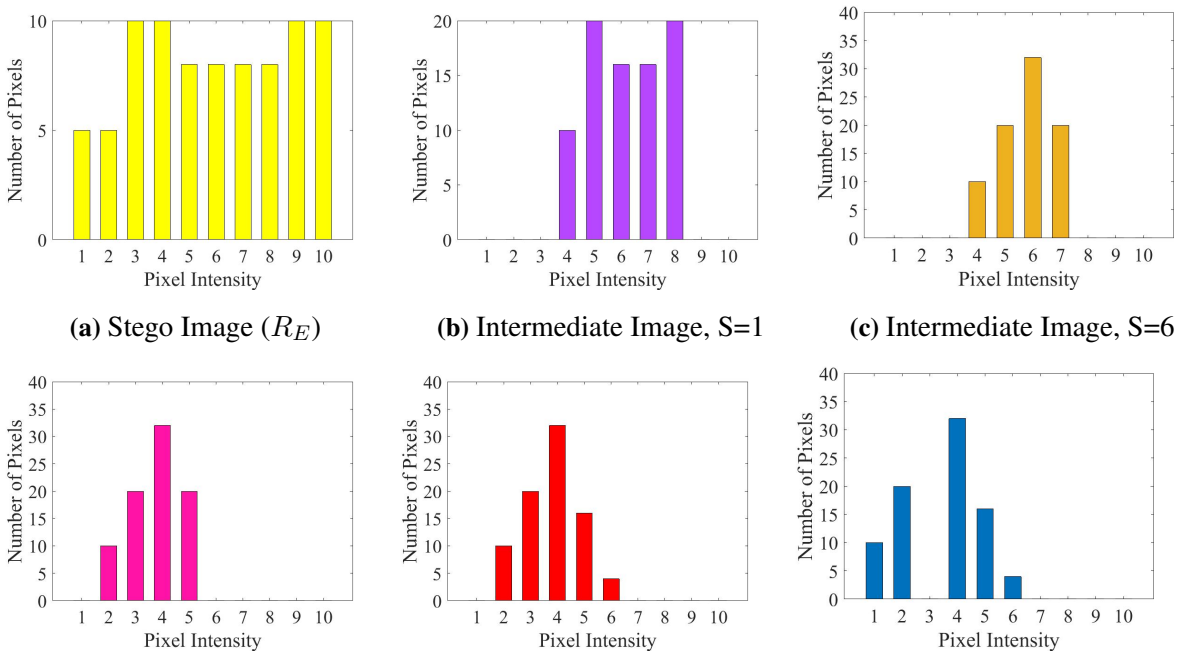
$$R_P(i,j) = \begin{cases} R_E(i,j) + 1 & \text{if } R_E(i,j) < p_{max}, D_s = -1 \\ R_E(i,j) - 1 & \text{if } R_E(i,j) > p_{max}, D_s = 1 \\ R_E(i,j) & \text{if } R_E(i,j) > p_{max}, D_s = -1 \\ R_E(i,j) & \text{if } R_E(i,j) < p_{max}, D_s = 1 \\ R_E(i,j) & \text{if } R_E(i,j) = p_{max} \end{cases} \tag{3.7}$$

Step 4: Step 3 is repeated for the S number of iterations to get the pre-processed image R_P and the complete secret data is recovered simultaneously.

Step 5: Recover the final shifting value ($toshift$) and its direction from the auxiliary information. Perform shifting by value ($toshift$) in the opposite direction using Eq. (3.8).

$$R_P(i,j) = R_P(i,j) + toshift \tag{3.8}$$

Step 6: Recover the bin value n, Direction of overlapping (O_D), and the bit stream(B_S) from the auxiliary information. Raster scan the image, obtained using step 5, and de-overlap



(a) Stego Image (R_E) **(b)** Intermediate Image, S=1 **(c)** Intermediate Image, S=6

(d) Shifting ($toshift$=-2) **(e)** De-overlapping, n=5 **(f)** Shifting (Extracted Image)

Fig. 3.4: Example histograms generated during extraction and recovery of ROI

the bins n and n-1/n+1 by using Eq. (3.9)

$$R_P(i,j) = \begin{cases} R_P(i,j) & \text{if } B_S(K) = 1, R_P(i,j) = n \\ R_P(i,j) + 1 & \text{if } B_S(K) = 0, R_P(i,j) = n, O_D = -1 \\ R_P(i,j) - 1 & \text{if } B_S(K) = 0, R_P(i,j) = n, O_D = 1 \\ R_P(i,j) & \text{otherwise} \end{cases} \tag{3.9}$$

Step 6 is repeated for $S - Ebins_{[0,255]}$ number of times.

Step 7: Shift the empty bins to their original positions with the help of E_P using Eq. (3.10).

$$R(i,j) = \begin{cases} R_P(i,j) + 1 & \text{if } R_P(i,j) >= E_P, E_P >= 127 \\ R_P(i,j) - 1 & \text{if } R_P(i,j) <= E_P, E_P <= 127 \\ R_P(i,j) & \text{otherwise} \end{cases} \tag{3.10}$$

Step 7 is repeated for all empty bins E_P, and the resultant image replicates the original image.

To ease the understanding of the image recovery steps, the same example histogram (illustrated in Fig. 3.3) is extended for the extraction and recovery analysis and is illustrated in Fig. 3.4. Fig. 3.4a represents the received stego image. From the retrieved auxiliary information, S = 6, p_{max} = 2, and D_s = -1 are obtained. For the first iteration, raster scanning of the entire image is performed, and the secret data is recovered using Eq. 3.6. Accordingly, if R_E = 2,

then b_i = 0 and if R_E = 1, then b_i = 1. Simultaneously, an intermediate pre-processed image is obtained for S = 1 using Eq. 3.7. As D_s = -1, R_P = R_E, for pixels $R_E \geq 2$; and R_P = 1 - D_s, for pixels $R_E < 2$. The resultant intermediate image histogram is illustrated in Fig. 3.4b. As S=6, Step 3 is repeated for 6 iterations, and the resultant histogram is presented in Fig. 3.4c.

Once the secret data is completely recovered from the stego image R_E, and the pre-processed image R_P is obtained after completing 6 iterations, the next step is to reverse the pre-processing to fully recover the original image.

In the illustrated example, at Step 5, the value of $(toshift)$ is -2. The updated histogram obtained after modifying the pixels R_P as per Eq. (3.8) is illustrated in Fig. 3.4d. In Step 6, the recovered overlapped bin, $n = 5$, and overlapping direction, $O_D = -1$. Thus, with the help of the stored bit stream B_S, de-overlapping is performed using Eq. (3.9). The resulting image histogram is presented in Fig. 3.4e, wherein bin 6 is de-overlapped from bin 5.

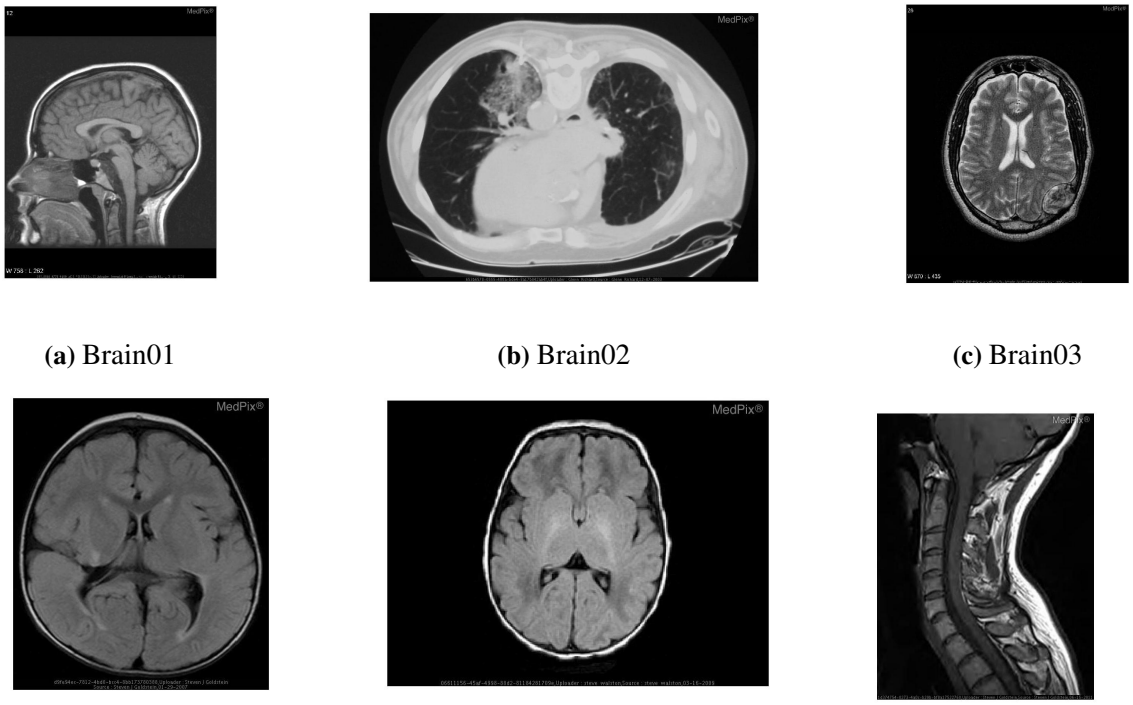
In Step 7, positions E_P of 5 empty bins are extracted as (3, 7, 8, 9, 10) from the auxiliary information. According to the shifting procedure defined in Eq. (3.10), pixels with values 2 and 3 are shifted to 1 and 2, respectively, creating an empty bin at 3. As empty bins 7, 8, 9, and 10 are already at the extreme right, these bins have no change concerning Eq.3.10. From the resultant example histogram (Fig. 3.4f) and original example histogram (Fig. 3.3a), it is evident that the original image is completely recovered from the stego image.

3.4 Experimental Evaluation

This section discusses the experimental results of the proposed method and their comparison with the SOTA methods on six different medical images, namely Brain01, Brain02, Brain03, Brain04, Brain05, and Chest1, as shown in Fig. 3.5. The images are taken from Medpix [2] and NBIA [3], which are free online medical image databases. The testing environment for conducting the experiments is MATLAB 2021, running on Windows 10, Intel(R) Core (TM) i5 2.60 GHz CPU, and 8.0 GB RAM.

For the performance evaluation and comparisons, popular and widely used metrics such as peak signal-to-noise ratio (PSNR), structural similarity index (SSIM), relative contrast error (RCE), and relative mean brightness error (RMBE) [1] have been considered. In the aforementioned metrics, PSNR and SSIM are used to evaluate the resultant image quality against the original image, where the higher the value of PSNR, the lower the objective difference between the original and the resultant image. Similarly, the value of SSIM closer to 1 indicates more structural similarity between the original and the resultant image. To measure the brightness, the Relative mean brightness error (RMBE) is used, which varies between 0 and 1. The RMBE calculates the difference between the mean brightness of the original image and the enhanced image. It is to be noted that the RMBE closer to 1 indicates an insignificant difference between the mean brightness of the original image and the enhanced image.

During the analysis of the SOTA methods, it has been observed that for some of the images,



(a) Brain01 **(b)** Brain02 **(c)** Brain03

(d) Brain04 **(e)** Brain05 **(f)** chest1

Fig. 3.5: Six medical test images from Medpix [2] and NBIA [3]

the data embedded in the ROI region is negligible in comparison to the data embedded in the NROI region. Accordingly, the quality parameters for these images, when calculated for the entire image, will be misleading, as the impact of the embedding on the ROI region remains insignificant due to low embedding in ROI. Hence, we have performed the comparative analysis in two separate sub-sections. In the Sub-section 3.4.1, the embedding performance in the entire image is discussed, whereas the embedding performance in ROI is separately evaluated in Sub-section 3.4.2. Furthermore, we have also evaluated the robustness of the proposed method using RS steganalysis in Sub-section 3.4.3. We then conducted a computational complexity analysis of the proposed approach compared to the SOTA methods in Subsection 3.4.4.

3.4.1 Performance Analysis when Embedding is Performed in Entire Image

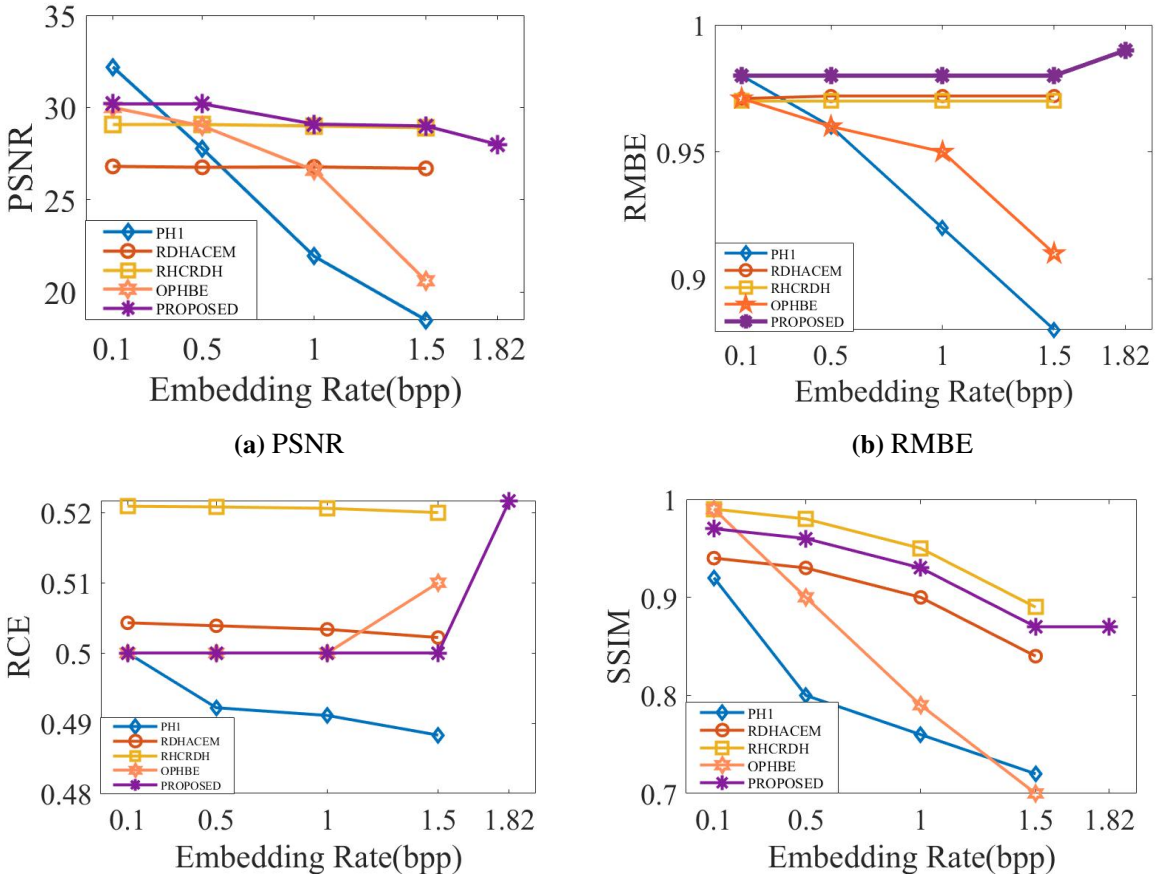
In this sub-section, we first assess the contrast of the stego-images obtained after embedding by the proposed method and the SOTA methods, such as PH1 [120], RHCRDH [122], RDHACEM [1], and OPHBE [116]. For this, the test image 'Brain01' has been considered, and results have been taken at an ER of 1.5 bpp. The resultant stego-images are shown in Fig. 3.6. From visual inspection, it is evident that the stego-images of three methods, namely RDHACEM [1], RHCRDH [122], and the proposed method, have achieved contrast enhancement, while PH1 [120] has the poorest contrast. Furthermore, it can also be observed that PH1 and OPHBE also

lack brightness preservation.

(a) PH1 **(b)** RDHACEM **(c)** RHCRDH **(d)** OPHBE **(e)** Proposed

Fig. 3.6: Stego-images of Brain01 at 1.5 bpp.

To showcase the performance of the proposed method on other evaluation parameters such as PSNR, SSIM, RCE, and RMBE, the analysis is performed on test images Fig. 3.5a to 3.5d, and experimental results are provided in Fig. 3.7 to Fig. 3.10. A comprehensive analysis and individual assessment for each image is provided below.



(a) PSNR **(b)** RMBE

(c) RCE **(d)** SSIM

Fig. 3.7: Performance comparison for 'Brain01'

For the test image 'Brain01', it is evident from Fig. 3.7 that the proposed method achieves the highest PSNR in comparison to the aforementioned SOTA methods. Further, the maxi-

Table 3.1: Experimental Results for the test image 'Brain01'

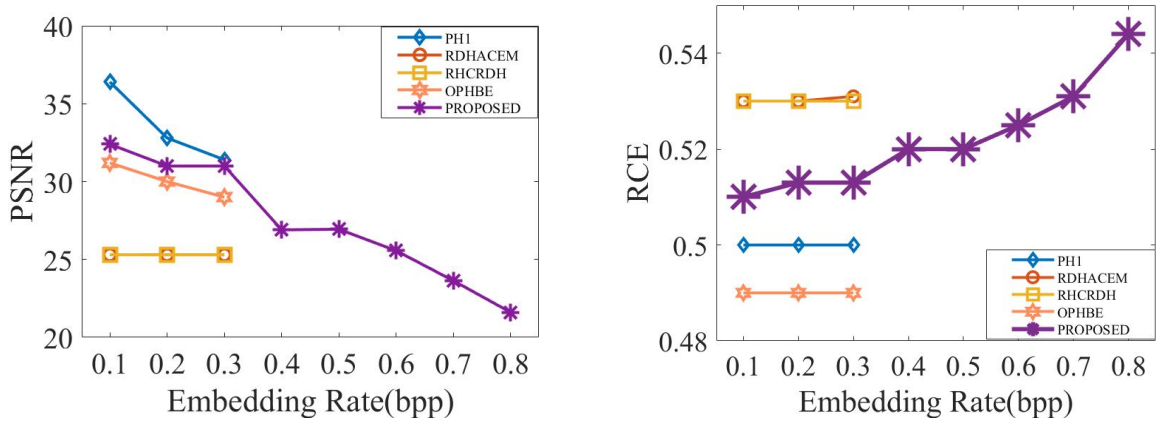
Method	bpp	PSNR	SSIM	RCE	RMBE	Time(s)	Average time(s/bit)
PH1	0.1	32.2	0.92	0.5	0.98	0.38	
	0.5	27.7	0.80	0.49	0.96	0.49	8.49×10^{-7}
	1.0	21.9	0.76	0.49	0.92	0.63	
	1.5	18.4	0.72	0.48	0.88	0.70	
RDHACEM	0.1	26.8	0.94	0.50	0.97	4.14	
	0.5	26.7	0.93	0.50	0.97	4.11	7.39×10^{-6}
	1.0	26.7	0.90	0.50	0.97	4.19	
	1.5	26.7	0.84	0.50	0.97	5.11	
RHCRDH	0.1	29.0	**0.99**	0.52	0.97	5.75	
	0.5	29.0	0.98	0.52	0.97	6.00	1.06×10^{-5}
	1.0	29.0	0.95	0.52	0.97	6.16	
	1.5	28.9	0.89	0.52	0.97	6.38	
OPHBE	0.1	30.0	0.99	0.50	0.99	1.26	
	0.5	29.0	0.90	0.50	0.98	1.61	2.47×10^{-6}
	1.0	26.6	0.79	0.50	0.95	1.62	
	1.5	20.0	0.70	0.51	0.91	1.92	
	1.6	17.24	0.64	0.47	0.87	1.99	
Proposed	0.1	30.2	0.97	0.50	0.98	0.76	
	0.5	30.2	0.96	0.50	0.98	0.83	
	1.0	**29.1**	0.93	0.50	0.98	0.80	1.22×10^{-6}
	1.5	**29.0**	0.87	0.50	0.98	1.00	
	1.8	**28.0**	0.87	**0.52**	**0.99**	1.50	

Table 3.2: Experimental Results for the test image 'Brain02'

Method	Capacity	PSNR	SSIM	RCE	RMBE	Time(s)	Average time(s/bit)
PH1	0.1	36.4	0.98	0.50	**0.98**	0.33	
	0.2	32.8	0.97	0.50	0.97	0.42	2.23×10^{-6}
	0.3	31.4	0.96	0.50	0.97	0.41	
RDHACEM	0.1	25.3	0.99	0.53	0.95	10.2	
	0.2	25.3	0.99	0.53	0.95	9.72	5.57×10^{-5}
	0.3	25.3	0.99	0.53	0.95	9.68	
RHCRDH	0.1	25.3	0.99	0.53	0.95	8.83	
	0.2	25.3	0.99	0.53	0.95	9.00	7.53×10^{-5}
	0.3	25.3	0.97	0.53	0.95	9.00	
OPHBE	0.1	31.2	0.98	0.49	0.98	1.68	
	0.2	30.0	0.97	0.49	0.98	1.58	8.59×10^{-6}
	0.3	29.0	0.96	0.49	0.97	1.98	
Proposed	0.1	32.4	0.99	0.51	**0.98**	0.60	
	0.2	31.0	0.99	0.51	0.97	0.63	3.13×10^{-6}
	0.3	31.0	0.99	0.51	0.97	0.63	
	0.4	**26.9**	**0.98**	0.52	0.96	0.78	
	0.8	**21.6**	**0.94**	**0.54**	0.95	0.90	

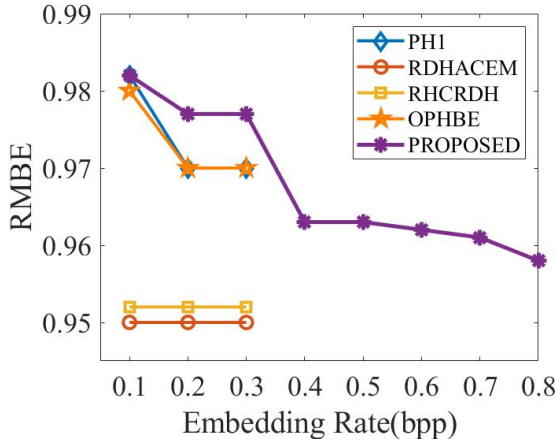
mum achievable ER with the proposed method is 1.82 *bpp* while maintaining good contrast and PSNR, whereas, for the other existing methods, the maximum achievable ER is only up to 1.5 *bpp*.

As evident from Table 3.1, for the 'Brain01' image, RDHACEM and RHCRDH methods provide constant PSNR and RCE irrespective of ER. OPHBE achieves higher PSNR at lower ER and SSIM, however, the image quality significantly deteriorates at higher ER. Further, it should be noted that though OPHBE [116] improves the contrast at lower ER, it usually results with over-enhanced stego-images at higher ERs and also introduces additional distortion. The proposed method achieves comparable RCE at lower ER; however, it significantly improves at high ER. From Fig. 3.7b, it is discernible that, with the highest RMBE, the proposed method performs better than RDHACEM and RHCRDH methods in preserving the image brightness, which further improves with the increase in ER.

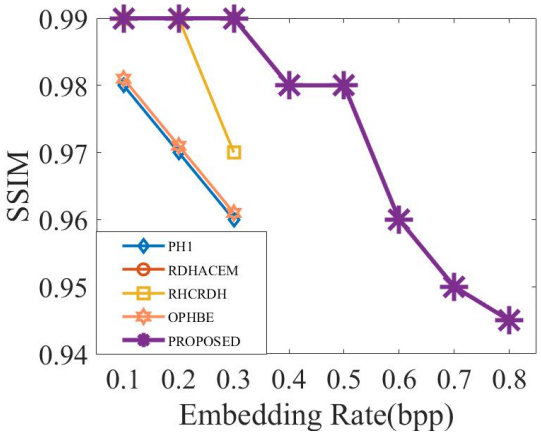


(a) PSNR

(b) RCE



(c) RMBE

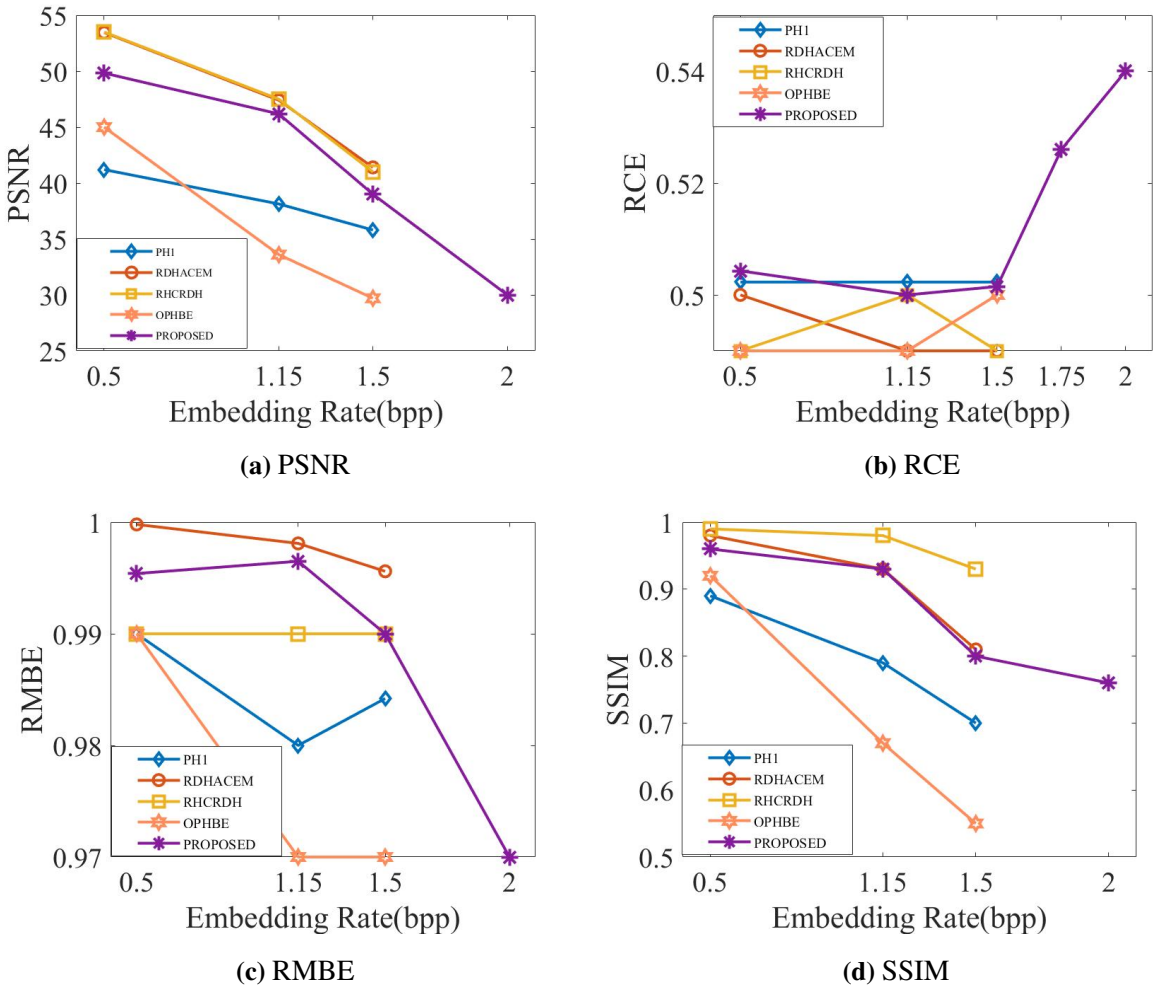


(d) SSIM

Fig. 3.8: Performance comparison for 'Brain02'

For image 'Brain02', as evident from Table 3.2, the proposed method, OPHBE, and PH1 achieve the highest RMBE. Moreover, RDHACEM and the proposed method achieve the highest SSIM that remains stable even with the change in ER. At lower embedding rates, RD-

HACEM and RHCRDH achieve the highest RCE. However, the maximum achievable ER is limited to 0.4 *bpp* only. PH1 and OPHBE methods provide the lowest contrast enhancement, wherein the OPHBE method provides negative contrast enhancement. In contrast, the proposed method not only achieves an ER of 0.9 *bpp*, but the image contrast also improves with the increase in ER with an excellent RCE of 0.54 at 0.8 *bpp*.

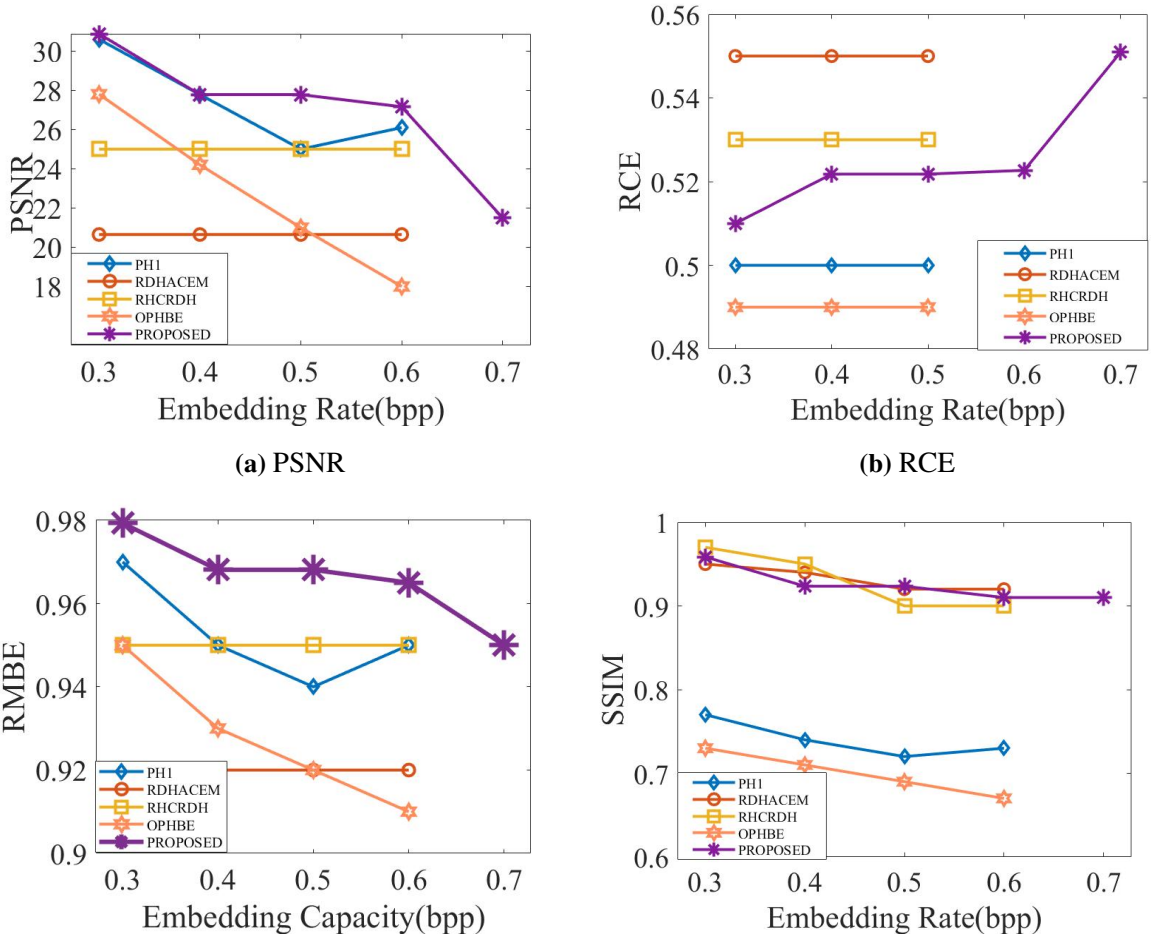


(a) PSNR **(b)** RCE

(c) RMBE **(d)** SSIM

Fig. 3.9: Performance comparison for 'Brain03'

Further extending our analysis to the 'Brain03' image, it can be observed from Fig. 3.9. that the proposed method achieves the best RCE and RMBE among all the compared methods. The RCE value obtained by RDHACEM and RHCRDH is only 0.49, which means the image is negatively enhanced. The OPHBE method improves the contrast enhancement at higher embedding rates but at the cost of poor PSNR, RMBE, and SSIM. It shows that the aforementioned SOTA methods do not analyze the characteristics of the histogram distribution well and, hence, do not improve the contrast. Additionally, the SOTA methods have limited ER, i.e., up to 1.5 *bpp*, whereas the proposed method can achieve a high ER of 2 *bpp* with high contrast enhancement at an RCE of 0.55. Similarly, for the 'Brain04' image, the proposed method achieves the highest RMBE and SSIM at all the embedding rates. Though RDHACEM and RHCRDH



(a) PSNR **(b)** RCE

(c) RMBE **(d)** SSIM

Fig. 3.10: Performance comparison for 'Brain04'

methods achieve good contrast at lower embedding rates, it is still comparable with the proposed method at the ER of 0.7 *bpp* and further improves at higher ER. Thus, it can be concluded that the proposed method's performance is superior to all the aforementioned SOTA methods. More specifically, the proposed method can achieve the highest EC and contrast enhancement for all the studied medical images of varying image characteristics in nature. Additionally, the proposed method preserves the brightness for all the images while achieving comparable PSNR. The SOTA methods, RDHACEM and RHCRDH, appear to perform well at lower embedding rates, but these methods have limited EC overall. Moreover, in some scenarios, the SOTA methods negatively impact the contrast. In the next sub-section, an analysis of performance when embedding is done only in the ROI is provided.

3.4.2 Performance Analysis when Embedding is Performed only in the ROI

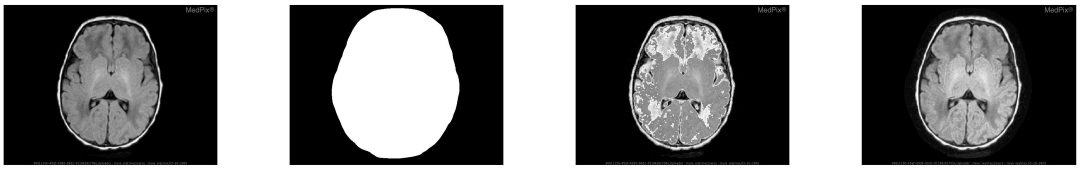
In this section, we evaluate the performance of the proposed method against the SOTA methods when the embedding is done only in the ROI region of the images. For this, two test images, namely 'Chest1' and 'Brain05' as shown in Fig. 3.11a and Fig. 3.12a, have been considered. Firstly, the images are partitioned into ROI and NROI regions using the ATD [130] to get the threshold images as shown in Fig. 3.11b and Fig. 3.12b, where the white region represents the ROI and the black region represents the NROI.

(a) Orignal **(b)** Threshold **(c)** RDHACEM **(d)** Proposed

Fig. 3.11: Experimentation with the test image 'Chest1'

Table 3.3: Experimental Results for the test image 'Brain05'

Method	BPP	RMBE	PSNR	SSIM	RCE
RDHACEM	0.1	0.95	17.5	0.83	0.56
	0.2	0.95	17.5	0.83	0.56
	0.3	0.95	17.5	0.83	0.56
	0.4	0.95	17.5	0.83	0.56
RHCRDH	0.0001	1.00	∞	1.00	0.50
Proposed	0.1	0.92	18.5	0.86	0.57
	0.2	0.92	18.6	0.86	0.57
	0.3	0.92	18.6	0.86	0.57
	0.4	0.92	18.2	0.86	0.57
	0.5	0.92	18.4	0.86	0.57
	0.6	0.92	18.5	0.86	0.57
	0.7	0.93	18.7	0.87	0.57
	0.8	0.93	18.7	0.87	0.54
PH1	-	-	-	-	-



(a) Orignal **(b)** Threshold **(c)** RDHACEM **(d)** Proposed

Fig. 3.12: Experimentation with the test image 'Brain05'

Afterward, embedding is performed in the ROI regions of the image. The resultant images after embedding the maximum possible bits of the secret data by RDHACEM [1] and the proposed method are shown in Fig. 3.11c, 3.12c, and Fig. 3.11d, 3.12d, respectively, for the 'Chest1' and 'Brain05'. The experimental results show that the proposed method can embed 1,52,357 and 1,67,918 bits into the ROI of 'Chest1' and 'Brain05'. In comparison, RDHACEM can embed only 1 and 55,513 bits in the images 'Chest1' and 'Brain05', respectively, as there are no or very minimal empty bins. However, this minimal embedding capacity of the RDHACEM in the ROI helps in keeping very low distortion but also results in negative RCE. In contrast, regardless of the number of empty bins, the proposed method has the ability to embed a large number of secret data bits in the region of interest (ROI). Additionally, as can be observed in Figures 3.11d and 3.12d, it significantly improves the contrast of the images.

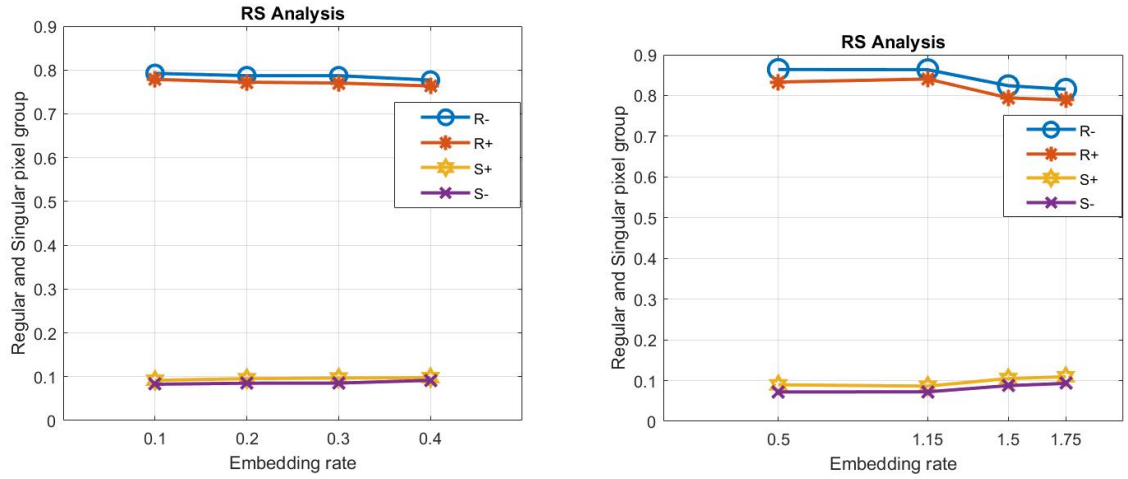
Table 3.4: Experimental Results Comparison for the Test Image 'Chest1'

Method	BPP	RMBE	PSNR	SSIM	RCE
RDHACEM	0.0001	1	∞	1	0.49
RHCRDH	0.0001	1	∞	1	0.49
Proposed	0.6	0.95	23.9	0.93	0.53
PH1	-	-	-	-	-

To perform a detailed analysis of the performance, we have included the experimental results for various metrics such as ER (in *bpp*), RMBE, PSNR, SSIM, and RCE in Table 3.3 and Table 3.4 for the test images 'Brain05' and 'Chest1', respectively. The proposed method achieved an RCE of 0.57 and 0.54 for the two images, whereas RDHACEM achieved an RCE of 0.56 and 0.49, respectively. Based on the results, it can be concluded that the proposed method offers better embedding capacity while providing better contrast.

3.4.3 Security Analysis

In this sub-section, we evaluate the robustness of the proposed method by employing steganalysis attacks that can detect if there is any hidden data within a stego-media (image, video, or audio) or not. For this, RS steganalysis [132, 133]has been chosen as it is the most widely used for security analysis in the RDH domain. RS steganalysis involves classifying pixels of an image into three categories: Regular pixels (R+ and R-), Singular pixels (S- and S+), and unusable pixels. If the pixels R+ are closer to pixels R- and pixels S+ are closer to pixels S-, the method concludes with no hidden data within the image. In another way, the farther the distance between R+ and R-, and S+ and S-, the method concludes with identifying hidden data within the image.



(a) Brain02 **(b)** Brain03

Fig. 3.13: RS-Analysis Results for the Proposed Method

The resistance against statistical RS-analysis is evident in the results depicted in Fig. 3.13 for the two test images, namely Brain02 and Brain03. In Fig.3.13a-3.13b, the x-axis shows ER in terms of *bpp* within the stego-image, while the y-axis displays the relative percentage of regular (R+, R-) and singular (S+, S-) groups. The figures show that the stego-images generated by the proposed method possess comparable values for both singular and regular parameters, even when the ER is increased from 0.1 to 1.5 *bpp*. This indicates that the proposed method has excellent protection against RS steganalysis.

Therefore, it can be concluded that the findings from this study underscore the robustness and effectiveness of the proposed RDH approach against the prominent RS steganalysis.

3.4.4 Computational Complexity

As explained in Section 3.3, the proposed method broadly comprises two stages: pre-processing and data embedding, where embedding is first done in ROI and then NROI. The ROI embedding is performed in peak bins iteratively, while the NROI embedding is done by replacing

pre-defined LSBs with the secret data bits. Let us assume that the cover image contains p pixels, ROI contains r pixels, and the NROI contains $(p-r)$ pixels. Assuming S is the total number of iterations performed to embed the data in ROI, the computational complexity for data embedding in ROI will be $\mathcal{O}(r \times S)$, as all the pixels of ROI are scanned for data embedding and histogram shifting. However, since the NROI embedding does not involve multiple iterations, the computational complexity for data embedding in NROI is $\mathcal{O}(p - r)$.

Furthermore, we have evaluated the performance of the proposed method in terms of the overall running time (seconds) and average running time per embedded bit (seconds) with the SOTA methods. For this, the results are provided in Table 3.1 and Table 3.2 for multiple test images. It is evident from the tables that the proposed method exhibits a shorter running time compared to the method discussed in [120] [122] [1], indicating that it boasts a lower implementation complexity and is well-suited for real-world applications. However, its performance is a bit slower in comparison to the PH1 [120]

Therefore, it can be clearly stated that the proposed method outperforms the aforementioned SOTA methods on almost all the performance metrics. However, its performance is somehow interlinked with the segmentation approach, i.e., ATD. Moreover, the proposed method needs to find the optimal number of iterations as a large number may introduce big distortion, while a low number may not provide sufficient embedding capacity.

3.5 Summary

This chapter presents a novel RDH method with high-embedding capacity, contrast enhancement, and brightness preservation for medical images. The proposed method first partitions the original images into ROI and NROI, embedding secret data based on their sensitivity. Prior to embedding in ROI, a novel pre-processing technique is employed to adaptively create and shift empty bins, providing high EC while reducing distortion and avoiding issues like overflow and underflow. Consequently, the proposed method achieves less overall distortion and produces higher-quality images compared to the SOTA methods. Experimental results also confirm its superiority, demonstrating higher PSNR, RCE, RMBE, and EC compared to existing methods.

CHAPTER 4

REVERSIBLE DATA HIDING FOR COLOR IMAGES USING A NOVEL SELF-ATTENTION BASED CNN PREDICTOR AND ERROR ADJUSTMENT

In prediction-based reversible data hiding (RDH), achieving high prediction accuracy is essential for obtaining high-fidelity marked images with increased embedding capacity. To achieve it, this chapter proposes two novel techniques: a self-attention-based convolutional neural network predictor (SA-CNNP) and an error adjustment strategy for color images. The SA-CNNP effectively captures both local characteristics and global pixel dependencies, resulting in comprehensive coverage and improved prediction accuracy. The error adjustment strategy further enhances accuracy by refining the prediction errors of two color channels using the error distribution of a reference channel, thereby promoting inter-channel consistency. Experimental results demonstrate that these innovations lead to significant improvements. More specifically, the proposed SA-CNNP achieves a sharper prediction error histogram with approximately 8% improvement in MSE over the best known state-of-the-art predictor. Additionally, the error adjustment strategy increases prediction accuracy for 'Kodim09' color image by 58%. Consequently, the proposed RDH approach achieves an average PSNR gain of around 1.2 dB compared to existing methods for color images.

4.1 Introduction

In the realm of RDH for color images [134–138], there has been surprisingly limited scholarly interest despite the widespread use of color images on digital platforms. This lack of focus has led to minimal research on inter-channel correlation in color images. Notably, Li et al. [139] proposed a method leveraging similar edge information across channels to enhance embedding performance. Ou et al. [140] introduced an adaptive embedding scheme that assigns variable payload sizes to different channels based on their PEH, and then utilizes the other channels to

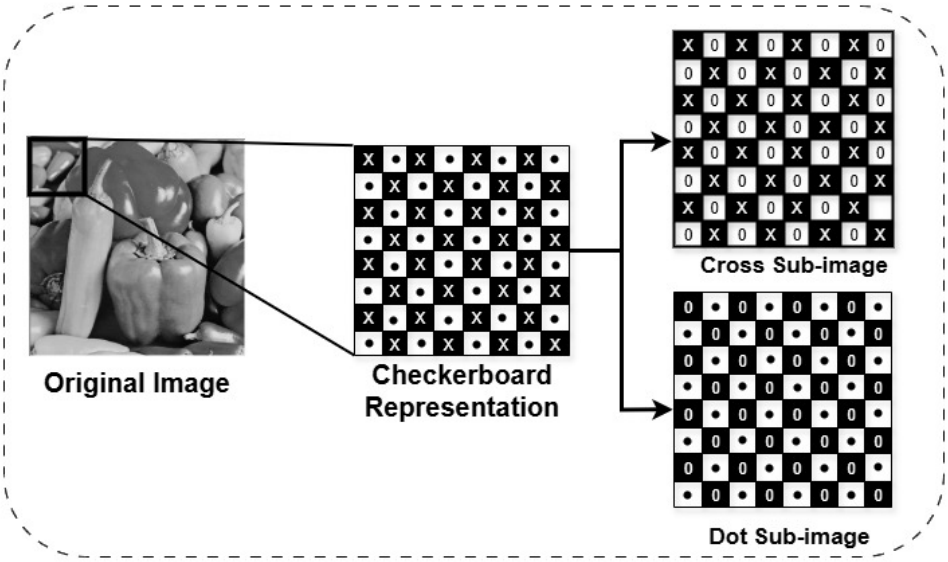


Fig. 4.1: Illustration of image partitioning into cross and dot sub-images

calculate pixel smoothness, thereby reducing embedding distortion. Yao et al. [141] developed a guided filtering-based RDH that exploits inter-channel correlation for more precise prediction, thereby improving embedding performance. Tang et al. [142] proposed a two-layer embedding scheme using image interpolation to generate prediction errors for the first level, followed by local complexity for the second level of embedding. Xu et al. [143] discussed a three-dimensional (3D) prediction error expansion-based RDH scheme that calculates error triplets for the three channels and modifies them for embedding. Recently, Mao et al. [144] introduced an RDH scheme that utilizes channel correlation to improve performance at three stages: prediction, block selection, and capacity allocation. Chang et al. [98] proposed an adaptive 3D histogram modification technique that uses PEH frequency ranking to iteratively adjust the reversible mapping, though the required optimization introduces high computational complexity.

Due to the limitations of traditional predictors, deep learning based predictors have gathered due attention in recent days. Hu et al. [80] introduced a convolutional neural network-based predictor (CNNP) that leverages the multi-receptive field and complex computation capabilities of CNN. The CNNP first partitions the original image into two independent sets as shown in Fig. 4.1 and then predicts one from the other, thus exploiting a larger set of context pixels along with the nonlinear relationship among them.

To enhance performance, Hu et al. [81] modified the image-division strategy used in CNNP training and implemented prediction error ordering (PEO) for embedding, resulting in a method known as CNNPEO. More specifically, CNNPEO divides the image into four independent sub-images, ensuring that context pixels from the other three sub-images are available to predict the target sub-image. This increased the availability of context pixels, leading to improved prediction accuracy. However, this resulted in a twofold increase in training and computation time. To further enhance performance, the image division strategy was revised to produce four

independent images, which undergo pre-processing before being input into a newly designed convolutional neural network (NCNN) model [145]. This approach provides a larger pixel context during prediction, resulting in improved prediction accuracy. Building on this idea, Wu et al. [146] introduced a novel image division pattern that increases the number of context pixels within a 5×5 block to 18, further boosting prediction performance. Continuing this trend, Luo et al. [147] enhanced the CNNP [80] model by incorporating an optimizer that refines the predicted pixel using the average of four diagonal neighbors, subsequently combining it with the original prediction to form OCNNP, which yields improved accuracy. To achieve even greater precision, Ma et al. [148] proposed a multi-scale fusion network that adaptively adjusts prediction through texture-based analysis, complemented by a multi-scale interpolation network. While these methods significantly enhance performance, they remain limited by the lack of global contextual information.

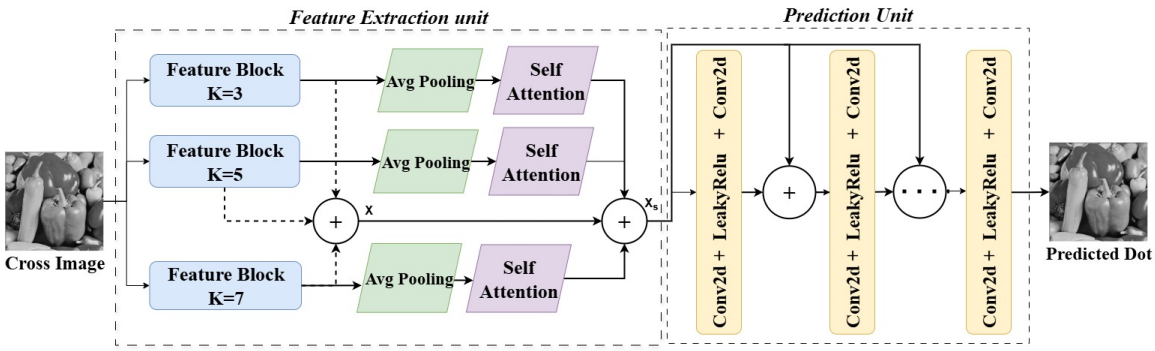
In summary, while existing predictors [149, 150] have achieved good prediction accuracy, this often comes at the expense of higher computational complexity due to added pre-processing steps. Moreover, most efforts to enhance context information have predominantly concentrated on local regions, thereby restricting the full potential of prediction accuracy. To overcome these limitations, this chapter proposes a novel self-attention-based CNN predictor (SA-CNNP), which incorporates self-attention mechanisms [151] into the CNN framework to effectively capture long-range dependencies and global contextual information.

Additionally, despite recent progress in RDH method targeting color images, the utilization of inter-channel correlation remains limited. Moreover, deep learning approaches for predicting color images have received minimal attention in this domain. To address this gap, this chapter introduces a new RDH method for color images by efficiently utilizing inter-channel correlation and integrating the SA-CNNP for improved embedding performance. The primary contributions of this chapter are outlined below:

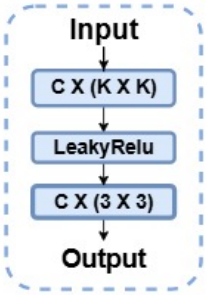
- **Introduction of SA-CNNP:** A self-attention-based CNN predictor (SA-CNNP) is introduced, which integrates self-attention mechanisms into the CNN framework. This allows the model to capture both local and global contextual information by attending to long-range pixel dependencies across the image, thereby significantly improving prediction accuracy.

- **Error Adjustment Strategy:** A novel error adjustment strategy is proposed that refines the prediction errors of two color channels by utilizing the error distribution of a reference channel. This leads to a even sharper prediction error histogram (PEH), resulting in improved embedding capacity and reduced distortion.

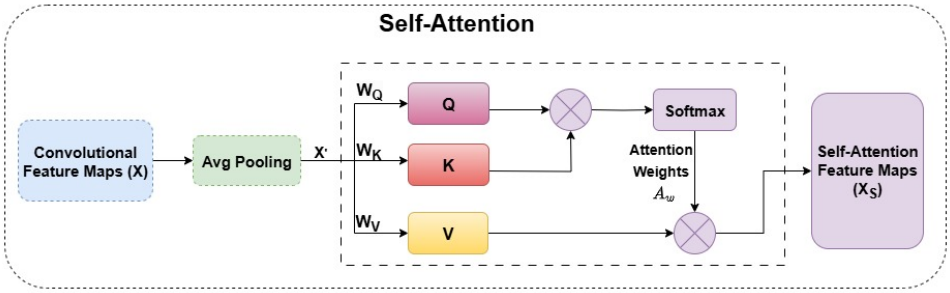
- **Performance Validation:** Experimental results demonstrate that the combination of SA-CNNP and error adjustment significantly outperforms existing state-of-the-art methods in both prediction accuracy and embedding performance.



(a) Overview of SA-CNNP



(b) Feature block



(c) Structure of the Self-attention block

Fig. 4.2: Overview of the proposed architecture (a) SA-CNNP; (b) Feature block; (c) Self-Attention block.

4.2 Proposed Method

This section is organized into three subsections. Subsection 4.2.1 discusses the proposed SA-CNNP in detail. Subsection 4.2.2 presents the newly developed RDH method for color images. Finally, Subsection 4.2.3 explains the extraction and recovery process associated with the proposed RDH method.

4.2.1 The Proposed SA-CNNP

Initially, original image (I) is pre-processed using a checkerboard pattern to generate two complementary sub-images: Cross (I_C) and Dot (I_D), as shown in Fig. 4.1. In I_C, pixels at the Cross positions retain their original values while the remaining are set to zero; conversely, in I_D, only the Dot-position pixels retain their values and the rest are set to zero. Thus, two mutually exclusive sub-images (I_C, I_D), that can be used to predict each other as the input-output training pair, are obtained.

SA-CNNP Architecture

The proposed SA-CNNP model is introduced to improve prediction quality by integrating multi-scale feature extraction with spatially adaptive attention. The architecture of the proposed SA-

CNNP is presented in Fig. 4.2a, which contains two primary units: a feature extraction unit and a prediction unit. The feature extraction unit, which is responsible for capturing features from an image, consists of three feature blocks, with kernel sizes of 3×3, 5×5, and 7×7, connected in parallel to leverage multi-receptive features. Each of the feature blocks is followed by an average pooling layer and a self-attention block. The detailed structure of the feature block is presented in Fig. 4.2b. On the other hand, the prediction unit comprises seven blocks connected in series, wherein each block consists of two convolution layers, each of kernel size 3×3, and a LeakyRelu activation function in between.

Unlike existing deep learning-based predictors [80,81,145,146], the SA-CNNP incorporates an average pooling layer along with a self-attention module, effectively capturing both local characteristics and global dependencies of pixels within an image. The detailed framework of the self-attention module is illustrated in Fig. 4.2c. The input to the self-attention block is the feature maps X, which are first passed through an average pooling layer, which helps in reducing the size of the feature maps X so that the self-attention block can effectively focus on the most relevant features with reduced computational load. The resultant feature map $X^{'}$ is then linearly transformed into three vectors: Query (Q), Key (K), and Value (V), using the following Eq. 4.1.

$$Q = X^{'}W_Q, K = X^{'}W_K, V = X^{'}W_V \tag{4.1}$$

Here, W_Q, W_K, and W_V are learnable weight matrices. The dot product of Q and K is passed through the softmax function to normalize the values, resulting in attention weights (A_w). These weights determine the extent to which each pixel's value contributes to the prediction of the current pixel. Finally, the Value (V) and attention weights (A_w) are multiplied to obtain the self-attention feature vector (X_S), which provides enhanced contextual understanding of the pixels while capturing the dependencies between distant pixels.

Training details

In order to train the proposed SA-CNN predictor, a dataset containing 3000 images is selected from the extensive ImageNet [152]. The dataset was divided into training and validation subsets using an 80:20 ratio, where 80% of the images were used for training and 20% for testing the model's performance. Before training, all the images are first transformed into grayscale images of 512×512 pixels. During the training process, the following loss function is used.

$$\text{loss} = \frac{1}{t}\sum_{i=1}^{t}(\hat{I}_i - I_i)^2 + \lambda\|\omega\|^2 \tag{4.2}$$

where t is the number of training images, I_i is the original image, $\hat{I}_i$ is the predicted image and ω represents the weights in the network. In order to mitigate over-fitting and speed up the network convergence, the weight decay (λ) is set to 10^{-3}. To optimize the proposed predictor, back-propagation and Adam optimizer are used with a batch size of 4. The proposed SA-CNNP

is trained on an Intel Core i5 CPU (2.40GHz) with 8 GB of RAM using an NVIDIA Tesla V100 GPU on Colab Pro.

4.2.2 Proposed RDH Method

Chang et al. [98] demonstrated that the channels, Red (R), Green (G), and Blue (B), of a color image are typically correlated. To leverage this correlation for improved embedding performance, the proposed RDH method exploits reference errors obtained from the highly accurate SA-CNNP. The complete workflow of the proposed RDH method for color images, illustrated in Fig. 4.3, entails three key steps: a Reference Channel Selection, Reference Channel Embedding, and Reference Error Adjustment. Each of these steps is discussed in subsequent sections.

Reference Channel Selection

The original color image (I) is initially decomposed into three separate images (I^R, I^G, and I^B), each representing one color component. Then, joint entropy, which quantifies the correlation between channels, is calculated using Eq. 4.3 across channel pairs (R-G, R-B, and G-B).

$$H(X,Y) = -\sum_{i=1}^{n}\sum_{j=1}^{m} p(x_i, y_j)\log_2(p(x_i, y_j)) \tag{4.3}$$

where $H(X,Y)$ denotes the joint entropy between two channels X and Y, n and m denote the number of unique intensity levels in X and Y, respectively, and $p(x_i, y_j)$ represents the joint probability mass function of X and Y, giving the probability that $X = x_i$ and $Y = y_i$. It is to be noted that a lower joint entropy $H(X,Y)$ between two channels signifies a higher degree of correlation. Therefore, the channel exhibiting the lowest joint entropy with the other two channels is chosen as the reference channel. This will significantly help in reducing the magnitude of prediction errors of the other two channels based on the learning of the reference channel prediction and its accuracy. To avoid redundant information and streamline the representation, the reference channel is denoted as I^r, and the other two channels are commonly denoted as I^x throughout this study.

Reference Channel Embedding

Now, the reference channel I^r is pre-processed and transformed to get the cross I^r_C and dot I^r_D sub-images as in Fig. 4.1. Next, the proposed SA-CNNP is employed on I^r_C to get the predicted sub-image $\hat{I}^r_D$ and the PE between the original and predicted sub-images are calculated as $e^r_D = I^r_D - \hat{I}^r_D$. Finally, the embedding of secret data is performed in the I^r_D using the embedding strategy of the classical RDH method [26] as detailed in Eq. 4.4 to get the marked sub-image I'^r_D.

$$I'_{i,j} = \hat{I}_{i,j} + E_{i,j} \tag{4.4}$$

where $E_{i,j}$ is the expanded error for $I(i,j)$, obtained using Eq. 4.5, which is defined as follows:

$$E_{i,j} = \begin{cases} e_{i,j} + b, & \text{if } e_{i,j} = T_p \\ e_{i,j} - b, & \text{if } e_{i,j} = T_n \\ e_{i,j}, & \text{if } e_{i,j} < T_p \text{ and } e_{i,j} > T_n \\ e_{i,j} + 1, & \text{if } e_{i,j} > T_p \\ e_{i,j} - 1, & \text{if } e_{i,j} < T_n \end{cases} \tag{4.5}$$

where T_p and T_n are the two most frequent errors, and b is the secret data bit to be embedded.

Next, the same process is repeated for I^r_C by employing the SA-CNNP model on the marked sub-image I'^r_D and calculating the PE between the predicted cross and the original cross sub-images as $e^r_C = I^r_C - \hat{I}^r_C$. Subsequently, the marked cross sub-image I'^r_C is obtained by employing Eq. 4.4. Thus, a marked reference image I'^r, which is synthesis of I'^r_D and I'^r_C, along with two error matrices (e^r_D, e^r_C) are obtained.

Reference Error Adjustment

Once the embedding for cross and dot sub-images for the reference channel is complete, the embedding for the remaining two channels is performed using the same steps, except for the reference error adjustment. It is to be noted that e^r_D and e^r_C serve as reference errors for the remaining two channels' dot and cross sub-images, respectively. In other words, the obtained PEs (during the secret data embedding process) of the remaining channels are corrected to adjust the probable prediction errors made by the SA-CNNP based on the corresponding PE of the reference channel. Therefore, the PEs for the dot and cross sub-images of the remaining channels (represented as I^x) are modified using the following Eqs. 4.6 and 4.7, respectively:

$$e'^x_D = e^x_D - e^r_D \tag{4.6}$$

$$e'^x_C = e^x_C - e^r_C \tag{4.7}$$

To theoretically validate that the prediction error from one channel supports the enhancement of prediction error for other channels, we have calculated Pearson correlation between two error matrices, as represented in Eq. 4.8. Here, ρ_{e^r,e^x} represents Pearson correlation coeeficient between PEs of reference channel r and the other channel x. $\mathrm{Cov(e^r, e^x)}$ is the covariance between e^r and e^x, σ_{e^r} and σ_{e^x} are the standard deviation of e^r and e^x, respectively. In general, ρ ranges between -1 and 1, wherein, $\rho_{e^r,e^x} > 0$ represents positive correlation between the PEs of

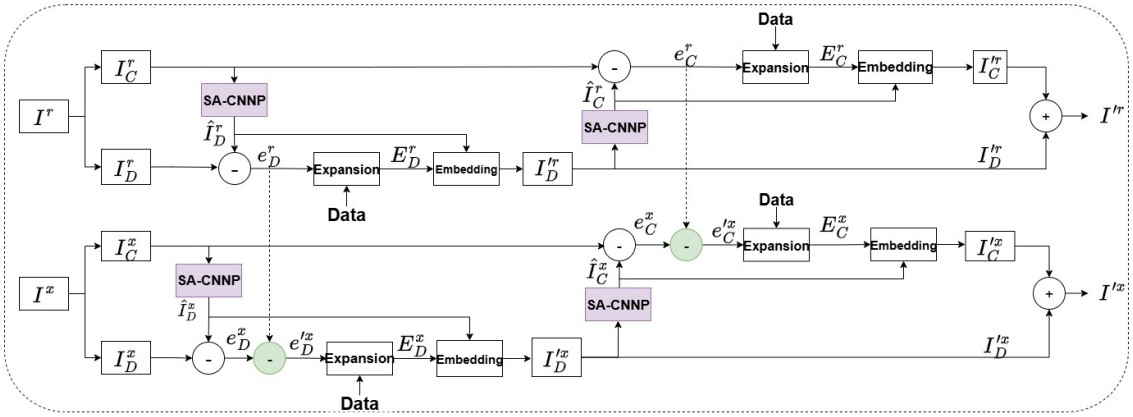


Fig. 4.3: Workflow of the proposed RDH embedding method for color images

two error channels.

$$\rho_{e^r,e^x} = \frac{\mathrm{Cov}(e^r, e^x)}{\sigma_{e^r}\,\sigma_{e^x}} \tag{4.8}$$

Next, the secret data is embedded into the remaining channels' sub-images using Eq. 4.4, and the final marked image ($I^{'}$) after synthesis of the cross set and dot set images in a channel-wise manner is obtained. The detailed algorithm demonstrating step-by-step approach of the proposed method is provided as Algorithm 1. It should be noted that, in order to focus primarily on our novel contributions, we have omitted details about auxiliary information and the location map, which are largely similar to those in [26].

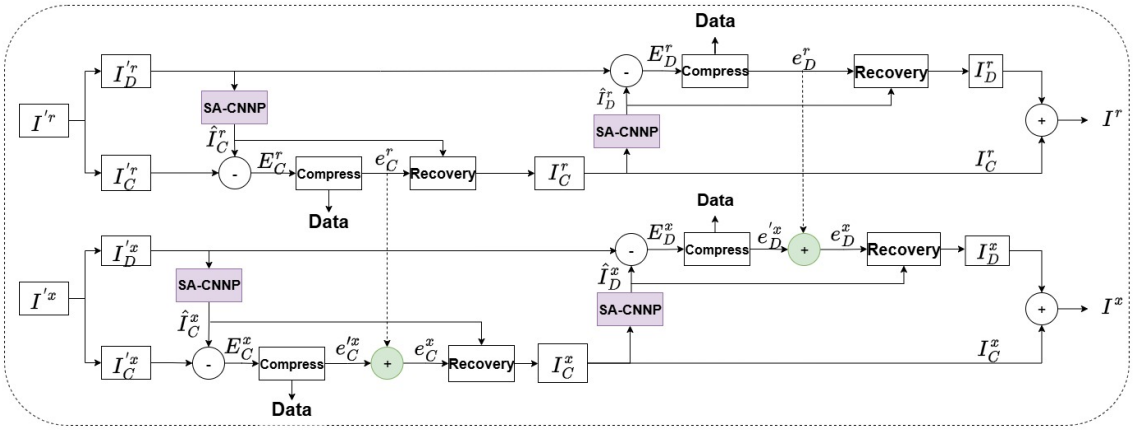


Fig. 4.4: Workflow of the proposed extraction and recovery method for color images

4.2.3 Extraction and Recovery

The extraction of secret data and recovery of the original image from the marked image ($I^{'}$) is the inverse of the embedding procedure and is depicted in Fig. 4.4. At first, similar to the embedding procedure, the reference channel $I^{'r}$ is pre-processed and transformed to get the

Algorithm 1: Embedding ALgorithm

Input: I^R, I^G, I^B (RGB channels), b (Secret data) **Output:** Embedded color image I'

Steps:

1. Compute joint entropy $H(I^R, I^G)$, $H(I^R, I^B)$, and $H(I^G, I^B)$ between channels using Eq. 4.3.
2. Select Reference channel I^r such that $H(I^r, I^{x_1})$ and $H(I^r, I^{x_2}) < H(I^{x_1}, I^{x_2})$.
3. Partition I^r into I^r_D and I^r_C as shown in Fig. 4.1
4. Predict $\hat{I}^r_D \leftarrow SACNNP(I^r_C)$
5. Calculate PE: $e^r_D = I^r_D - \hat{I}^r_D$.
6. Obtain Expanded error E^r_D using Eq. 4.5
7. Embedding: I'^r_D = $\hat{I}^r_D$ + E^r_D
8. Repeat steps 3–7 with I'^r_D as input and embedded cross I'^r_C as output
9. Obtain final Embedded Reference channel: I'^r = I'^r_D + I'^r_C
10. For the two non-reference channels x_1 and x_2, repeat steps 3–9 to obtain I'^{x_1} and I'^{x_2}, with adjustment of prediction errors $e^{x_i}_c$ and $e^{x_i}_D$ in step 5 using Eqs. 4.6 and 4.7
11. Finally, the marked color image is obtained as $I' = I'^r + I'^{x_1} + I'^{x_2}$.

cross I'^r_C and dot I'^r_D sub-images as in Fig. 4.1. Next, the proposed SA-CNNP is employed on I'^r_D to get the predicted sub-image $\hat{I}^r_C$ and the errors between the original and predicted sub-images are calculated as $E^r_C = I'^r_C - \hat{I}^r_C$. Subsequently, to reverse the expansion effect produced during the embedding procedure, the errors E^r_C are modified to generate e^r_C and extract the secret data (b) using Eq. 4.9.As given in Eq. 4.9, the conditions $E_{i,j} = T_p$ and $E_{i,j} = T_n$ correspond to the embedding rule $(e_{i,j} \pm b)$ described in Eq. 4.9, where no change in the pixel value indicates that $b = 0$. In contrast, if $E_{i,j} = T_p + 1$ or $E_{i,j} = T_n - 1$, it implies that $b = 1$, as T_p and T_n were shifted by one during the embedding process in accordance with Equation (5).

$$(b, e_{i,j}) = \begin{cases} (-,\ E_{i,j}+1) & \text{if } E_{i,j} < T_n - 1 \\ (1,\ E_{i,j}+1) & \text{if } E_{i,j} = T_n - 1 \\ (0,\ E_{i,j}) & \text{if } E_{i,j} = T_n \\ (-,\ E_{i,j}) & \text{if } T_n < E_{i,j} < T_p \\ (0,\ E_{i,j}) & \text{if } E_{i,j} = T_p \\ (1,\ E_{i,j}-1) & \text{if } E_{i,j} = T_p + 1 \\ (-,\ E_{i,j}-1) & \text{if } E_{i,j} > T_p + 1 \end{cases} \tag{4.9}$$

Finally, the original image I_C^r is recovered from the marked sub-image using Eq. 4.10. Next, the same process is repeated for $I_D^{'r}$ to extract the secret data and recover the sub-image I_D^r.

$$I_{i,j} = \hat{I}_{i,j} + e_{i,j} \tag{4.10}$$

Thus, a recovered reference image I^r which is synthesis of I_D^r and I_C^r, along with two error matrices (e_D^r, e_C^r) are obtained.

The extraction and recovery procedure for the remaining channels (I^x) is similar to that of I^r except for the reference error adjustment introduced during the embedding. Similar to the embedding procedure, the error matrices e_D^r and e_C^r, obtained during I^r recovery, serve as reference errors for the remaining two channels' extraction, and the probable PE at these two channels are corrected using Eqs. 4.11 and 4.12, respectively.

$$e_D^x = e_D^{'x} + e_D^r \tag{4.11}$$

$$e_C^x = e_C^{'x} + e_C^r \tag{4.12}$$

Once the extraction for the two remaining channels is complete, the final recovered image (I) is obtained by concatenation of the recovered images for the three channels R, G, and B.

4.3 Experimental Evaluation

This section empirically analyses the performance of the proposed predictor and our RDH method for color images. First, we evaluate the prediction accuracy of the proposed SA-CNNP. Next, the performance of the proposed RDH method is evaluated and compared with the state-of-the-art methods. For this, two well-known data sets, SIPI [69] and Kodak [153], have been considered as test images. At last, the security of the proposed method is evaluated.

4.3.1 Performance Evaluation of the SA-CNNP

To gauge the prediction accuracy of the SA-CNNP, an average PEH is plotted as shown in Fig.4.5 over a test set consisting of 32 gray-scale standard images, including 8 from the USC SIPI dataset [69] and 24 from the Kodak dataset [153]. The SIPI images are standard grayscale images with a resolution of 512×512 pixels. The Kodak images are color images with dimensions of either 512×768 or 768×512 pixels. For consistency in grayscale prediction error analysis, the color images were converted to grayscale while retaining their original dimensions.

The PEH clearly shows that the proposed predictor, with 68533 occurrences, achieves the highest peak at zero among all the existing predictors, such as CNNPEO [81], CNNP [80], DP [30], and MEDP [31]. Additionally, the counts in adjacent bins such as -1 (58377), +1 (55150), indicate highly accurate predictions. Moreover, 98% of the pixels are concentrated within the range of [-2, 2]. In other words, the SA-CNNP increases the pixel concentration within the range of [-2,2] by 13% and 20% compared to the best-performing existing predictors, CNNPEO and CNNP, respectively. The overall PEH distribution shows that the majority of prediction errors are tightly clustered around zero, highlighting the strong predictive performance and generalization capability of the proposed method across both datasets.

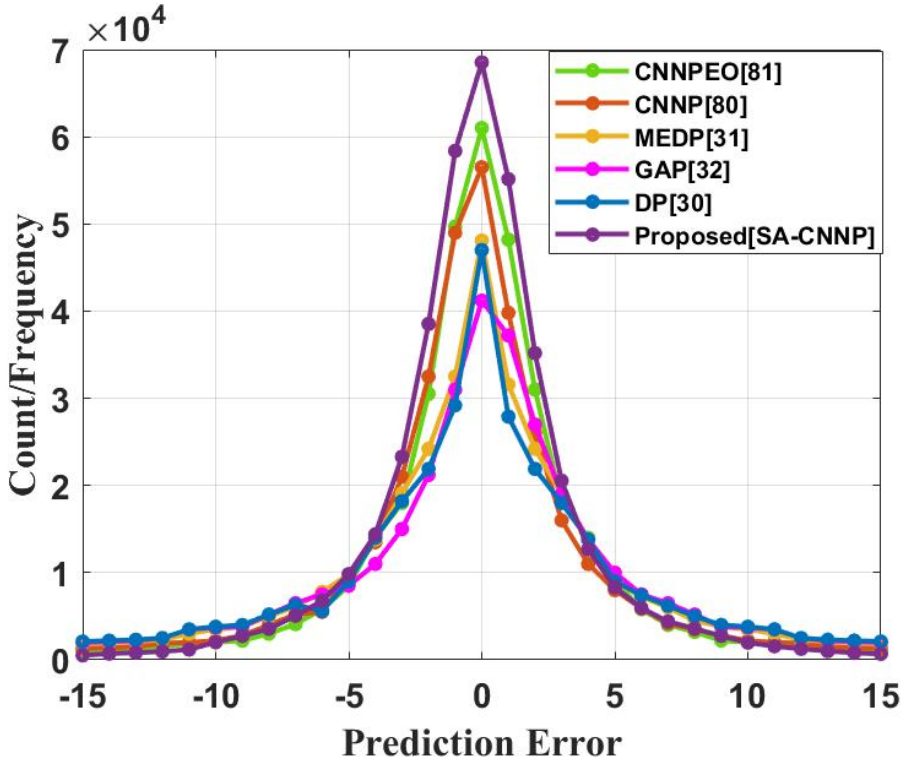


Fig. 4.5: Average PEH comparison over gray-scale images from USC-SIPI and Kodak datasets

To further augment the analysis, a color image (Airplane) from the MS COCO dataset has been used, and the prediction error histogram (PEH) comparison is presented in Fig. 4.6. The proposed SA-CNNP, combined with the reference error adjustment strategy, achieves the highest frequency at zero prediction error, showing an improvement of approximately 110% over

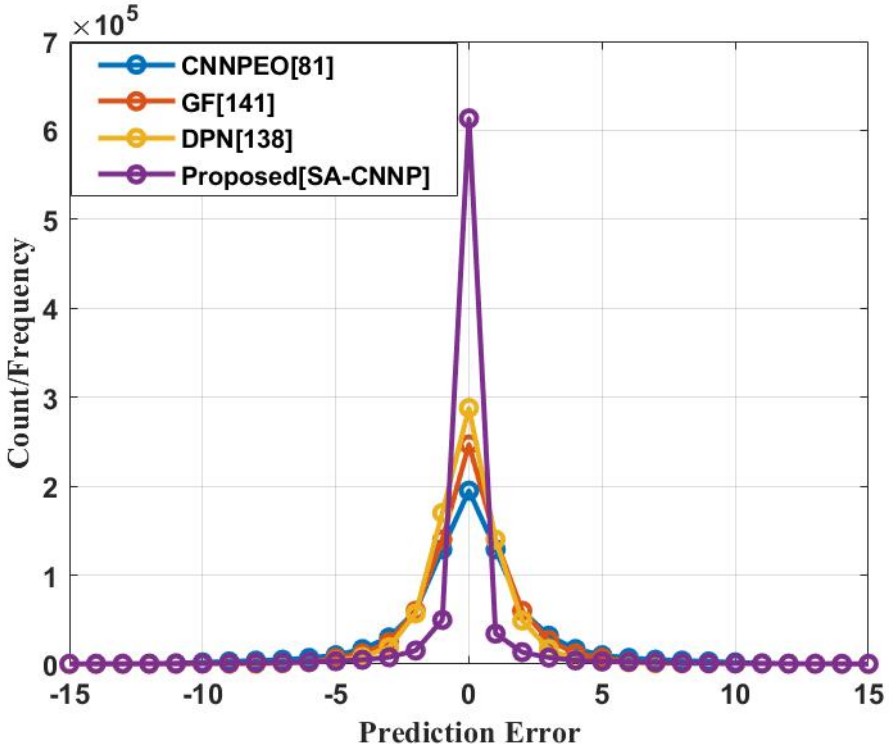


Fig. 4.6: PEH comparison for color image (Airplane) from MS-COCO dataset

Table 4.1: Comparison of Average Mean, Variance, and MSE on USC-SIPI dataset

Metric	Methods			
	[145]	[81]	[80]	Proposed
Mean	3.29	3.31	4.21	2.99
Variance	27.56	20.29	32.63	19.27
MSE	31.86	34.24	52.47	29.44

the best-performing existing method. The error distribution remains tightly concentrated around zero, indicating more accurate and stable predictions.

To further evaluate the effectiveness of the proposed SA-CNNP, we analyzed quality parameters such as mean, variance, and mean square error (MSE) of the absolute prediction errors (PEs) and compared them with state-of-the-art (SOTA) predictors for the standard SIPI test images [69]. As shown in Table 4.1, the proposed SA-CNNP achieves the lowest average mean of 2.99, indicating that the PEs generated by the SA-CNNP are significantly smaller compared to those produced by SOTA methods. Additionally, the average variance of 19.27 suggests that the PEs are tightly clustered, reflecting minimal volatility in prediction performance. Furthermore, the MSE, which measures the squared difference between actual and predicted values, confirms the high prediction accuracy of the proposed predictor. These results underscore the superior performance of the SA-CNNP in comparison to existing SOTA methods.

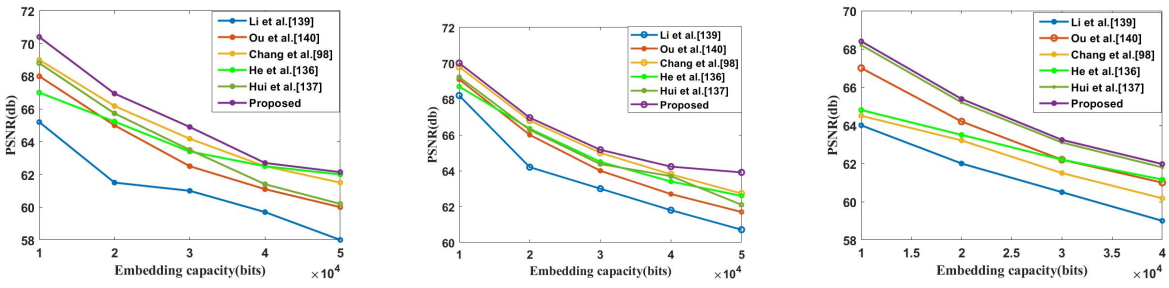
Table 4.2: Average PSNR (dB) comparison over 32 test images at different payloads, highlighting the performance of the proposed method and its ablation setups in relation to existing approaches.

Methods	Payload (bits)				
	10,000	**20,000**	**30,000**	**40,000**	**50,000**
RP [26]	57.0	54.1	51.1	50.9	48.6
CNNP [80]	58.4	55.1	52.9	51.2	49.8
CNNPEO [81]	61.17	57.56	55.43	53.84	52.24
NCNN [145]	61.48	57.86	55.79	54.13	52.73
UCANet [149]	58.25	55.75	54.02	52.07	50.21
Ablation Setup-1	60.27	56.42	55.27	54.07	52.58
Ablation Setup-2	61.95	57.69	55.86	54.14	52.90
Proposed	**62.68**	**58.72**	**56.18**	**54.66**	**53.84**

Furthermore, the embedding performance of the classical prediction-based RDH method [26] was evaluated by integrating it with the SA-CNNP and comparing against existing predictors, such as RP [26], CNNP [80], CNNPEO [81], NCNN [145], and UCANet [149]. Using a comprehensive test set of 32 images comprising 8 SIPI [69] images and 24 Kodak images [153], Table 4.2 presents the average PSNR values at various payloads. The results indicate that the classical RDH method, when combined with the proposed SA-CNNP, achieves an average PSNR value of approximately 1.2 dB higher than when combined with the SOTA predictors. This demonstrates the superior embedding performance of the SA-CNNP-enhanced RDH method.

Table 4.3: Count of prediction errors before and after error adjustment with % change for Kodim09

Prediction Error (E)	E^r (Reference Channel) G	E (Before Adjustment)		E' (After Adjustment)		Change (%)	
		R	**B**	**R**	**B**	**R**	**B**
-2	11788	11807	12667	4165	6131	-64.72	-51.60
-1	34151	34733	38589	30888	38862	-11.07	0.71
0	119624	117187	114153	188519	177413	60.87	55.42
1	33817	35331	33186	32321	30696	-8.52	-7.50
2	11874	12103	12179	4144	5403	-65.76	-55.64



(a) Airplane **(b)** House **(c)** Splash

Fig. 4.7: PSNR comparison on SIPI color images (512×512) across existing methods and the proposed approach.

To study the impact of the average pooling layer and self-attention in the CNN predictor,

an ablation study has been performed with two additional experimental setups in addition to the proposed model architecture. The ablation study set-up 1 includes a CNN predictor, with an average pooling layer added to it but without self-attention. Whereas, the ablation study set-up 2 includes a self-attention-based CNN model, without an average pooling layer. The results of the ablation study have been added in Table 4.2. The experimental results show that the proposed SA-CNNP model outperforms the ablation study set-ups 1 and 2, as well as the SOTA methods. Thus, it can be concluded that the proposed architecture effectively captures a comprehensive feature set, encompassing both local and global features, which significantly enhances prediction accuracy.

4.3.2 Performance Evaluation of the Proposed RDH Method

In this sub-section, we evaluate the efficacy of the proposed RDH scheme for color images. For this, we first assess the performance of our error adjustment strategy using the 'Kodim09' color image of size 512×512. As outlined in the proposed RDH method, the image is decomposed into three separate channels (R, G, and B). The joint entropy values, calculated using Eq. 4.3, are 10.88 for R-G, 11.50 for R-B, and 11.01 for G-B. Channel G, being common to the two pairs with the lowest joint entropy values, is selected as the reference channel for subsequent error adjustment.

The experimental results, presented in Table 4.3, show the count of PEs within the range [-2, 2] for the three channels (R, G, and B) before and after performing the error adjustment as per Eqs. 4.6 and 4.7. Before adjustment, the count of '0' PEs is 1,17,187 for R, 119624 for G, and 114153 for B. After performing the error adjustment, the count of '0' PEs for channels R and B increases to 188519 and 177,413, respectively, representing an enhancement of 61% and 55%.

To evaluate the embedding performance of the proposed RDH method for color images, the PSNR achieved by the proposed RDH method is compared against the SOTA methods by considering both SIPI and Kodak image datasets. For the wide range of embedding capacities, Fig. 4.7 represents the PSNR comparison for the three standard SIPI images. Across all three images, the proposed method consistently shows superior performance by maintaining higher PSNR values at varying embedding capacity. Specifically, for the Airplane image, the proposed method showcase remarkable improvement in embedding performance at low embedding capacity, while for the House image the proposed method demonstrate notable enhancment at higher embedding capacity, indicating better image quality preservation across entire range of embedding capacity.

Similarly, Fig. 4.8 illustrates the PSNR comparison for 24 Kodak images at an EC of 50,000 bits. The proposed method achieves higher PSNR values (except few cases having low inter-channel correlation) compared to the best-known SOTA methods for most images, demonstrating its generalizability. Security Analysis This sub-section evaluates the robustness

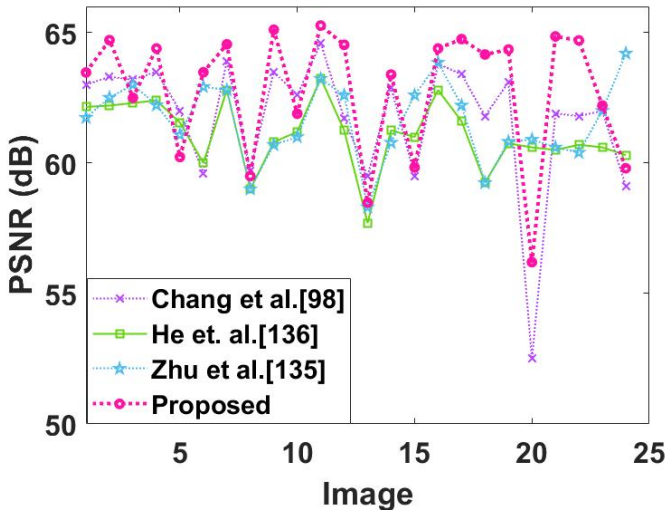


Fig. 4.8: PSNR comparison for Kodak color images at EC of 50,000 bits

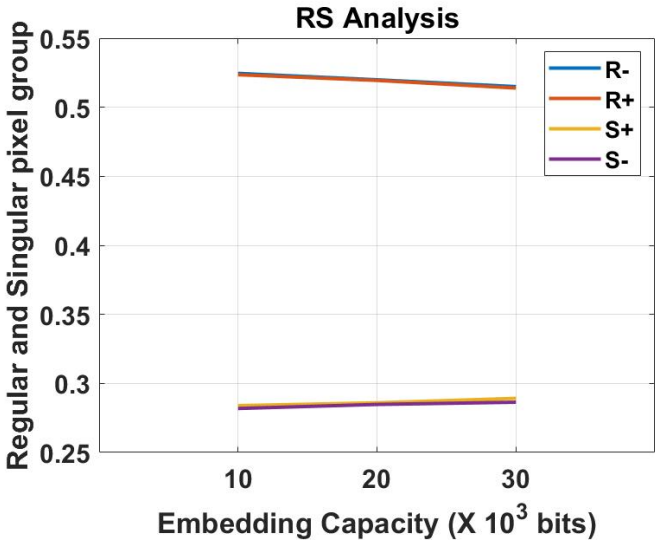


Fig. 4.9: RS steganalysis for kodim11

of the proposed method by employing RS steganalysis that classifies pixels of an image into three categories: Regular (R+ and R-), Singular (S- and S+), and unusable pixel groups. Fig. 4.9 illustrates the steganalysis for the embedded green channel that utilizes error adjustment while embedding. It is evident from the overlapping R+ and R- components, and S+ and S- components that the proposed method is highly robust against the RS steganalysis attacks.

4.4 Summary

In this chapter, new RDH method for color images has been proposed, emphasizing substantial innovation in the prediction component. The prediction employs a novel SA-CNNP, which effectively captures both local characteristics and global context for enhanced prediction accu-

racy. In addition, an error adjustment strategy is introduced to refine the PEH by leveraging inter-channel dependencies. Experimental results show that the proposed predictor outperforms existing SOTA predictors in terms of accuracy. When integrated with a classical RDH framework, the proposed SA-CNNP achieves an average PSNR improvement of approximately 1.2 dB over leading predictors. The overall RDH method for color images consistently surpasses SOTA techniques, highlighting its efficacy. While the proposed SA-CNNP achieves high prediction accuracy, the embedding performance is limited by the traditional data hiding approach utilising the traditional complexity calculation technique. The future research will focus on improving the embedding performance by utilising deep learning predictors for local complexity classification.

CHAPTER 5

UMANET: A TWO-STAGE INTERPOLATION-BASED REVERSIBLE DATA HIDING FRAMEWORK WITH ATTENTION-ENHANCED PREDICTION

Interpolation-based reversible data hiding (RDH) techniques have recently attracted significant attention due to their ability to enhance image resolution while ensuring secure data embedding. However, the effectiveness of these methods heavily depends on the quality of the interpolated cover images. Conventional interpolation techniques, typically based on linear models and limited local pixel contexts, often fail to generate high-quality cover images, thereby compromising the visual quality of the resulting stego images and limiting embedding capacity. To address these limitations, this chapter introduces a novel hybrid interpolation framework that combines bicubic interpolation with a deep learning-based predictor to construct a high-fidelity two-stage interpolation mechanism. Central to this framework is a newly proposed predictor, termed UMANeT, which leverages a broader contextual region for improved pixel prediction accuracy. By effectively capturing non-linear and long-range dependencies, UMANeT enhances the overall image quality used for data embedding. Experimental results demonstrate that the proposed method not only achieves superior embedding capacity but also generates cover and stego images of significantly higher visual quality compared to existing interpolation-based RDH techniques.

5.1 Introduction

Recently, interpolation-based RDH techniques [154–156] have gained significant attention as a promising approach for enhancing both embedding capacity and visual quality in data hiding applications. By leveraging interpolation, a low-resolution image is transformed into a high-resolution version by estimating intermediate pixel values based on surrounding data. This process not only improves image clarity but also provides additional embedding space without

distorting the original content. Unlike conventional RDH methods that directly modify original pixels, interpolation-based RDH selectively embeds SD in newly generated (non-reference) pixels while preserving the original (reference) pixels sampled from the input image. This strategy effectively mitigates visual distortions, significantly enhances embedding capacity, and ensures the full reversibility of the cover image upon data extraction. As a result, the field has witnessed significant advancements in recent years.

One of the earliest methods, Neighbor Mean Interpolation (NMI), introduced by Jung et al. [37], partitions the image into 2×2 blocks, designating the first pixel as a reference and predicting the remaining three. While simple and computationally efficient, NMI struggles to capture complex image structures due to its limited contextual awareness. To address this limitation, the Modified Neighbor Mean Interpolation (MNMI) method [156] extended the reference area to four surrounding pixels and incorporated a more refined weighting scheme. In a separate development, C.A. Hall [157] independently proposed Bicubic Interpolation [158], which focuses on expanding the contextual neighborhood rather than relying on predefined reference pixels. The Bicubic interpolation (BI) is an advanced version of bilinear interpolation [159] that expands the calculation to 16 neighboring pixels captured by a 4×4 grid surrounding the target pixel position. This modification improved prediction accuracy but remained constrained by its reliance on linear interpolation, which is inadequate for handling intricate image details. To overcome these shortcomings, Zhang et al. [160] introduced the Parabolic Interpolation (PI) method, employing a non-linear approach by fitting a parabolic curve through three reference points to predict two additional pixels. While PI enhanced adaptability to complex structures, it also introduced higher computational costs. More recently, Quadratic Bezier Interpolation (QBI), proposed by Hassan et al. [161], sought to balance accuracy and efficiency by transforming the three-parameter PI model into a single-parameter system using Bezier curves. This adaptation retained the benefits of non-linearity while reducing computational overhead.

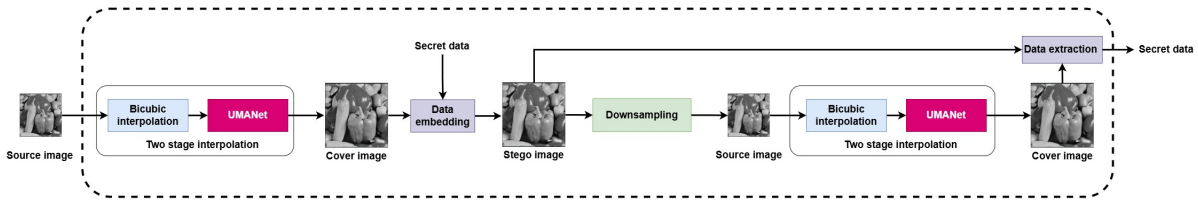


Fig. 5.1: Framework of the proposed two-stage interpolation-based RDH scheme.

However, researchers have not explored the bicubic interpolation technique for implementing RDH. We discovered that the bicubic Interpolation technique produces good-quality interpolated images. However, it alone struggles in producing high-quality stego images as the coverage of neighboring pixels captured by the method is still limited. On the other hand, with its capability of calculating non-linear relationships between neighboring pixels, deep learning techniques have proved themselves to be of wide use in various image processing tasks such as image classification. Building upon the foundation of traditional prediction methods, re-

cent research [80, 145, 149, 162, 163] has explored the integration of deep learning techniques within RDH to further enhance image prediction and reconstruction. Hu et al. [80] introduced a CNN-based predictor (CNNP) where the cover image is partitioned into two sub-images: cross sub-image and dot sub-image, wherein each sub-image comprises half of the pixels available at alternate locations with zero at the remaining positions. CNNP is then trained to predict one sub-image from another. With the capabilities of calculating non-linear relationships between neighboring pixels, CNNP achieves higher prediction accuracy than the traditional SOTA methods. Yang et al. [145] further improve the prediction accuracy of CNNP by partitioning the image into four sub-images. With this partitioning scheme, for predicting the pixels of one sub-image, pixels from the other three sub-images are available as neighboring pixels, thus providing more context for prediction and hence improving prediction accuracy.

To summarize, the existing interpolation-based data hiding methods typically employ traditional interpolation techniques, such as NMI and MNMI, that do not provide a promising resultant image as they lack adaptability. Similarly, the conventional deep learning based predictors being used in RDH cannot be directly applied to downsampled images, as the reduced number of pixels provides insufficient contextual information, leading to degraded reconstruction quality. To address these limitations, this chapter presents two main innovations. First, a two-stage framework is introduced in which bicubic interpolation is used to predict the center pixel in each block, providing an initial yet consistent estimation for reconstruction. Bicubic interpolation was particularly chosen over other techniques such as advanced edge-aware interpolation methods as it offers a controlled input that avoids introducing high-frequency artifacts. Second, building on this initial prediction, a novel deep learning model integrating channel attention (CA) and multi-head attention (MHA) in parallel [151] is proposed to simultaneously capture local and global dependencies. This enables the model to adaptively focus on the most relevant regions, refining the interpolation output and thereby enhancing reconstruction quality, robustness, and generalization. The key contributions of this work are as follows:

- **Two-Stage Interpolation-Based RDH Technique :** A sequential framework integrating bicubic interpolation with a novel deep learning–based predictor. The first stage applies bicubic interpolation to specific target pixels, forming an intermediate image that encodes crucial structural details. This intermediate output provides enhanced contextual cues for the second stage, where the predictor refines the interpolated result to produce a high-quality reconstruction.

- **UMANeT: A Multi-Head Attention-Enhanced Predictor:** The proposed UMANeT is a U-Net inspired architecture that integrates channel attention (CA) and multi-head attention (MHA) in parallel. Unlike conventional deep learning based predictors that either limit global spatial context or assign equal importance to all feature maps, CA prioritizes the most informative ones while suppressing less relevant features, and MHA captures long-range spatial relationships across the entire feature map. Combining these

complementary dependencies yields refined features that improve prediction accuracy for reconstructing the complete interpolated image.

- **Improved Image Quality and Embedding Capacity:** The proposed two-stage strategy with UMANeT significantly improves the quality of interpolated images, enabling high embedding capacity while maintaining superior visual quality. Extensive experiments on diverse datasets (USC-SIPI, Bossbase, and BOWS-2) validate the method's superiority and broad applicability across varied image characteristics. In particular, the proposed method achieves an average PSNR of 33.51 dB across BossBase and BOWS-2 datasets, representing a 4.29% gain over the best-performing state-of-the-art method.

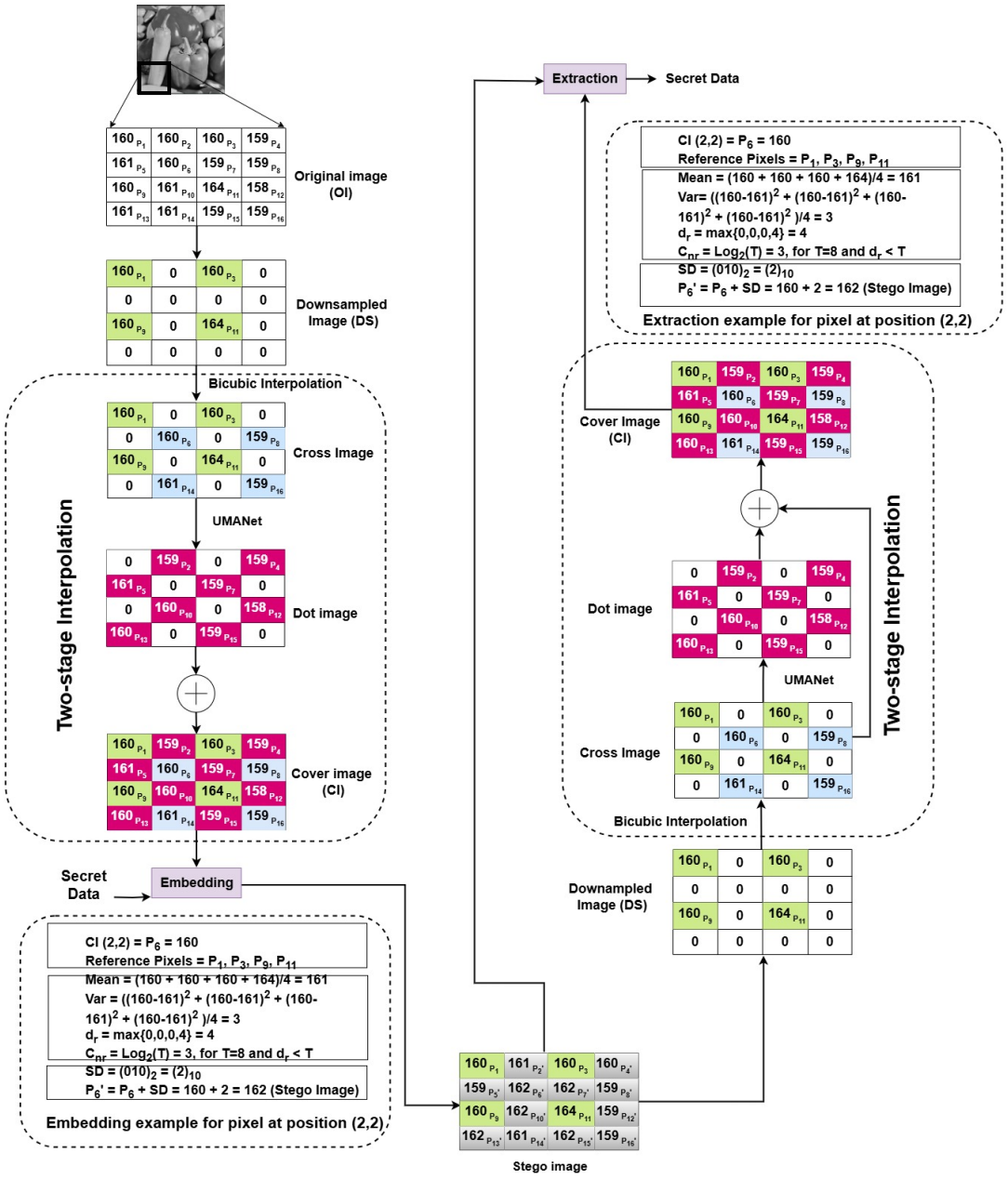


Fig. 5.2: Illustrative example to showcase the working of the proposed interpolation technique

5.2 Proposed Method

In this section, we present a novel interpolation-based RDH technique that introduces an innovative two-stage interpolation approach. This method effectively integrates bicubic interpolation with a newly developed deep learning-based predictor. The proposed RDH technique is designed to achieve a high embedding capacity while maintaining superior image quality, ensuring that the original image can be perfectly restored, thus preserving its reversibility. Fig. 5.1 outlines the basic structure of the proposed RDH technique, where the first stage utilizes bicubic interpolation to create baseline input for the second stage. The second stage utilizes a novel multi-head attention-based U-Net predictor that, when combined with bicubic interpolation, produces a high-quality interpolated image. Fig. 5.2 illustrates an example framework of the proposed RDH technique that comprises three primary components: 1. Interpolation, 2. Data embedding, and 3. Data extraction, that are discussed in detail in section 5.2.1, section 5.2.2, and section 5.2.3, respectively.

5.2.1 Two-Stage Interpolation

In interpolation-based RDH techniques, the high quality of the interpolated image is imperative to achieve high embedding performance. For this, a novel two-stage image interpolation method is introduced, which uses Bicubic interpolation [158] at the first stage and a deep learning predictor, namely UMANET, for the second stage. A schematic representation of the proposed interpolation-based RDH technique is illustrated in Fig. 5.2. The process begins with an original image (OI) of size $m \times n$, composed of pixels denoted as P_i. This image is downsampled to produce a reduced image (DS) of size $\frac{m}{2} \times \frac{n}{2}$. In Fig. 5.2, the pixels P_1, P_3, P_9, and P_{11} are shown in green. These pixels are retained from the original image and serve as reference pixels, while the remaining positions are designated as non-reference pixels. The data is embedded in the cover image (CI) at non-reference pixel positions only, leaving the reference pixels unaltered. This is important to losslessly recover the OI during the extraction procedure. The proposed two-stage approach of the interpolation and the training details of the proposed predictor are discussed in the following subsections.

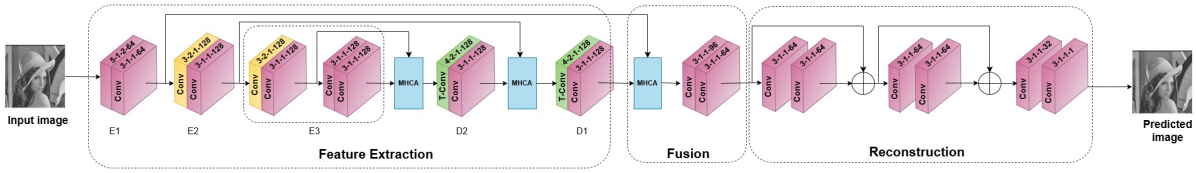


Fig. 5.3: Architectural design of UMANet: Conv block indicates a convolutional block with c output channels, a specified kernel size, stride s, and a ReLU activation layer.

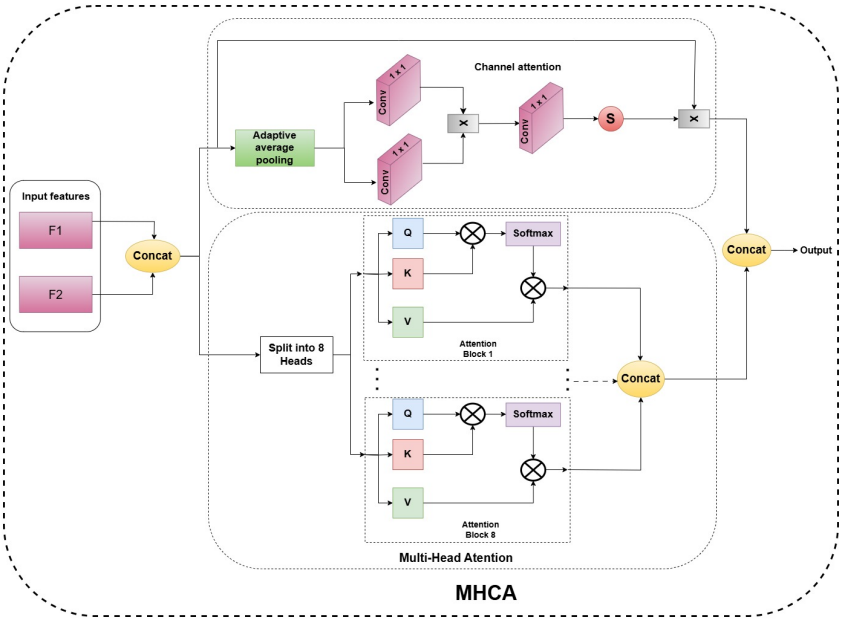


Fig. 5.4: Schematic representation of MHCA architecture featuring channel attention and multi-head attention mechanism.

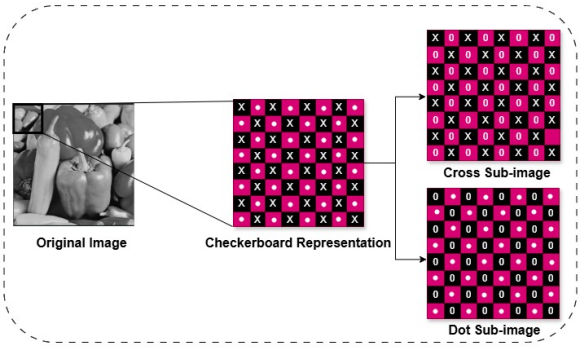


Fig. 5.5: Illustration of image division into cross and dot sub-images.

Bicubic Interpolation

In the first stage of the proposed framework, Bicubic Interpolation [158] is applied to the down-sampled image DS to estimate the pixels intensities at selective positions only, such as P_6, P_8, P_{14}, P_{16} etc. (highlighted in blue in Fig. 5.2). This strategic estimation ensures that the image transforms into a checkerboard pattern, where empty positions correspond to white pixels, while reference and interpolated pixels (via Bicubic Interpolation) are represented as black pixels. This structured arrangement serves as a crucial pre-processing step, enabling the subsequent UMANeT predictor to accurately infer the intensities of white pixels in the checkerboard layout.

The Bicubic interpolation operates by performing cubic interpolation along both horizontal

and vertical axes within a 4×4 neighborhood of reference pixels, as defined by Eq. 5.1

$$f(x, y) = \sum_{i=-1}^{2} \sum_{j=-1}^{2} p_{i,j} \cdot w(x - i) \cdot w(y - j) \tag{5.1}$$

where $f(x, y)$ represents the interpolated intensity at location (x, y). $p_{i,j}$ represents pixels intensity of the reference pixel at position (i, j) within the 4×4 grid. $w(x-i)$ and $w(y-j)$ are the weighting factors that are dependent on the distance of pixel $p_{i,j}$ from the location (x, y).

Following this interpolation step, the resultant black sub-image, containing all alternate pixels, is obtained. This sub-image subsequently serves as the input to the UMANet predictor, which further refines the pixel estimations and enhances reconstruction accuracy.

UMANet predictor

In this sub-section, the proposed U-Net-like network with multi-head attention, namely UMANeT, is discussed. The UMANet is a predictor that is employed by the proposed method at the second stage of interpolation to estimate the remaining (white location) values of the image. As shown in Fig. 5.3, the UMANet consists of a 3-layer architecture that comprises a feature extraction module, a feature fusion module, and an image reconstruction module. Convolution layers in each block are represented in conv $k - s - p - c$ format, where k represents Kernel size, s represents Stride, p represents the amount of padding, and c represents the number of output channels. The feature extraction module consists of a 3-stage encoder ($E_1 - E_3$) and a 2-stage decoder (D_1, D_2). Each encoder consists of standard convolution layers with the depth of feature maps controlled by the number of output channels as 64 in E_1, and 128 in E_2 as well as E_3. A convolution layer with stride 2 is used in E_2 and E_3 to reduce the size of feature maps without discarding any information. In the proposed architecture, instead of the commonly used channel adaptive attention module, a multi-head attention-based channel adaptive module (MHCA) is utilized to concatenate the output of encoders with the output of respective decoders. While the role of channel attention is to emphasize selective feature channels based on their relevance, integrating multi-head attention into channel attention improves the model's performance by introducing pixel-level spatial attention within each channel, enabling the model to capture finer details.

Fig. 5.4 presents the architecture of the proposed MHCA that integrates channel attention (CA) with multi-head attention (MH). At first, the input features are processed and concatenated along the channel dimensions for parallel processing in the MH and CA blocks. In the CA block, a pooling layer is applied to the features to obtain a feature vector that captures global context. The feature vector is passed through convolutional layers to introduce non-linearity and learn channel dependencies, and then element-wise multiplication is performed to compute refined weights. The refined weights are then passed through the Sigmoid function to normalize the weights in the range of 0 to 1, which represents the importance of each channel. The output

of the Sigmoid function is then combined with input features via skip connection to refine the features while balancing the impact of channel attention. The final feature map output of the CA is of the same shape as that of the input feature maps but represents enhanced channel-wise focus. The MHCA leverages the parallel operation of CA and MH operations, wherein CA focuses on channel dependencies and MH focuses on spatial dependencies. As shown in Fig. 5.4, the MH attention consists of 8 parallelly connected attention blocks, each focusing on the different regions of the input feature map. Using query (Q), key (K), and Value (V) vectors, attention weights of the feature map are calculated that signify which spatial regions are more crucial for the prediction. Diverse spatial patterns learned by each head are concatenated to create a refined, richer spatial pattern. Finally, concatenating the learned channel-level dependencies (from CA) and spatial-level dependencies (from MH) enables the model to calculate a more robust refined feature map.

The UMANet predictor is trained to predict the dot sub-image (the remaining half of the image) from the cross sub-image. The pixels corresponding to the dot sub-image are represented with pink colour in the figure. 5.2. The two sub-images, cross and dot, are combined to generate a complete interpolated image, also known as the cover image (CI).

UMANet training details

After constructing the network structure of the UMANet predictor, the network is trained as follows. The training methodology encompasses several key components:

- Runtime Environment: The UMANet predictor was implemented on an Intel Core i5 CPU (2.40GHz) alongside 16 GB of RAM with the help of the NVIDIA Tesla A100 GPU on Colab Pro.

- Dataset preparation and Pre-processing: A total of 3000 images are randomly selected from the BOWS2 dataset to construct the training set, complemented by a validation set of 600 images. To ensure consistency, all images underwent pre-processing to convert them to grayscale images and were standardized to dimensions of 512×512 pixels. Each image is then divided into cross and dot sub-images as presented in Fig. 5.5.

- Training Implementation: The model is trained to predict the dot sub-image from the cross sub-image. The training strategy utilizes the back-propagation algorithm to optimize the network parameters. The optimization process leverages the Adam optimizer [164] with a weight decay (λ) set to 10^{-3}. During the training process, the following loss function is used.

$$\text{loss} = \frac{1}{N}\sum_{i=1}^{N}(\hat{y}_i - y_i)^2 + \lambda\|\omega\|^2 \tag{5.2}$$

  where N is the batch size, set to 4, and y_i is the i^{th} original image, $\hat{y}_i$ is the corresponding predicted image, and ω represents the weights in the network.

5.2.2 Data Embedding

The cover image (CI) obtained after performing the proposed integrated interpolation steps comprises $(m \times n)/4$ reference pixels and $(3 \times m \times n)/4$ non-reference pixels. To ensure the reversibility of the image during data extraction, the reference pixels are kept unaltered while performing data embedding in the non-reference pixels. For data embedding, the embedding strategy proposed in [165] has been employed to ensure high-fidelity. As a first step, non-reference pixels are sorted based on increasing local complexity. This complexity is quantified by computing the variance of neighboring reference pixels located within a defined circular radius, as outlined in [165]. The variance is calculated using the mean square error function, as shown in Eq.5.3.

$$Var(i,j) = \frac{1}{N}\sum \left(DS(x,y) - \frac{1}{N}\sum DS(x,y)\right)^2 \tag{5.3}$$

where $DS(x,y)$ denotes the set of reference pixels derived from the downsampled image that fall within the radius r from the target pixel at location (i,j), and N is the number of such pixels. Next, the deviation d_r of each non-reference pixel is calculated from its neighbouring reference pixels as defined in Eq.5.4,

$$d_r(i,j) = \begin{cases} \max\left(|DS(i{+}1,j) - D|,\ |DS(i{-}1,j) - D|\right) \\ \qquad\qquad \text{; if i=2:2:m-3, j=3:2:n-3} \\ \max\left(|DS(i,j{-}1) - D|,\ |DS(i,j{+}1) - D|\right) \\ \qquad\qquad \text{; if i=3:2:m-3, j=2:2:n-2} \\ \max\left\{ \begin{array}{l} |DS(i{-}1,j{-}1) - D|,\ |DS(i{-}1,j{+}1) - D|, \\ |DS(i{+}1,j{-}1) - D|,\ |DS(i{+}1,j{+}1) - D| \end{array} \right\} \\ \qquad\qquad \text{; if i=2:2:m-2, j=2:2:n-2} \end{cases} \tag{5.4}$$

Here, $D = DS(i,j)$ denotes the non-reference pixel located at index (i,j). The format $a:s:b$ denotes a range where the index starts at a, increments by a step size s, and continues up to b. Based on the deviation value d_r, the number of bits (C_{nr}) that can be embedded in each non-reference pixel is calculated using Eq. 5.5.

$$C_{nr}(i,j) = \begin{cases} \lceil \log_2(T) \rceil & \text{if } d_r \le T, \\ \lceil \log_2(d_r) \rceil & \text{otherwise} \end{cases} \tag{5.5}$$

where, T is the minimum threshold that defines the minimum number of bits to be embedded in a pixel.

Post sorting the non-reference pixels in the order of their increasing complexity, ($C_{nr}(i,j)$) bits are embedded in each non-reference pixel at location (i,j) to obtain stego image (SI). To ensure the minimum distortion due to embedding, the direction of pixel modification is dynamically chosen depending on the direction difference between the original pixel in OI and

the interpolated pixel in CI. An exemplary illustration of the data embedding process for the pixel at position $(2, 2)$ is depicted in Fig. 5.2. For the target pixel P_6 in the cover image, and with a circular radius of $r = 2$, the reference pixels P_1, P_3, P_9, and P_{11} are located within the defined neighborhood. Based on a threshold value of $T = 8$, the number of bits to be embedded is determined using Eq. 5.5, resulting in $C_{nr} = 3$. Subsequently, three bits from the secret data sequence $(010)_2$ are embedded into P_6, producing the modified stego pixel value $P'_6 = 162$.

(a) Baboon **(b)** Couple **(c)** Lena **(d)** Boat **(e)** Elaine **(f)** Pepper

Fig. 5.6: Six standard images from the USC-SIPI dataset.

5.2.3 Data Extraction and Recovery

The Steps for data extraction as well as recovery of the cover image are similar to the steps performed during interpolation and data embedding. As shown in Fig.5.2, downsampling is performed on the SI received at the receiver end. The downsampling is similar to what was performed at the transmitter where the CI of size $m \times n$, which is again represented by pixels Pi, is downsampled into an image (DS) of size $\frac{m}{2} \times \frac{n}{2}$. Similar to the transmitter end, the reference pixels (represented in green) are retained in the DS image. Next, Bicubic interpolation is applied to the DS image to obtain the cross image. The cross image is provided as an input to the proposed UMANet predictor, which provides a predicted Dot image at the output. After combining the cross and dot images, the CI, which is identical to the CI at the transmitter end, is obtained at the receiver. Next, the complexity of the non-reference pixels is calculated using Eq.5.3 and the pixels are sorted in increasing order of their complexity. Subsequently, the number of bits (C_{nr}) embedded in each non-reference pixel is determined using Eq. 5.5. Based on the computed C_{nr} values, the difference between the corresponding non-reference pixels of the CI and the SI is calculated. This difference is then represented using 'C_{nr}' binary bits, which are used to extract the embedded secret data (SD) An illustrative example of the extraction procedure for a pixel located at position $(2, 2)$ is shown in Fig. 5.2. In this example, the received stego pixel is $P'_6 = 162$. Following the same interpolation strategy used at the embedding end, the corresponding interpolated pixel value is determined as $P_6 = 160$, and the number of embedded bits is calculated to be $C_{nr} = 3$. Consequently, the three-bit secret data can be extracted by computing the difference between the stego and interpolated pixel values, yielding $SD = (P'_6 - P_6) = (2)_{10} = (010)_2$.

5.3 Experimental Evaluation

This section empirically analyses the performance of the proposed UMANet-driven interpolation against distinct interpolation approaches [37, 155, 156, 165, 166] and Interpolation-based RDH methods [154, 165, 167–169]. The UMANet predictor was implemented on an Intel Core i5 CPU (2.40GHz) alongside 16 GB of RAM with the help of the NVIDIA Tesla A100 GPU on Colab Pro. To ensure the comprehensive performance evaluation of the proposed predictor, a diverse dataset comprising six images namely 'Lena', 'Baboon', 'Couple', 'Boat', 'Elaine', and 'Pepper', as illustrated in Fig.5.6, exhibiting distinct textural complexity and complex patterns, is selected from the SIPI database [69]. To validate the superiority of the UMANet predictor, the performance comparison is performed against five distinct SOTA methods: NMI [37], INP [155], ENMI [166], MNMI [156], and CPI [165]. The experimental analysis has been divided into two sections. Section 5.3.1 evaluates the performance of the proposed UMANet predictor by evaluating the quality of interpolated images (CI) against known SOTA methods. However, Section 5.3.2 evaluates the efficacy of the UMANet predictor by comparing its embedding performance against that of the SOTA methods. .

Table 5.1: PSNR comparison between original and cover images across various interpolation techniques on the USC-SIPI dataset, including % gain achieved by the proposed method over the best-performing SOTA method.

Image	NMI [37]	INP [155]	ENMI [166]	MNMI [156]	CPI [165]	Proposed	% Gain over the Best SOTA method
Lena	31.80	30.70	33.44	31.97	32.83	35.23	**5.37**
Peppers	29.88	28.81	31.47	30.48	31.31	32.19	**2.28**
Baboon	22.22	21.71	22.67	22.43	22.95	23.41	**3.15**
Elaine	30.99	30.33	31.80	31.61	31.75	31.89	**0.28**
Couple	25.67	25.02	26.55	26.05	26.93	27.96	**5.32**
Average	28.12	27.32	29.18	28.46	29.09	29.97	**2.71**

5.3.1 Effectiveness of UMANet Predictor based two-stage Interpolation Method

To achieve efficient embedding performance in an Interpolation-based RDH method, it is imperative to achieve a high-quality cover image during the interpolation stage. Peak Signal-to-Noise Ratio (PSNR), which serves as the primary metric for measuring the quality of an output image against a target image, is utilized to gauge the performance of the proposed UMANet predictor as an interpolation method. Table 5.1 represents the PSNR comparison of the obtained CI against the OI, achieved by the proposed method against the SOTA methods for the six standard images. It is evident from Table 5.1 that, among all the SOTA methods, the proposed UMANet predictor achieves the highest PSNR consistently for all six standard images [69].

Empirical results show that, for 'Lena' and 'Couple' images, the proposed method achieves approximately 2 dB improvement in PSNR compared to best-performing SOTA methods, which is equivalent to % gain of more than 5% for both the images against the best-performing SOTA

Table 5.2: PSNR comparison of various interpolation methods across BOSSBase and BOWS-2 image datasets, including % gain achieved by the proposed method over the best-performing SOTA method.

Image	Mandal et al. [170]	Jung et al. [37]	Lee et al. [155]	Chang et al. [166]	Malik et al. [171]	Zhong et al. [172]	Proposed –	% Gain over Best SOTA Method
BossBase [71]	26.95	30.58	29.73	31.65	30.65	32.72	33.13	**1.25**
BOWS-2 [173]	27.54	31.17	30.36	32.20	31.24	31.55	33.92	**6.90**
Average	27.25	30.88	30.04	31.92	30.95	32.13	33.51	**4.29**

Table 5.3: Performance comparison of maximum EC, bits-per-pixel (bpp), and PSNR for the proposed method and various SOTA methods across standard test images from the USC-SIPI dataset.

Images	Metrics	Jung et al. [37]	Lee et al. [155]	Xiong et al. [167]	Hassan et al. [154]	Chen et al. [168]	Bai et al. [169]	Zhang et al. [165]	Proposed
Lena	EC (bits)	1,99,091	3,45,446	3,19,712	3,85,680	3,05,644	5,85,225	6,22,812	**6,51,113**
	ER (bpp)	0.759	1.317	1.219	1.471	1.165	2.232	2.375	**2.483**
	PSNR (dB)	31.14	29.53	32.99	29.68	34.87	29.94	34.36	**35.88**
Baboon	EC(bits)	4,37,311	6,16,981	5,14,112	6,79,982	5,61,306	5,85,225	7,21,443	**8,38,394**
	ER (bpp)	1.668	2.353	1.961	2.593	2.141	2.232	2.752	**3.198**
	PSNR (dB)	21.91	21.18	22.53	21.27	24.08	22.02	24.84	**31.52**
Boat	EC(bits)	2,57,201	4,12,668	3,68,284	4,74,606	3,77,764	5,85,225	6,43,888	**7,00,655**
	ER (bpp)	0.981	1.574	1.404	1.810	1.441	2.232	2.456	**2.672**
	PSNR (dB)	27.71	26.49	28.86	26.39	32.24	26.88	30.92	**34.08**
Elaine	EC(bits)	2,34,020	3,83,043	3,38,430	4,43,813	3,55,971	5,85,225	6,12,120	**6,61,582**
	ER (bpp)	0.892	1.461	1.291	1.693	1.357	2.232	2.335	**2.523**
	PSNR (dB)	30.66	29.75	31.58	29.88	32.88	29.99	33.69	**34.02**
Peppers	EC(bits)	1,96,094	3,47,610	3,13,117	3,90,295	3,07,957	5,85,225	6,21,170	**6,52,981**
	ER (bpp)	0.748	1.326	1.194	1.488	1.174	2.232	2.369	**2.490**
	PSNR (dB)	29.76	28.63	31.14	28.78	33.35	28.55	33.40	**35.44**
Couple	EC(bits)	2,25,338	3,75,348	3,47,097	4,32,475	3,38,031	5,85,225	6,35,743	**6,84,869**
	ER (bpp)	0.859	1.431	1.324	1.649	1.289	2.232	2.425	**2.612**
	PSNR (dB)	25.87	25.46	26.42	25.04	28.04	25.34	28.57	**34.71**

method. Overall, when evaluating the performance over six images, the average gain achieved by the proposed method, in terms of PSNR, is 2% more than the average of the PSNR achieved by the best-performing SOTA method.

To assess the robustness of the proposed predictor and demonstrate the effectiveness of the interpolation scheme across diverse scenarios, performance has been evaluated on multiple benchmark datasets. Table 5.2 presents the PSNR comparison between the proposed interpolation approach and various SOTA methods on two widely used datasets: BossBase [71] and BOWS-2 [173], both of which consist of a large and diverse collection of images. The results clearly indicate that the proposed method consistently outperforms existing techniques, achieving an average PSNR improvement exceeding 4% compared to the best-performing SOTA alternatives.

5.3.2 Analyzing the Embedding Performance of two-stage Interpolation Method

To demonstrate the superiority of the proposed UMANet predictor-based two-stage Interpolation methodology, we performed a comprehensive comparative analysis against six SOTA methods. The performance metrics primarily focused on EC as a quantitative measure of secret data concealment efficiency, and PSNR as a measure of visual quality. This performance differential tested across six standard benchmark images [69] is visualized through the EC trajectory illustrated in Fig. 5.7. As evident from Fig. 5.7, the proposed method demonstrates remarkable

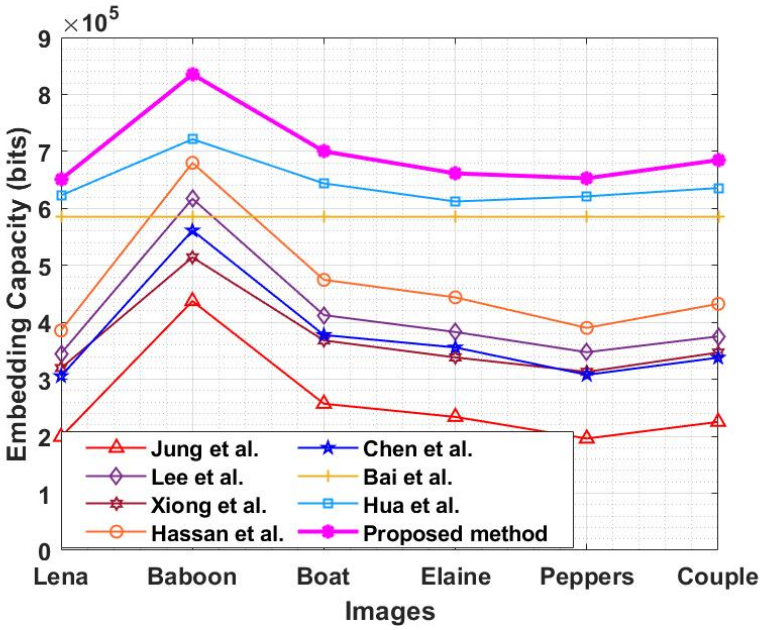
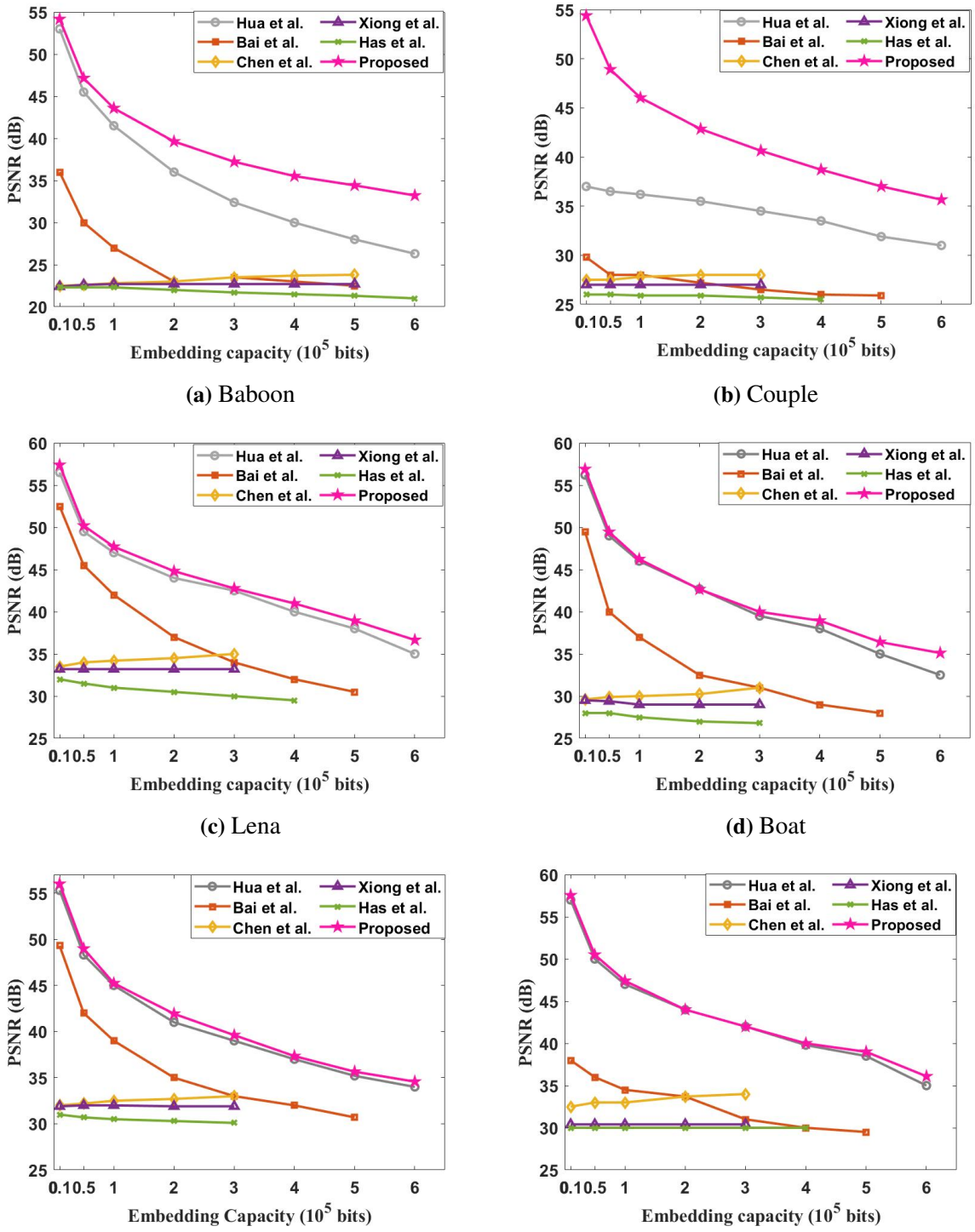


Fig. 5.7: Comparative analysis of maximum EC on the USC-SIPI image dataset, including the proposed method and SOTA methods.

embedding performance across all test images. Further, Table 5.3 represents the maximum EC and corresponding PSNR achieved by the proposed as well as the SOTA Interpolation-based RDH methods. For the Lena image, the method achieves an embedding capacity of 651,113 bits, representing a significant improvement of 4.54% compared to [165]. Similarly, for the Baboon image, our method reaches an exceptional capacity of $838,394$ bits, marking a 16.21% increase over [165]. The visual quality metrics, measured through PSNR, show equally impressive results. The proposed method consistently maintains superior image fidelity across all test images. For the Lena image, achieving a PSNR of 35.88 dB, representing a 4.4% improvement over [165]. The Couple image achieves a PSNR of 34.71 dB, marking a substantial 21.48% improvement compared to [165]. As evident from the Table 5.3, the empirical results demonstrate that the proposed methodology consistently outperforms SOTA methods [154, 165, 167–169] across all test images. On average, our method achieves an 8.42% improvement in EC and an 11.67% enhancement in PSNR values. The most notable improvements are observed in the Baboon image, with a 16.21% increase in capacity and a 26.87% enhancement in visual quality, highlighting the robust performance of our approach across varying image characteristics. These quantitative metrics definitively establish that our method achieves superior performance in both aspects of RDH schemes - maintaining exceptional visual quality while achieving significantly higher EC compared to contemporary approaches.

Fig. 5.8 represents the PSNR comparison of stego images obtained by the SOTA methods and the proposed method for the six SIPI images, at the distinct level of EC ranging from 0.1×10^5 bits to 6×10^5 bits. As evident from fig. 5.8, in comparison to the well-known SOTA methods, the proposed method showcases superior performance at all embedding capac-



(a) Baboon **(b)** Couple

(c) Lena **(d)** Boat

(e) Elaine **(f)** Pepper

Fig. 5.8: Evaluation of PSNR performance across varying EC for the proposed method and SOTA methods on USC-SIPI standard images.

ities across all six images. For the Baboon image, the proposed method achieves a PSNR of 54.17 dB at the EC of 10^4 bits, which is a 2.2% improvement over the best-performing SOTA method [165]. The widening of the gap with the increase of EC represents that the embedding performance is sustainable at all EC and even improves at higher EC in comparison to

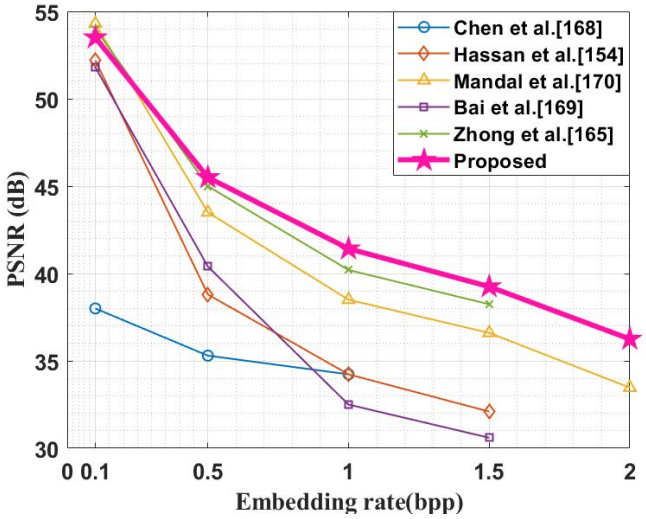


Fig. 5.9: Comparative analysis of PSNR versus embedding rate (bpp) on BOWS-2, contrasting the proposed method with SOTA methods.

the SOTA methods. For the 'Couple' image, the proposed method witnesses drastic improvement over SOTA methods, especially at the lower ECs. The PSNR (54.4 dB) achieved by the proposed method at EC of 10^4 bits is 47% higher than the best-performing SOTA method and 67.4% higher than the average PSNR obtained by the 5 SOTA methods. For most of the images, the embedding performance achieved by the proposed method at lower EC is superior to the SOTA methods, with performance enhancing progressively with increasing EC. These results indicate that the proposed approach not only achieves high EC but also effectively optimizes the trade-off between embedding capacity and visual quality.

To evaluate the embedding performance of the UMANet predictor over diverse characteristic images, the chapter studies its performance over another popular database, Bows-2 [173], with 50 images randomly selected for the analysis. Fig. 5.9 represents the comparative analysis of average PSNR achieved by various interpolation-based techniques at different embedding rates. It is evident from the Fig. 5.9 that the proposed method achieves superior PSNR in comparison to all the SOTA methods, which further improves at higher embedding rates (ER). The maximum average EC achieved by the proposed method over the Bows-2 database is 2.47 bpp, which is the highest among all the well-known interpolation-based SOTA methods. These results highlight the superiority of the proposed UMANet predictor-based two-stage interpolation methodology.

5.4 Summary

In this chapter, we proposed a novel predictor, UMANeT, which integrates the global context modeling capabilities of multi-head attention into a U-Net-like deep learning architecture. Building on this, we introduced a two-stage interpolation-based reversible data hiding

framework that combines traditional bicubic interpolation with the UMANeT predictor. This integrated approach enhances the quality of the interpolated cover images, which directly contributes to improved data embedding performance. Experimental evaluations demonstrate that the proposed method outperforms existing state-of-the-art techniques, achieving both higher embedding capacity and superior visual quality of the stego images.

CHAPTER 6

SCAM-NET: SPATIAL-CHANNEL ATTENTION MULTI-SCALE NETWORK FOR REVERSIBLE DATA HIDING IN ENCRYPTED IMAGES

The rapid pace of digitization, coupled with the widespread use of cloud storage, has significantly increased data transmission over networks, raising serious concerns about security and confidentiality. This is mainly because digitization converts records into formats that can be easily accessed or stolen if unprotected, while cloud storage keeps data online, exposing it to potential cyber-attacks from anywhere. These combined factors amplify the risk of unauthorized access, underscoring the need for stronger protective measures. To address this, this chapter introduces SCAM-Net, a Spatial-Channel Attention Multi-Scale Network, within a RDHEI method. SCAM-Net strengthens security by using spatial attention to highlight key image regions and channel attention to refine features across multiple scales, boosting the prediction precision. This heightened accuracy, in turn, improves data embedding effectiveness, ensuring robust concealment within encrypted images. Experimental results show that the proposed method achieves the highest embedding capacity for most of the USC-SIPI images, surpassing the best-performing SOTA method by 7.4%. Our proposed method not only exhibits superior performance across diverse datasets (USC-SIPI and BOWS-2) but also achieves a notable enhancement in embedding performance (approximately 66%) over medical images randomly chosen from The Cancer Imaging Archive TCIA dataset.

6.1 Introduction

The traditional RDH techniques struggle with a critical challenge in preserving the confidentiality of the cover image, which often remains exposed in its plain form. This becomes crucial in applications, such as military and healthcare, where the images themselves are of primary concern and need to be protected from unauthorized access. To overcome this vulnerability,RDHEI

was introduced to enhance confidentiality and privacy by first encrypting the cover images and then allowing embedding in these encrypted images. This not only protects the secret data but also keeps the cover media confidential. In recent years, RDHEI has gained significant traction due to its widespread applicability, particularly in cloud security [174, 175].

Primarily, RDHEI methods are distinguished into two key divisions based on their embedding approach: VRAE [176] and RRBE [177]. VRAE methods entail initially encrypting the original image and subsequently assigning aside space for embedding. In contrast, RRBE techniques employ an alternative strategy by initially exploiting the natural image correlations through pre-processing, thereby reserving space for embedding before the image undergoes encryption. Consequently, spatial dependencies within natural images are efficiently exploited, allowing these methods to achieve high-capacity data embedding.

Ma et al. [178] introduced one of the earliest RRBE-based RDHEI methods, pre-processing the image usingHS to create room for embedding, where the first three Least Significant Bits (LSBs) are shifted between regions, achieving an Embedding Rate (ER) of up to 0.5 bpp. However, this method does not ensure perfect reconstruction of the original image. Puteaux et al. [87] pioneered the first PE driven RDHEI method, utilizing the MED to calculate PEs and embed data in the MSBs of image pixels. While this approach achieves an average ER of approximately 1 bpp, reconstructing the original image necessitates identifying non-embeddable Most Significant Bits (MSBs) through an 8-bit flag, potentially compromising data restoration. Building upon the work of [87],

Yin et al. [34] introduced an innovative methodology that first utilizes the MED to approximate pixel values and subsequently performs bit-level evaluations to derive PEs, which are then compressed via Huffman encoding to allocate additional embedding space. Gao et al. [179] employed multilinear regression for pixel prediction and managed the absolute PEs via a dual-level label-map approach. Although accurate PEs can enhance EC, the regression technique substantially extends the computation time.

Most of these method target to improve embedding performance by improving the prediction performance, however, the prediction accuracy remains limited due to limitation of linear calculation. To improve upon this, leveraging the non-linear capability of CNN, Ping et al. [94] presented an asymmetric CNN-based predictor with a two-stage embedding model that leverages neighboring pixel correlations, enabling flexible and independent data embedding and image recovery through an innovative prediction technique. However, the overall prediction performance remains limited as the number of neighbouring pixels considered for prediction remains limited. In this chapter, we propose a novel predictor that leverages multi-scale feature extraction, focusing on larger feature set and further utilizing attention mechanism to refine the feature maps.

The key contributions of the chapter include the followings:

- A novel SCAM-Net predictor has been designed, incorporating channel attention and spatial attention to initially emphasize important regions for feature selection out of the

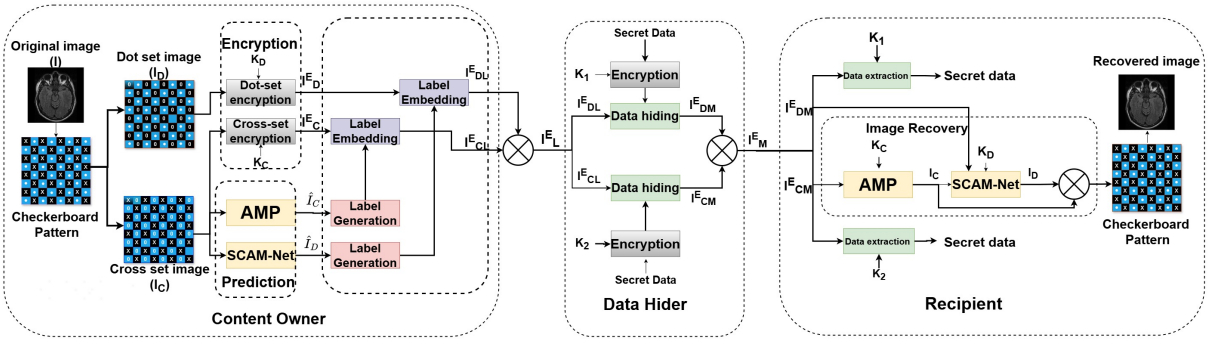


Fig. 6.1: Overall framework of the proposed method.

multi-scale feature extraction, and subsequently refine feature maps adaptively, enhancing prediction performance.

- Experimental results demonstrate that the proposed predictor attains high prediction accuracy, resulting in increased embedding capacity.

The subsequent sections are structured as follows. Section 6.2 delves into the details of our proposed methodology. In Section 6.3, we present a comprehensive analysis of our experimental results. Lastly, Section 6.4 concludes the chapter.

6.2 Proposed Work

This section discusses the proposed RDHEI method in detail. Fig. 6.1 represents the overall framework of the proposed method that involves three key entities: content owner, data hider, and recipient. The content owner is the one who possesses the original image and performs encryption to ensure that the content remains private. Data hider hides the data in the encrypted image while ensuring reversibility. The recipient receives the marked image and performs data extraction and content recovery depending upon the access to the data and content keys. The proposed method allows the separation of two roles at the recipient so that two different users, data recipient and content recipient, can retrieve the data and content, respectively. The workflow includes three principal components: Prediction, Data embedding, and recovery. These components are discussed in detail in subsequent sub-sections.

6.2.1 Prediction

For achieving high-embedding capacity in RDHEI methods, pixel prediction is one of the most crucial steps, as the prediction accuracy directly correlates to the embedding efficiency. For this, the original image I is divided into two sub-images, cross I_C and dot I_D. As illustrated in Fig. 6.1, the cross-set image contains half of the pixels located at alternate positions, and the dot-set

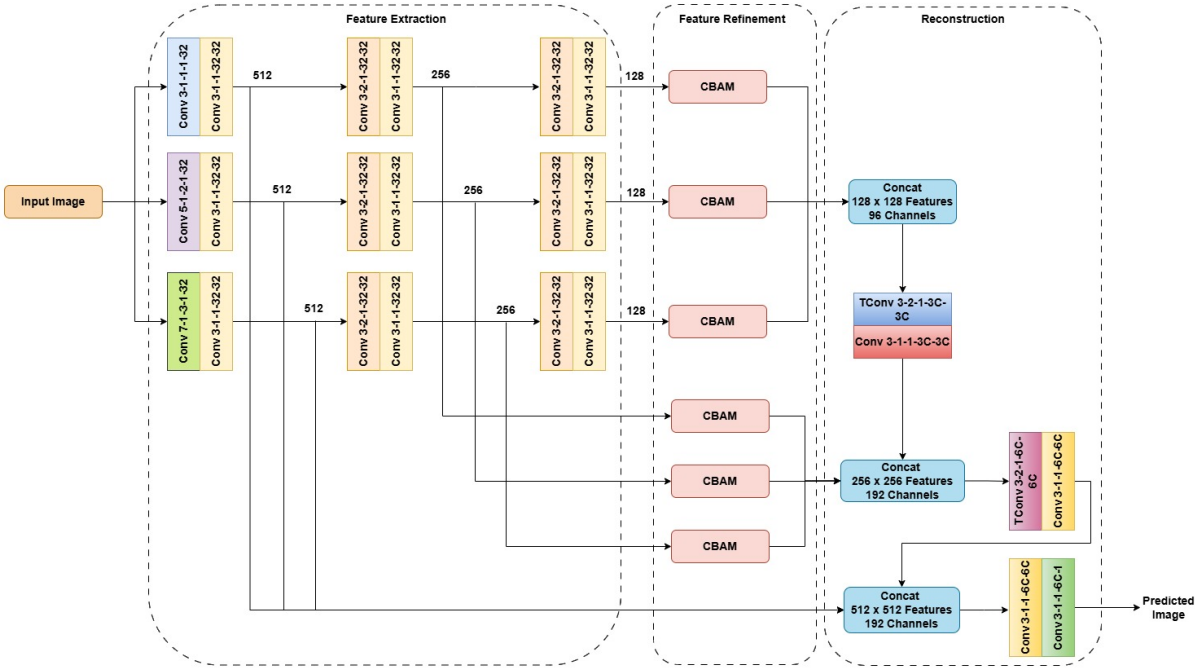


Fig. 6.2: Schematic representation of SCAM-Net predictor

image contains the other half of the pixels with zeros at the position of cross pixels. The two sub-images are independent of each other but are highly correlated and can be used to predict each other. Next, the prediction is done in two stages as shown in the proposed framework. At first, we utilize the novel SCAM-Net predictor to predict the dot-set image $\hat{I}_D$ from the cross-set image. Subsequently, the adaptive mean predictor [94] is employed to predict the cross-set image $\hat{I}_C$ from itself. In the following sub-section, the proposed SCAM-Net predictor is discussed in detail.

SCAM-Net Predictor

As shown in Fig. 6.2, the SCAM-Net (Spatial-Channel Attention Multi-Scale Network) predictor consists of a feature extraction module, a feature refinement module, and a reconstruction module. Each convolution block within the model is represented in a $k-s-p-cin-cout$ format that represents kernel size, stride, padding, input channels, and output channels, respectively. The feature extraction module comprises multi-scale feature extraction with three parallel convolutional blocks, each with a different kernel size (3×3, 5×5, and 7×7), capturing granular details as well as broader details. The obtained feature-set is then downsampled to create a hierarchical structure capturing multi-scale features. The proposed model utilizes Convolutional Block Attention Module (CBAM) illustrated in Fig. 6.3 for feature refinement. The CBAM module leverages a combination of channel attention and spatial attention, leading to improved feature refinement. The reconstruction module uses concatenation to fuse the refined features and perform upsampling to recover the original dimensions. The module further utilizes skip connections to ensure the high-resolution details are utilized for the final prediction output.

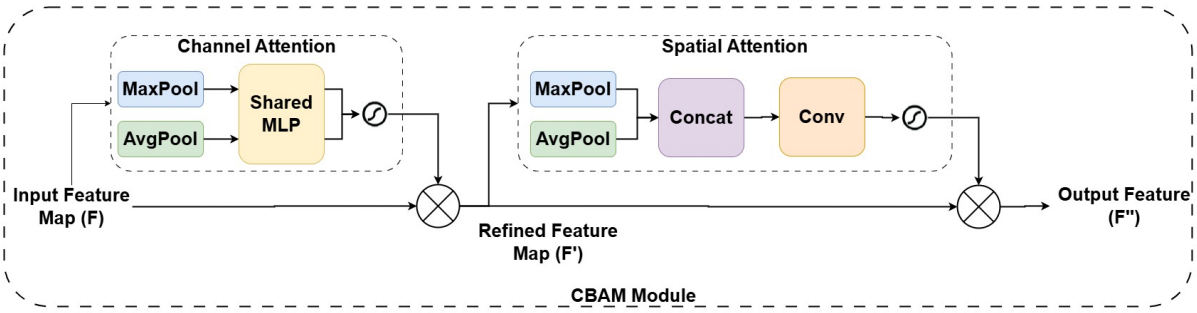


Fig. 6.3: Structural illustration of CBAM

To train the proposed SCAM-Net predictor, the training and validation datasets are constructed by randomly extracting $3,000$ images and 750 images from the BOWS2 [173] dataset, respectively. Prior to training, the color images in the train and validation sets underwent preprocessing to convert them to grayscale images and were standardized to dimensions of 512×512 pixels. The loss function L for the proposed predictor is depicted as follows:

$$\text{loss} = \frac{1}{N}\sum_{i=1}^{N}(\hat{o}_i - o_i)^2 + \lambda\|\omega\|^2 \tag{6.1}$$

where N is the batch size, set to 4, and o_i is the i^{th} original image, $\hat{o}_i$ is the corresponding predicted image, and ω encapsulates all the weights in the network. The proposed predictor is trained on an Intel Core $i5$ CPU (2.40GHz) alongside 16 GB of RAM with the help of the NVIDIA Tesla A100 GPU on Colab Pro, optimized via backpropagation and the Adam optimization algorithm.

Adaptive mean predictor (AMP)

The proposed method employs the AMP for predicting cross-set image from itself as done in [94]. For this, the first row and first column are reserved as the reference pixels. The predictor calculates the mean of the two neighboring pixels in four directions (top, left, right, and bottom) of each target pixel location. Subsequently, the predicted image with minimum mean square error (MSE) out of the four directions is selected as the predicted cross-set image.

6.2.2 Encryption and Data Hiding

In the proposed method, content encryption and data hiding are performed independently in the two sub-images. The entire process is discussed in the following subsections. More comprehensive analysis of encryption and data hiding procedure can be found in [34].

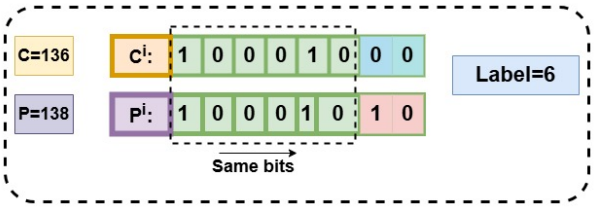


Fig. 6.4: An instance of pixel labeling

Image Encryption

The proposed method ensures privacy by encrypting the image before hiding data in it. The content owner uses two keys, K_C and K_D, to encrypt the cross-set and dot-set images and obtain ${I^E}_C$ and ${I^E}_D$, respectively. Each key (K_s) is used to generate a matrix of random numbers of the same size as that of the sub-images, and then the sub-images are encrypted as per Eq. 6.2.

$$E(i,j) = I_s(i,j) \oplus K_s(i,j) \tag{6.2}$$

where $s = D$ for the dot-set image, and $s = C$ for the cross-set image.

Label map generation

To calculate the number of bits that can be losslessly embedded into the cover image, labels are generated for each pixel in the two sub-images. For this, a bit-by-bit comparison of the original and predicted pixels of the sub-images is performed, as shown in Fig. 6.4, where the count of bits that are the same from the Most Significant Bit (MSB) to the Least Significant Bit (LSB) is noted as the label for that pixel. A label with value n, where $0 \leq n \leq 8$, represents that $n+1$ bits, where $n+1 \leq 8$, can be embedded in the pixel.

Label map embedding

To recover the secret data at the recipient, the label map needs to be embedded into the sub-images. The proposed method utilizes Huffman codes to compress the labels. Once the Huffman coding rule is generated for the two sub-images, the sub-images are raster scanned, and a binary sequence is generated, replacing the labels with their respective codes. The binary-coded sequence represents the compressed labels and is embedded into the sub-images along with the secret data. ${I^E}_{DL}$ and ${I^E}_{CL}$ represent the sub-images obtained after embedding the labels into the respective encrypted sub-images.

Data hiding

Once the two sub-images are encrypted and label maps are embedded into the reference pixels by the content owner, the data hider hides the secret data in the encrypted images. First, the

secret data is split into two sequences based on the available embedding capacity of each sub-image, calculated from the labels. Then, the data hider encrypts the secret data using two encryption keys ($K1$ and $K2$), known to the data hider only. The data hider then reads the label (n) for each pixel and hides $n + 1$ bits in the MSB of each pixel. Combining the two marked sub-images provides a final marked encrypted image ${I^E}_M$. It is important to note that in the dot-set image, the pixels designated for storing the Huffman coding rule remain unmodified, and, in the cross-set image, the reference pixels (depending on the direction of prediction) are preserved without alteration.

(a) Baboon **(b)** Jetplane **(c)** Lena **(d)** Pepper

Fig. 6.5: Benchmark images from the USC-SIPI dataset

6.2.3 Data Extraction and Image Recovery

The proposed RDHEI method enables the recipient to perform data extraction and image recovery separately. To extract the data, the recipient splits the received marked image ${I^E}_M$ into cross and dot sub-images. The recipient reads the reference pixels to extract the length of the secret data, the coding rules, and the label maps for the two sub-images. Next, the pixels in the two sub-images are raster-scanned to read the $n+1$ bits of each pixel according to their respective label value n recovered from the label map. The recovered binary sequences, one each for cross-set and dot-set images, represent the encrypted secret data, which is then decrypted with the corresponding data keys ($K1$ and $K2$) to recover the secret data.

For image recovery, the recipient follows the same steps as the transmitter but in reverse order. First, the cross-set image is predicted by applying the adaptive mean predictor on the reference pixels. For each pixel, replace the n MSB bits with the corresponding bits of the predicted pixels and $(n + 1)th$ bit by the inverse of the $(n + 1)_{th}$ bit of the predicted value. Decrypt the remaining bits using the key K_C. Once the cross-set image is fully recovered, predict the dot-set image from it using SCAM-Net predictor. Replace the $n + 1$ MSB bits of the pixels according to the respective labels. Decrypt the remaining bits using the key K_D. Combining the recovered cross-set image and the dot-set image gives the fully recovered cover image.

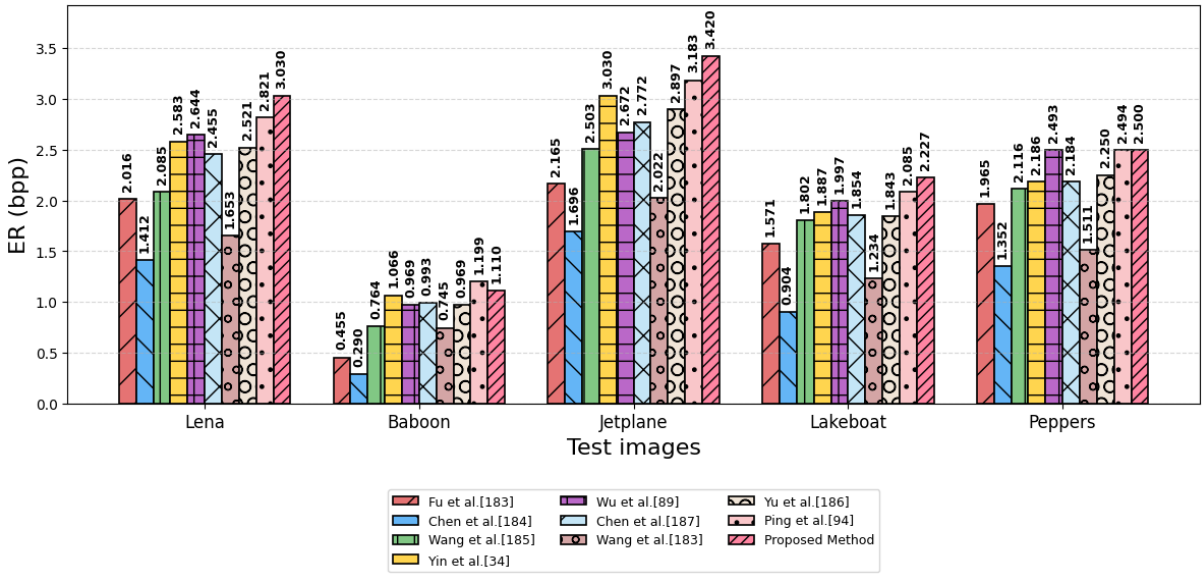


Fig. 6.6: Quantitative assessment of various SOTA methods on test images

6.3 Experimental Evaluation

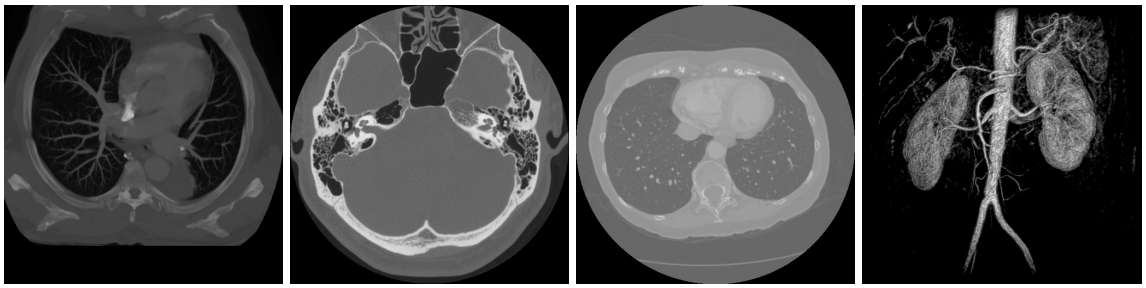
This section entails the experimental findings of the proposed RDHEI scheme and compares its performance with the well-known SOTA methods. We first, assess the performance of the proposed SCAM-Net predictor. Subsequently, we conduct an embedding capacity analysis of our proposed RDHEI approach. To ensure the analysis of images having diverse characteristics, we have utilized the widely used three datasets, BOWS2 [173], USC- SIPI [69], and TCIA dataset [4] for the experimental analysis.

Table 6.1: Statistical measures (Mean, VAR, MSE) across different predictors using the test set

Method	**Mean**	**VAR**	**MSE**
GAP [32]	7.71	174.23	247.73
MEDP [180]	6.12	84.29	129.66
RP [26]	5.14	52.59	84.59
CNNP [80]	4.16	53.52	81.05
CNNPEO [81]	3.31	20.29	34.20
Proposed	**2.97**	**19.40**	**30.54**

Table 6.2: Statistical assessment of embedding capacity and payloadbpp for USC-SIPI test images.

Test images	**Total EC (bits)**	**Label maps (bits)**	**Auxillary info (bits)**	**Pure EC (bits)**	**Payload (bpp)**
Lena	15,64,683	7,70,151	256	7,94,276	3.029
Baboon	10,85,614	7,93,112	256	2,92,246	1.114
Jetplane	16,47,532	7,48,332	256	8,98,944	3.429
Lakeboat	13,81,254	7,97,050	256	5,83,948	2.227
Peppers	14,38,241	7,82,175	256	6,55,810	2.500



(a) Med1 **(b)** Med2 **(c)** Med3 **(d)** Med4

Fig. 6.7: Medical images from the TCIA dataset [4]

Table 6.3: Performance evaluation on medical images [4]

Method	Medical image				Average
	med1	med2	med3	med4	
Bhardwaj et al. [181]	2.24	2.24	2.24	2.24	2.24
Xiuli et al. [182]	2.49	2.24	2.63	2.24	2.39
Proposed	5.10	4.19	2.66	3.99	3.98

6.3.1 Prediction Performance

Table 6.1 presents a performance comparison of the proposed SCAM-Net predictor against the various state-of-the-art (SOTA) predictors [26, 32, 80, 81, 180], evaluated on a test dataset comprising 8 standard images from the USC-SIPI dataset and 24 images from the Kodak dataset, using metrics such as Mean, Variance (VAR), and MSE. Among all the predictors, the SCAM-Net predictor achieves the most notable results, recording the lowest values for Mean (2.97), Variance (19.40), and MSE (30.54). It is evident from the superior results that the multi-scale feature extraction helps the proposed predictor to capture a larger feature set comprising both finer details and global details. Additionally, the CBAM module first helps in identifying the relevant feature maps using channel attention and then adaptively refines the filtered feature maps using spatial attention. Thus, capturing the rich hierarchical features and then adaptively refining the most critical features helps the proposed SCAM-Net predictor achieve remarkable prediction accuracy.

6.3.2 Embedding Performance Evaluation

To illustrate the effectiveness of our method, we assess its embedding capacity, measured in bpp, against nine SOTA methods [34, 89, 94, 103, 183–187]. Table 6.2 summarizes the total EC, the length of label maps, and the pure EC for the six test images. It is evident from the results that the higher prediction accuracy leads to a smaller length of label map, resulting in improved payload embedding. Fig. 6.6 demonstrates a comparative analysis for different methods on five standard test images from the USC-SIPI dataset, illustrated in Fig. 6.5. The results demonstrate

the superior performance of our approach. Notably, our method achieved the highest ERs on four out of five test images, with particularly impressive results on the 'Lena' image (3.0299 bpp) and 'Jetplane' image (3.42 bpp), surpassing the best performing SOTA method by 7.4%.

6.3.3 Robust Embedding Performance in Medical Images

Table 6.3 presents a comparison of the ER for various methods applied to medical images from the TCIA dataset, as shown in Fig. 6.7. The proposed approach surpasses all SOTA methods, delivering notably higher ER values, especially for med1 (5.10) and med2 (4.19), and achieving the highest overall average of 3.985.

Therefore, the experimental findings show that the proposed method exceeds SOTA methods by delivering the highest ER across all test images. This steady enhancement highlights its promise for real-world uses, such as secure image transmission, medical imaging, and digital watermarking, where maintaining high EC with minimal distortion is crucial.

6.4 Summary

This chapter introduces a novel RDHEI scheme that utilizes SCAM-Net predictor for generating highly accurate predicted images. The SCAM-Net predictor features a multi-scale extraction module that captures both fine and broad details effectively. Additionally, an integrated CBAM module enhances feature refinement by leveraging spatial and channel attention mechanisms. Experimental results reveal that this predictor outperforms existing state-of-the-art alternatives with approximately 11% improvement in MSE over best known SOTA method. The improved prediction performance led to improved embedding performance which is evident from the enhanced embedding capacity achieved by the proposed method (approximately 7.4%) for the majority of USC-SIPI test images. Furthermore, it shows substantial gains in embedding capacity across a diverse collection of medical images, underscoring its broad practical value.

CHAPTER 7

CONCLUSION, FUTURE SCOPE AND SOCIETAL APPLICATIONS

With the increase of digitization in the Healthcare industry, the importance of content protection and information security has increased. RDH entails an important role in the realm of information security, especially in content authentication and copyright protection. It involves concealing confidential/secret information within a cover medium in such a manner that it remains imperceptible to the human eye. Its dual capability, enhanced security and lossless recovery, has made RDH methods highly popular among security researchers, particularly for applications like medical and military fields, where preserving content accuracy and integrity is paramount.

Recent breakthroughs in RDH have led to promising techniques such as DE, HS, and prediction-error expansion (PEE) for minimizing distortion. Some of techniques rely on designing accurate predictors to predict pixel values, such as the DP, MED, GAP, and Bilinear interpolation (BIP). Other popular areas within RDH involve RDHCE and RDHEI that focus on improving contrast of the stego images or data hiding in encrypted images respectively.

It was observed that traditional RDH methods achieve low EC or result in poor quality stego images, in other words, exhibiting a trade-off between EC and image quality. To mitigate this trade-off, this thesis has presented several novel RDH methods that significantly advance the SOTA by achieving higher EC, improved visual quality, optimum contrast enhancement, brightness preservation and enhanced security. The contributions of the thesis are outlined in the subsequent subsections.

7.1 Research Contributions

- **Survey of reversible data hiding: Statistics, current trends, and future outlook:** The evolving area of RDH warrants a comprehensive review and analysis to map out the in-

tellectual development of the field, identify emerging trends, and highlight promising research directions. As a result, this study offers an extensive review of state-of-the-art RDH techniques along with their trend analysis, key contributions, and evolution. The study presented a thorough quantitative analysis of research patterns through visualizations of publication timelines, citation trends, and statistical modeling across years. The study incorporated modern tools such as co-occurrence analysis and keyword clustering to identify hot topics and recent trends, where researchers have contributed more frequently.

Along with the quantitative analysis, an in-depth qualitative analysis of various techniques has been conducted to present a systematic, highlighting major research contributions for each category. The study presented a taxonomy tree of different categories of RDH showing the domain evolution till date. Additionally, a comparative overview of different techniques within different domains has been presented while highlighting their originators, advantages, and limitations. Overall, the study presents a comprehensive guide for the researchers embarking on the evolution of RDH since its very inception. It presented an in-depth analysis of the existing literature's strengths and weaknesses, offering insights into the various approaches categorized by their operational strategies, such as RDHCE, RDHEI, prediction methods, interpolation, and robustness, their application areas, and future directions.

- **A high-capacity reversible data hiding with contrast enhancement and brightness preservation for medical images:** In the era of RDHCE, which is primarily popular in healthcare industry, the SOTA methods lag in managing the trade-off between image quality and EC. This work proposed a novel pre-processing technique for the ROI of a medical image that focuses on intelligently creating and managing the empty bins, resulting in low distortion. The study focused on not only enhancing the contrast of the image but also preserving the image brightness. The method followed an adaptive approach to allow different ER to be chosen for ROI and NROI regions, further minimizing the tradeoff between image quality and EC.

  Experimental results show that the proposed method effectively preserves brightness with 0.99 RMBE while providing higher EC, exceeding the best known SOTA methods by 20%. Furthermore, the method enhances contrast with minimal distortion in the output stego images. The experimental analysis further underscored the robustness and effectiveness of the proposed RDH approach against the prominent RS steganalysis.

- **Reversible Data Hiding for Color Images using a Novel Self-attention-based CNN predictor and Error Adjustment:** This study introduced a new RDH method for color images by efficiently utilizing inter-channel correlation of the three channels for improved embedding performance. Self attention based CNN (SA-CNN) predictor is introduced, which integrates self-attention mechanisms into the CNN framework. This integration

allowed the predictor to capture both local and global contextual information by attending to long-range pixel dependencies across the image, thereby significantly improving prediction accuracy and thus achieving improved Prediction error histogram (PEH). In addition to the novel predictor, an error adjustment strategy is proposed that refines the prediction errors of two color channels by utilizing the error distribution of a reference channel. This lead to a even sharper PEH, representing further enhanced prediction performance.

The experimental analysis showcased that the PEH obtained using the proposed SA-CNN achieves the highest peak at zero among all the existing predictors. In other words, the SA-CNN increased the pixel concentration within the range of [-2,2] by 13% and 20% compared to the best-performing SOTA predictors. A further experimentation studied that the proposed SA-CNN, combined with the reference error adjustment strategy, achieves the highest frequency at zero prediction error, showing an improvement of approximately 110% over the best-performing existing method. Experimentation around the quality parameters highlighted that the proposed SA-CNNP achieves the lowest average mean of 2.99, indicating that the PEs generated by the SA-CNN are significantly smaller compared to those produced by SOTA methods. Furthermore, with an achievement of an average PSNR value of approximately 1.2 dB higher than the SOTA method, the SA-CNN-enhanced RDH method demonstrated superior embedding performance.

- **UMANeT: A Two-Stage Interpolation-Based Reversible Data Hiding Framework with Attention-Enhanced Prediction:** Interpolation-Based RDH methods have gained popularity in applications where size of the original cover images is relatively small, limiting their embedding capacity. In the realm RDH, the interpolation techniques play crucial role in enhancing the resolution of the cover image, providing more space for data hiding, without causing distortion to the original image. However, the performance of the SOTA methods is limited due to their limitation of using traditional interpolation techniques that do not provide promising results. To improve upon these limitations, this work introduced a two-stage approach that utilized bicubic interpolation in first stage followed by a deep learning based interpolation in second stage. The study introduced a novel predictor UMANet, which is a U-Net-inspired architecture that integrates Channel Attention (CA) and multi-head attention (MHA) in parallel.

  Experimental results demonstrated that the proposed UMANet predictor achieves the highest PSNR consistently for all test images, with approximately 2dB improvement in PSNR for popular images, such as Lena and Couple. The proposed method further demonstrated remarkable embedding performance across all standard test images, demonstrating the superiority of the proposed UMANet predictor-based two-stage Interpolation methodology.

- **SCAM-Net: Spatial-Channel Attention Multi-Scale Network for Reversible Data**

Hiding in Encrypted Images: In RDHEI techniques, the prediction accuracy directly translates into embedding performance as the pixels with more accurate predictions can be compressed into smaller number of bits. However, the traditional techniques lack the prediction accuracy due to their limitation of using linear calculations for prediction. The proposed study mitigated this limitation by proposing a novel predictor that utilizes deep learning based predictor. The proposed SCAM-Net predictor incorporates channel attention and spatial attention to initially emphasize important regions for feature selection, and subsequently refining feature maps adaptively, enhancing prediction performance.

Experimental results demonstrated that, among all the best-known predictors, the SCAM-Net predictor achieved superior performance by recording the lowest values for Mean, Variance, and MSE. The higher embedding capacity along with high PSNR achieved by the proposed RDHEI technique demonstrated that the higher prediction accuracy lead to a smaller length of label map, resulting in improved payload embedding.

7.2 Rationale for multiple Deep Learning Architectures

The thesis introduces three deep learning architectures: SA-CNNP, UMANet, and SCAM-Net, each designed to address a different category of prediction challenges encountered within the RDH. While the universal applicability of the three predictors is validated by extensive experimental analysis over diverse characteristic test images, each predictor is designed to cater varying need of different application areas. SA-CNNP, a self-attention-based CNN predictor, is suitable for tasks requiring a balance of local convolutional processing and spatially distant context, which is typically important in medical images that contain repetitive texture. UMA-Net, a multi-head attention-based U-Net type architecture that helps the predictor to capture long range dependencies, is particularly beneficial for images that preserve structural continuity. SCAM-Net, that utilizes multi-scale network combined with spatial and channel attention, providing deeper feature refinement, thus focusing on the tasks demanding fine-grained feature refinement. Using specialized architectures ensures the adaptability of the proposed methods to medical and other critical application areas.

7.3 Future Scope

The domain of RDH continues to expand with promising directions emerging through the integration of advanced deep learning models, improved prediction mechanisms, and adaptive embedding strategies. While substantial progress has been achieved, several avenues remain open for further research and innovation.

- **Enhancing Prediction Accuracy through Advanced Deep Learning Models-** Recent studies employing convolutional neural networks have shown remarkable improvement

in prediction-based embedding performance. However, advanced architectures such as attention mechanisms, transformers, and causal predictors remain underexplored for RDH tasks. Future research can focus on integrating these models to capture complex spatial dependencies, enabling more precise prediction and higher embedding efficiency.

- **Leveraging Inter-Channel Correlation for Color Image RDH-** With the dominance of color imagery in digital platforms, exploiting inter-channel dependencies among RGB components offers significant potential. Future work may develop deep learning–driven frameworks that jointly analyze and embed data across channels, instead of treating them independently. Such approaches could substantially improve embedding capacity and maintain visual fidelity, particularly in color image authentication and secure sharing scenarios.

- **Deep Learning–Based Image Segmentation for Medical Applications-** In medical imaging, separating regions of interest (ROI) from non-regions of interest (NROI) is crucial for adaptive data hiding. Conventional segmentation techniques, such as Otsu [188] and ATD [189], often struggle with heterogeneous image characteristics. Deep neural networks, especially encoder–decoder or transformer-based segmentation models, can be utilized to achieve precise region delineation, allowing customized embedding strategies and enhanced reversibility in RDH for medical images.

- **Advancing RDH in Encrypted Images (RDHEI)-** In RDHEI, efficient space reservation before or after encryption remains a key challenge. Deep learning can play a pivotal role in learning optimal embedding locations and improving space vacating strategies without compromising encryption integrity. Future research can focus on intelligent RDHEI frameworks that adaptively balance data hiding capacity, security, and reconstruction quality.

- **RDH for Adversarial and Secure AI Applications-** A growing application of RDH lies in generating reversible adversarial examples, which embed perturbations that mislead classification models while allowing perfect image recovery [34]. Future studies can explore adaptive adversarial embedding guided by deep generative or reinforcement learning models, aiming to maintain both strong adversarial effectiveness and high visual quality. Moreover, RDH can be further extended to support secure model training and watermarking in adversarial environments.

- **Deep Learning–Assisted RDH for Embedding Optimization and Overhead Reduction-** In most RDH frameworks, auxiliary information such as location maps introduces redundancy and limits embedding capacity [190]. Future directions include designing neural predictors that automatically infer embeddable regions, reducing or eliminating the need for explicit location maps. Additionally, learning-based adaptive embedding strategies

can be developed to dynamically adjust payload and distortion trade-offs, minimizing overhead while maintaining reversibility [191].

- **Security Validation against Deep-Learning-based Steganalysis-** The security validation conducted in this work primarily relies on the traditional steganalysis techniques, such as, RS (Regular–Singular) analysis method. Historically, RS steganalysis has been a highly effective technique to validate the robustness, it is now considered limited when compared to modern steganalysis approaches, particularly utilising deep learning–based CNN detectors. Off-lately, CNN and transformer-based steganalysis techniques have exhibited strong performance in identifying subtle embedding artifacts that traditional statistical methods such as RS analysis may miss. The future work should extend the security validation by incorporating advanced deep-learning-based steganalysis techniques.

- **Security Analysis of underlying Encryption Technique-** The proposed RDHEI method utilizes multi-scale attention based predictor, where prediction is applied before encryption. This ensures the effective utilization of pixel correlations, leading to improved embedding performance. The method utilizes traditional pixel-wise XOR encryption that secures the data against potential attacks. However, the future work could investigate whether the predictor-induced embedding pattern introduces any detectable statistical patterns prior to encryption, evaluating its robustness against plain-text attacks.

7.4 Societal Applications

The wider societal implication of this research lies in strengthening reliability in digitization of healthcare industry as well as promoting assurance in the integrity of sensitive data storage.

- **Safeguarding patient confidentiality-** Strengthening the confidentiality of patient records, nurturing confidence in digital healthcare systems.

- **National Security and Defense-** Strengthens covert exchange of sensitive information and reinforcing resilience against data tampering.

- **Digital Rights Management-** Combat plagiarism and digital forgery while ensuring original media reversibility.

- **Cloud Security-** Securing public trust in cloud-based services, e-governance, and content sharing platforms.

BIBLIOGRAPHY

[1] G. Gao, S. Tong, Z. Xia, B. Wu, L. Xu, and Z. Zhao, "Reversible data hiding with automatic contrast enhancement for medical images," *Signal Processing*, vol. 178, p. 107817, 2021.

[2] "The national library of medicine presents medpix," https://medpix.nlm.nih.gov/home.

[3] "National cancer institute, national biomedical imaging archive," https://imaging.nci.nih.gov.

[4] K. Clark, B. Vendt, K. Smith, J. Freymann, J. Kirby, P. Koppel, S. Moore, S. Phillips, D. Maffitt, M. Pringle *et al.*, "The cancer imaging archive (tcia): maintaining and operating a public information repository," *Journal of digital imaging*, vol. 26, pp. 1045–1057, 2013.

[5] H. J. Shim, J. Ahn, and B. Jeon, "Dh-lzw: lossless data hiding in lzw compression," in *2004 International Conference on Image Processing, 2004. ICIP'04.*, vol. 4. IEEE, 2004, pp. 2195–2198.

[6] G. Coatrieux, C. Le Guillou, J.-M. Cauvin, and C. Roux, "Reversible watermarking for knowledge digest embedding and reliability control in medical images," *IEEE Transactions on Information Technology in Biomedicine*, vol. 13, no. 2, pp. 158–165, 2008.

[7] J. M. Barton, "Method and apparatus for embedding authentication information within digital data," *United States Patent, 5 646 997*, 1997.

[8] Z. Ni, Y.-Q. Shi, N. Ansari, and W. Su, "Reversible data hiding," *IEEE Transactions on circuits and systems for video technology*, vol. 16, no. 3, pp. 354–362, 2006.

[9] R. Kumar, D. Sharma, A. Dua, and K.-H. Jung, "A review of different prediction methods for reversible data hiding," *Journal of Information Security and Applications*, vol. 78, p. 103572, 2023.

[10] A. M. Alattar, "Reversible watermark using the difference expansion of a generalized integer transform," *IEEE transactions on image processing*, vol. 13, no. 8, pp. 1147–1156, 2004.

[11] G. Kaur, S. Singh, R. Rani, and R. Kumar, "A comprehensive study of reversible data hiding (rdh) schemes based on pixel value ordering (pvo)," *Archives of Computational Methods in Engineering*, vol. 28, pp. 3517–3568, 2021.

[12] D. Hou, H. Wang, W. Zhang, and N. Yu, "Reversible data hiding in jpeg image based on dct frequency and block selection," *Signal Processing*, vol. 148, pp. 41–47, 2018.

[13] C.-C. Lin and P.-F. Shiu, "Dct-based reversible data hiding scheme," in *Proceedings of the 3rd International Conference on Ubiquitous Information Management and Communication*, 2009, pp. 327–335.

[14] T.-S. Nguyen, C.-C. Chang, and X.-Q. Yang, "A reversible image authentication scheme based on fragile watermarking in discrete wavelet transform domain," *AEU-International Journal of Electronics and Communications*, vol. 70, no. 8, pp. 1055–1061, 2016.

[15] F. Li, Q. Mao, and C.-C. Chang, "Reversible data hiding scheme based on the haar discrete wavelet transform and interleaving prediction method," *Multimedia Tools and Applications*, vol. 77, pp. 5149–5168, 2018.

[16] S. Lee, C. D. Yoo, and T. Kalker, "Reversible image watermarking based on integer-to-integer wavelet transform," *IEEE Transactions on information forensics and security*, vol. 2, no. 3, pp. 321–330, 2007.

[17] A. Shaik and V. Thanikaiselvan, "Comparative analysis of integer wavelet transforms in reversible data hiding using threshold based histogram modification," *Journal of King Saud University-Computer and Information Sciences*, vol. 33, no. 7, pp. 878–889, 2021.

[18] G. Kasana, K. Singh, and S. S. Bhatia, "Data hiding algorithm for images using discrete wavelet transform and arnold transform," *Journal of Information Processing Systems*, vol. 13, no. 5, pp. 1331–1344, 2017.

[19] A. M. Kapadia and P. Nithyanandam, "A recursive high payload reversible data hiding using integer wavelet and arnold transform." *Intelligent Automation & Soft Computing*, vol. 35, no. 1, 2023.

[20] R. Kumar, S. Chand, and S. Singh, "An optimal high capacity reversible data hiding scheme using move to front coding for lzw codes," *Multimedia Tools and Applications*, vol. 78, pp. 22 977–23 001, 2019.

[21] C.-C. Chang and C.-Y. Lin, "Reversible steganographic method using smvq approach based on declustering," *Information Sciences*, vol. 177, no. 8, pp. 1796–1805, 2007.

[22] L. Wang, Z. Pan, X. Ma, and S. Hu, "A novel high-performance reversible data hiding scheme using smvq and improved locally adaptive coding method," *Journal of Visual Communication and Image Representation*, vol. 25, no. 2, pp. 454–465, 2014.

[23] C.-C. Lin, X.-L. Liu, W.-L. Tai, and S.-M. Yuan, "A novel reversible data hiding scheme based on ambtc compression technique," *Multimedia Tools and Applications*, vol. 74, pp. 3823–3842, 2015.

[24] R. Kumar and K.-H. Jung, "A systematic survey on block truncation coding based data hiding techniques," *Multimedia Tools and Applications*, vol. 78, no. 22, pp. 32 239–32 259, 2019.

[25] A. Malik, G. Sikka, and H. K. Verma, "An ambtc compression based data hiding scheme using pixel value adjusting strategy," *Multidimensional Systems and Signal Processing*, vol. 29, pp. 1801–1818, 2018.

[26] V. Sachnev, H. J. Kim, J. Nam, S. Suresh, and Y. Q. Shi, "Reversible watermarking algorithm using sorting and prediction," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 19, no. 7, pp. 989–999, 2009.

[27] R. Kumar, S. Chand, and S. Singh, "An improved histogram-shifting-imitated reversible data hiding based on hvs characteristics," *Multimedia Tools and Applications*, vol. 77, pp. 13 445–13 457, 2018.

[28] B. Ma and Y. Q. Shi, "A reversible data hiding scheme based on code division multiplexing," *IEEE Transactions on Information Forensics and Security*, vol. 11, no. 9, pp. 1914–1927, 2016.

[29] B. Ma, B. Li, X.-Y. Wang, C.-P. Wang, J. Li, and Y.-Q. Shi, "Code division multiplexing and machine learning based reversible data hiding scheme for medical image," *Security and Communication Networks*, vol. 2019, no. 1, p. 4732632, 2019.

[30] J. Tian, "Reversible data embedding using a difference expansion," *IEEE transactions on circuits and systems for video technology*, vol. 13, no. 8, pp. 890–896, 2003.

[31] D. M. Thodi and J. J. Rodríguez, "Expansion embedding techniques for reversible watermarking," *IEEE transactions on image processing*, vol. 16, no. 3, pp. 721–730, 2007.

[32] D. Coltuc, "Improved embedding for prediction-based reversible watermarking," *IEEE Transactions on Information Forensics and security*, vol. 6, no. 3, pp. 873–882, 2011.

[33] G. Gao, S. Tong, Z. Xia, B. Wu, L. Xu, and Z. Zhao, "Reversible data hiding with automatic contrast enhancement for medical images," *Signal Processing*, vol. 178, p. 107817, 2021.

[34] Z. Yin, Y. Xiang, and X. Zhang, "Reversible data hiding in encrypted images based on multi-msb prediction and huffman coding," *IEEE Transactions on Multimedia*, vol. 22, no. 4, pp. 874–884, 2019.

[35] Z. Yinn, Y. Xiang, and X. Zhang, "Reversible data hiding in encrypted images based on multi-msb prediction and huffman coding," *IEEE Transactions on Multimedia*, vol. 22, no. 4, pp. 874–884, 2019.

[36] R. Kumar, A. K. Sharma *et al.*, "High capacity reversible data hiding with contiguous space in encrypted images," *Computers and Electrical Engineering*, vol. 112, p. 109017, 2023.

[37] K.-H. Jung and K.-Y. Yoo, "Data hiding method using image interpolation," *Computer Standards & Interfaces*, vol. 31, no. 2, pp. 465–470, 2009.

[38] K.-H. Jung, "A survey of interpolation-based reversible data hiding methods," *Multimedia Tools and Applications*, vol. 77, pp. 7795–7810, 2018.

[39] J. Liu, W. Zhang, K. Fukuchi, Y. Akimoto, and J. Sakuma, "Unauthorized ai cannot recognize me: Reversible adversarial example," *Pattern Recognition*, vol. 134, p. 109048, 2023.

[40] R. Kumar, R. Caldelli, K. Wong, A. Malik, and K.-H. Jung, "High-fidelity reversible data hiding using novel comprehensive rhombus predictor," *Multimedia Tools and Applications*, pp. 1–23, 2024.

[41] N. Kumar, R. Kumar, and R. Caldelli, "Local moment driven pvo based reversible data hiding," *IEEE Signal Processing Letters*, vol. 28, pp. 1335–1339, 2021.

[42] H.-W. Tseng and C.-P. Hsieh, "Prediction-based reversible data hiding," *Information Sciences*, vol. 179, no. 14, pp. 2460–2469, 2009.

[43] F. Di, M. Zhang, X. Liao, and J. Liu, "High-fidelity reversible data hiding by quadtree-based pixel value ordering," *Multimedia Tools and Applications*, vol. 78, pp. 7125–7141, 2019.

[44] S. Meikap, B. Jana, and T.-C. Lu, "Context pixel-based reversible data hiding scheme using pixel value ordering," *The Visual Computer*, vol. 40, no. 5, pp. 3529–3552, 2024.

[45] T. Zhang, X. Li, W. Qi, and Z. Guo, "Prediction-error value ordering for high-fidelity reversible data hiding," in *International Conference on Multimedia Modeling*. Springer, 2019, pp. 317–328.

[46] J. Qin and F. Huang, "Reversible data hiding based on prediction-error-ordering," in *Pattern Recognition and Computer Vision: Third Chinese Conference, PRCV 2020, Nanjing, China, October 16–18, 2020, Proceedings, Part I 3*. Springer, 2020, pp. 527–537.

[47] H.-T. Wu, J.-L. Dugelay, and Y.-Q. Shi, "Reversible image data hiding with contrast enhancement," *IEEE signal processing letters*, vol. 22, no. 1, pp. 81–85, 2014.

[48] H.-T. Wu, J. Huang, and Y.-Q. Shi, "A reversible data hiding method with contrast enhancement for medical images," *Journal of Visual Communication and Image Representation*, vol. 31, pp. 146–153, 2015.

[49] H.-T. Wu, S. Tang, J. Huang, and Y.-Q. Shi, "A novel reversible data hiding method with image contrast enhancement," *Signal Processing: Image Communication*, vol. 62, pp. 64–73, 2018.

[50] J. Huang, G. Huang, X. Zhang, X. Yuan, F. Xie, C.-M. Pun, and G. Zhong, "Black-box reversible adversarial examples with invertible neural network," *Image and Vision Computing*, vol. 147, p. 105094, 2024.

[51] L. Xiong, Y. Wu, P. Yu, and Y. Zheng, "A black-box reversible adversarial example for authorizable recognition to shared images," *Pattern Recognition*, vol. 140, p. 109549, 2023.

[52] Y. Wang, G. Xiong, and W. He, "High-capacity reversible data hiding in encrypted images based on pixel-value-ordering and histogram shifting," *Expert Systems with Applications*, vol. 211, p. 118600, 2023.

[53] Z. Han and C. Guohua, "Reversible data hiding in encrypted image with local-correlation-based classification and adaptive encoding strategy," *Signal Processing*, vol. 205, p. 108847, 2023.

[54] R. Kumar, A. K. Sharma *et al.*, "Link chain driven reversible data hiding in encrypted images for high payload," *Signal, Image and Video Processing*, pp. 1–16, 2024.

[55] X. Liu, G. Han, J. Wu, Z. Shao, G. Coatrieux, and H. Shu, "Fractional krawtchouk transform with an application to image watermarking," *IEEE Transactions on Signal Processing*, vol. 65, no. 7, pp. 1894–1908, 2017.

[56] X. Wang, X. Li, and Q. Pei, "Independent embedding domain based two-stage robust reversible watermarking," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 30, no. 8, pp. 2406–2417, 2019.

[57] R. Kumar and K.-H. Jung, "Robust reversible data hiding scheme based on two-layer embedding strategy," *Information Sciences*, vol. 512, pp. 96–107, 2020.

[58] Y. V. Singh, S. Khan, S. K. Shukla, and K.-H. Jung, "Local-moment-driven robust reversible data hiding," *Applied Sciences*, vol. 12, no. 22, p. 11826, 2022.

[59] R. Hu and S. Xiang, "Cover-lossless robust image watermarking against geometric deformations," *IEEE Transactions on Image Processing*, vol. 30, pp. 318–331, 2020.

[60] R. Hux and S. Xiang, "Lossless robust image watermarking by using polar harmonic transform," *Signal Processing*, vol. 179, p. 107833, 2021.

[61] Y. Tang, S. Wang, C. Wang, S. Xiang, and Y.-M. Cheung, "A highly robust reversible watermarking scheme using embedding optimization and rounded error compensation," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 33, no. 4, pp. 1593–1609, 2022.

[62] Y. Liu, L. Chen, M. Hu, Z. Jia, S. Jia, and H. Zhao, "A reversible data hiding method for h. 264 with shamir's (t, n)-threshold secret sharing," *Neurocomputing*, vol. 188, pp. 63–70, 2016.

[63] X. Jin, L. Su, and J. Huang, "A reversible data hiding algorithm based on secret sharing," *Journal of Information Hiding and Privacy Protection*, vol. 3, no. 2, p. 69, 2021.

[64] L. Xiong, X. Han, C.-N. Yang, and Z. Xia, "Rdh-des: Reversible data hiding over distributed encrypted-image servers based on secret sharing," *ACM Transactions on Multimedia Computing, Communications and Applications*, vol. 19, no. 1, pp. 1–19, 2023.

[65] Z. Ni, Y. Q. Shi, N. Ansari, W. Su, Q. Sun, and X. Lin, "Robust lossless image data hiding designed for semi-fragile image authentication," *IEEE Transactions on circuits and systems for video technology*, vol. 18, no. 4, pp. 497–509, 2008.

[66] X.-T. Zeng, L.-D. Ping, and X.-Z. Pan, "A lossless robust data hiding scheme," *Pattern Recognition*, vol. 43, no. 4, pp. 1656–1667, 2010.

[67] X. Xiong, S. Zhong, and Y. Lu, "Separable and reversible data hiding scheme for medical images using modified logistic and interpolation," *Biomedical Signal Processing and Control*, vol. 87, p. 105521, 2024.

[68] G. Gao, H. Zhang, Z. Xia, X. Luo, and Y.-Q. Shi, "Reversible data hiding-based contrast enhancement with multi-group stretching for roi of medical image," *IEEE Transactions on Multimedia*, 2023.

[69] A. G. Weber, "The usc-sipi image database: Version 5," *http://sipi. usc. edu/database/*, 2006.

[70] G. Schaefer and M. Stich, "Ucid: An uncompressed color image database," in *Storage and retrieval methods and applications for multimedia 2004*, vol. 5307. SPIE, 2003, pp. 472–480.

[71] P. Bas, T. Filler, and T. Pevnỳ, "Break our steganographic system: the ins and outs of organizing boss," in *International workshop on information hiding*. Springer, 2011, pp. 59–70.

[72] P. Bas and T. Furon, "BOWS-2 Web page," http://bows2.ec-lille.fr/, accessed: 18 Aug 2022.

[73] C. Chen *et al.*, *CiteSpace: a practical guide for mapping scientific literature*. Nova Science Publishers Hauppauge, NY, USA, 2016.

[74] N. J. van Eck and L. Waltman, "Software survey: Vosviewer, a computer program for bibliometric mapping," *Scientometrics*, vol. 84, no. 2, pp. 523–538, 2010.

[75] M. Mauri, T. Elli, G. Caviglia, G. Uboldi, and M. Azzi, "Rawgraphs: a visualisation platform to create open outputs," in *Proceedings of the 12th biannual conference on Italian SIGCHI chapter*, 2017, pp. 1–5.

[76] Y.-Q. Shi, X. Li, X. Zhang, H.-T. Wu, and B. Ma, "Reversible data hiding: Advances in the past two decades," *IEEE access*, vol. 4, pp. 3210–3237, 2016.

[77] R. Kumar, A. K. Sharma *et al.*, "Bit-plane based reversible data hiding in encrypted images using multi-level blocking with quad-tree," *IEEE Transactions on Multimedia*, 2023.

[78] P. Amrit, K. Singh, N. Baranwal, A. Singh, and J. Singh, "Deep learning-based segmentation for medical data hiding with galois field," *Neural Computing and Applications*, pp. 1–16, 2023.

[79] G. Gao, T. Han, B. Wu, J. Fu, and Z. Xia, "A hue preservation lossless contrast enhancement method with rdh for color images," *Digital Signal Processing*, vol. 136, p. 103965, 2023.

[80] R. Hu and S. Xiang, "Cnn prediction based reversible data hiding," *IEEE Signal Processing Letters*, vol. 28, pp. 464–468, 2021.

[81] R. Hun and S. Xiang, "Reversible data hiding by using cnn prediction and adaptive embedding," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 44, no. 12, pp. 10 196–10 208, 2021.

[82] S. Kim, R. Lussi, X. Qu, F. Huang, and H. J. Kim, "Reversible data hiding with automatic brightness preserving contrast enhancement," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 29, no. 8, pp. 2271–2284, 2018.

[83] Y. Yang, W. Zhang, D. Liang, and N. Yu, "A roi-based high capacity reversible data hiding scheme with contrast enhancement for medical images," *Multimedia Tools and Applications*, vol. 77, pp. 18 043–18 065, 2018.

[84] M. Shi, Y. Yang, J. Meng, and W. Zhang, "Reversible data hiding with enhancing contrast and preserving brightness in medical image," *Journal of Information Security and Applications*, vol. 70, p. 103324, 2022.

[85] D. Nuñez-Ramirez, D. Mata-Mendoza, and M. Cedillo-Hernandez, "Improving preprocessing in reversible data hiding based on contrast enhancement," *Journal of King Saud University-Computer and Information Sciences*, vol. 34, no. 8, pp. 5468–5477, 2022.

[86] H. Chen, J. Ni, W. Hong, and T.-S. Chen, "Reversible data hiding with contrast enhancement using adaptive histogram shifting and pixel value ordering," *Signal Processing: Image Communication*, vol. 46, pp. 1–16, 2016.

[87] P. Puteaux and W. Puech, "An efficient msb prediction-based method for high-capacity reversible data hiding in encrypted images," *IEEE transactions on information forensics and security*, vol. 13, no. 7, pp. 1670–1681, 2018.

[88] S. Yi and Y. Zhou, "Separable and reversible data hiding in encrypted images using parametric binary tree labeling," *IEEE Transactions on Multimedia*, vol. 21, no. 1, pp. 51–64, 2018.

[89] Y. Wu, Y. Xiang, Y. Guo, J. Tang, and Z. Yin, "An improved reversible data hiding in encrypted images using parametric binary tree labeling," *IEEE Transactions on Multimedia*, vol. 22, no. 8, pp. 1929–1938, 2019.

[90] A. Mohammadi, M. Nakhkash, and M. A. Akhaee, "A high-capacity reversible data hiding in encrypted images employing local difference predictor," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 30, no. 8, pp. 2366–2376, 2020.

[91] C. Yu, X. Zhang, X. Zhang, G. Li, and Z. Tang, "Reversible data hiding with hierarchical embedding for encrypted images," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 32, no. 2, pp. 451–466, 2021.

[92] K. Chen and C.-C. Chang, "High-capacity reversible data hiding in encrypted images based on extended run-length coding and block-based msb plane rearrangement," *Journal of Visual Communication and Image Representation*, vol. 58, pp. 334–344, 2019.

[93] S. Xu, J.-H. Horng, C.-C. Chang, and C.-C. Chang, "Reversible data hiding with hierarchical block variable length coding for cloud security," *IEEE transactions on dependable and secure computing*, vol. 20, no. 5, pp. 4199–4213, 2022.

[94] P. Ping, J. Huo, and B. Guo, "Novel asymmetric cnn-based and adaptive mean predictors for reversible data hiding in encrypted images," *Expert Systems with Applications*, vol. 246, p. 123270, 2024.

[95] X. Zhang, F. He, C. Yu, X. Zhang, C.-N. Yang, and Z. Tang, "Reversible data hiding in encrypted images with asymmetric coding and bit-plane block compression," *IEEE Transactions on Multimedia*, 2024.

[96] W. Hong, T.-S. Chen, and C.-W. Shiu, "Reversible data hiding for high quality images using modification of prediction errors," *Journal of Systems and Software*, vol. 82, no. 11, pp. 1833–1842, 2009.

[97] V. Conotter, G. Boato, M. Carli, and K. Egiazarian, "Near lossless reversible data hiding based on adaptive prediction," in *2010 IEEE International Conference on Image Processing*. IEEE, 2010, pp. 2585–2588.

[98] Q. Chang *et al.*, "Reversible data hiding for color images based on adaptive three-dimensional histogram modification," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 32, no. 9, pp. 5725–5735, 2022.

[99] A. Avramović and S. Savić, "Lossless predictive compression of medical images," *Serbian Journal of Electrical Engineering*, vol. 8, no. 1, pp. 27–36, 2011.

[100] C. Dragoi and D. Coltuc, "Improved rhombus interpolation for reversible watermarking by difference expansion," in *2012 Proceedings of the 20th European Signal Processing Conference (EUSIPCO)*. IEEE, 2012, pp. 1688–1692.

[101] B. Ma, X. Wang, Q. Li, B. Li, J. Li, C. Wang, and Y. Shi, "Adaptive error prediction method based on multiple linear regression for reversible data hiding," *Journal of Real-Time Image Processing*, vol. 16, pp. 821–834, 2019.

[102] H. J. Hwang, S. Kim, and H. J. Kim, "Reversible data hiding using least square predictor via the lasso," *EURASIP Journal on Image and Video Processing*, vol. 2016, pp. 1–12, 2016.

[103] Y. Wang and W. He, "High capacity reversible data hiding in encrypted image based on adaptive msb prediction," *IEEE Transactions on Multimedia*, vol. 24, pp. 1288–1298, 2021.

[104] T. Zong, Y. Xiang, I. Natgunanathan, S. Guo, W. Zhou, and G. Beliakov, "Robust histogram shape-based method for image watermarking," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 25, no. 5, pp. 717–729, 2014.

[105] D. Coltuc and J.-M. Chassery, "Distortion-free robust watermarking: a case study," in *Security, Steganography, and Watermarking of Multimedia Contents IX*, vol. 6505. SPIE, 2007, pp. 585–592.

[106] W. Wang, J. Ye, T. Wang, and W. Wang, "Reversible data hiding scheme based on significant-bit-difference expansion," *IET image processing*, vol. 11, no. 11, pp. 1002–1014, 2017.

[107] B. Ma, Z. Tao, R. Ma, C. Wang, J. Li, and X. Li, "A high-performance robust reversible data hiding algorithm based on polar harmonic fourier moments," *IEEE Transactions on Circuits and Systems for Video Technology*, 2023.

[108] L.-C. Huang, L.-Y. Tseng, and M.-S. Hwang, "A reversible data hiding method by histogram shifting in high quality medical images," *Journal of Systems and Software*, vol. 86, no. 3, pp. 716–727, 2013.

[109] Z. Ni, Y.-Q. Shi, N. Ansari, and W. Su, "Reversible data hiding," *IEEE Transactions on circuits and systems for video technology*, vol. 16, no. 3, pp. 354–362, 2006.

[110] O. M. Al-Qershi and B. E. Khoo, "High capacity data hiding schemes for medical images based on difference expansion," *Journal of Systems and Software*, vol. 84, no. 1, pp. 105–112, 2011.

[111] C. Qin, C.-C. Chang, and T.-J. Hsu, "Reversible data hiding scheme based on exploiting modification direction with two steganographic images," *Multimedia Tools and Applications*, vol. 74, pp. 5861–5872, 2015.

[112] Y.-Q. Shi, X. Li, X. Zhang, H.-T. Wu, and B. Ma, "Reversible data hiding: advances in the past two decades," *IEEE access*, vol. 4, pp. 3210–3237, 2016.

[113] L.-C. Huang, L.-Y. Tseng, and M.-S. Hwang, "A reversible data hiding method by histogram shifting in high quality medical images," *Journal of Systems and Software*, vol. 86, no. 3, pp. 716–727, 2013.

[114] H.-T. Wu, J.-L. Dugelay, and Y.-Q. Shi, "Reversible image data hiding with contrast enhancement," *IEEE signal processing letters*, vol. 22, no. 1, pp. 81–85, 2014.

[115] H.-T. Wu, J. Huang, and Y.-Q. Shi, "A reversible data hiding method with contrast enhancement for medical images," *Journal of Visual Communication and Image Representation*, vol. 31, pp. 146–153, 2015.

[116] H.-T. Wu, S. Tang, J. Huang, and Y.-Q. Shi, "A novel reversible data hiding method with image contrast enhancement," *Signal Processing: Image Communication*, vol. 62, pp. 64–73, 2018.

[117] S. Kim, R. Lussi, X. Qu, and H. J. Kim, "Automatic contrast enhancement using reversible data hiding," in *2015 IEEE International Workshop on Information Forensics and Security (WIFS)*. IEEE, 2015, pp. 1–5.

[118] S. Kim, R. Lussi, X. Qu, F. Huang, and H. J. Kim, "Reversible data hiding with automatic brightness preserving contrast enhancement," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 29, no. 8, pp. 2271–2284, 2018.

[119] G. Gao, X. Wan, S. Yao, Z. Cui, C. Zhou, and X. Sun, "Reversible data hiding with contrast enhancement and tamper localization for medical images," *Information Sciences*, vol. 385, pp. 250–265, 2017.

[120] D. Nuñez-Ramirez, D. Mata-Mendoza, and M. Cedillo-Hernandez, "Improving preprocessing in reversible data hiding based on contrast enhancement," *Journal of King Saud University-Computer and Information Sciences*, vol. 34, no. 8, pp. 5468–5477, 2022.

[121] Y. Liu, X. Qu, and G. Xin, "A roi-based reversible data hiding scheme in encrypted medical images," *Journal of Visual Communication and Image Representation*, vol. 39, pp. 51–57, 2016.

[122] Y. Yang, W. Zhang, D. Liang, and N. Yu, "A roi-based high capacity reversible data hiding scheme with contrast enhancement for medical images," *Multimedia Tools and Applications*, vol. 77, pp. 18 043–18 065, 2018.

[123] H.-T. Wu, W. Mai, S. Meng, Y.-M. Cheung, and S. Tang, "Reversible data hiding with image contrast enhancement based on two-dimensional histogram modification," *IEEE Access*, vol. 7, pp. 83 332–83 342, 2019.

[124] H. Chen, J. Yuan, W. Hong, J. Ni, and T.-S. Chen, "On performance improvement of reversible data hiding with contrast enhancement," *The Computer Journal*, vol. 63, no. 10, pp. 1584–1596, 2020.

[125] S. Mansouri, H. K. Bizaki, and M. Fakhredanesh, "Reversible data hiding with automatic contrast enhancement using two-sided histogram expansion," *Journal of Visual Communication and Image Representation*, vol. 81, p. 103359, 2021.

[126] M. Shi, Y. Yang, J. Meng, and W. Zhang, "Reversible data hiding with enhancing contrast and preserving brightness in medical image," *Journal of Information Security and Applications*, vol. 70, p. 103324, 2022.

[127] H. Chen, J. Ni, W. Hong, and T.-S. Chen, "Reversible data hiding with contrast enhancement using adaptive histogram shifting and pixel value ordering," *Signal Processing: Image Communication*, vol. 46, pp. 1–16, 2016.

[128] W. Lyu, Y. Yue, and Z. Yin, "Reversible data hiding based on automatic contrast enhancement using histogram expansion," *Journal of Visual Communication and Image Representation*, vol. 92, p. 103798, 2023.

[129] T. Zhang, T. Hou, S. Weng, F. Zou, H. Zhang, and C.-C. Chang, "Adaptive reversible data hiding with contrast enhancement based on multi-histogram modification," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 32, no. 8, pp. 5041–5054, 2022.

[130] P.-Y. Pai, C.-C. Chang, Y.-K. Chan, and M.-H. Tsai, "An adaptable threshold detector," *Information Sciences*, vol. 181, no. 8, pp. 1463–1483, 2011.

[131] N. Otsu, "A threshold selection method from gray-level histograms," *IEEE transactions on systems, man, and cybernetics*, vol. 9, no. 1, pp. 62–66, 1979.

[132] J. Fridrich, M. Goljan, and R. Du, "Reliable detection of lsb steganography in color and grayscale images," in *Proceedings of the 2001 workshop on Multimedia and security: new challenges*, 2001, pp. 27–30.

[133] P. C. Mandal, I. Mukherjee, and B. Chatterji, "High capacity steganography based on iwt using eight-way cvd and n-lsb ensuring secure communication," *Optik*, vol. 247, p. 167804, 2021.

[134] M. Hassaballah, M. A. Hameed, S. Aly, and A. AbdelRady, "A color image steganography method based on adpvd and hog techniques," in *Digital Media Steganography*. Elsevier, 2020, pp. 17–40.

[135] J. Chang, G. Zhu, H. Zhang, Y. Zhou, X. Luo, and L. Wu, "Reversible data hiding for color images based on adaptive 3d prediction-error expansion and double deep q-network," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 32, no. 8, pp. 5055–5067, 2022.

[136] D. He and Z. Cai, "Reversible data hiding for color images using channel reference mapping and adaptive pixel prediction," *Mathematics*, vol. 12, no. 4, p. 517, 2024.

[137] H. Wang, D. Wang, Z. Chu, Z. Rao, and Y. Yao, "Reversible data hiding for color images based on prediction-error value ordering and adaptive embedding," *Journal of Visual Communication and Image Representation*, vol. 103, p. 104239, 2024.

[138] R. Hu, Y. Wu, S. Xiang, X. Li, and Y. Zhao, "Deep prediction and efficient 3d mapping of color images for reversible data hiding," *IEEE Transactions on Information Forensics and Security*, 2025.

[139] J. Li, X. Li, and B. Yang, "Reversible data hiding scheme for color image based on prediction-error expansion and cross-channel correlation," *Signal Processing*, vol. 93, no. 9, pp. 2748–2758, 2013.

[140] B. Ou *et al.*, "Efficient color image reversible data hiding based on channel-dependent payload partition and adaptive embedding," *Signal Processing*, vol. 108, pp. 642–657, 2015.

[141] H. Yao, C. Qin, Z. Tang, and Y. Tian, "Guided filtering based color image reversible data hiding," *Journal of Visual Communication and Image Representation*, vol. 43, pp. 152–163, 2017.

[142] Z. Tang *et al.*, "Color image reversible data hiding with double-layer embedding," *IEEE Access*, vol. 8, pp. 6915–6926, 2020.

[143] M. Xu and J. Li, "3d pee mapping based reversible data hiding for color images," *Multimedia Tools and Applications*, vol. 78, no. 7, pp. 8003–8016, 2019.

[144] N. Mao *et al.*, "Reversible data hiding for color images based on pixel value order of overall process channel," *Signal Processing*, vol. 205, p. 108865, 2023.

[145] X. Yang and F. Huang, "New cnn-based predictor for reversible data hiding," *IEEE Signal Processing Letters*, vol. 29, pp. 2627–2631, 2022.

[146] M. Wu and S. Xiang, "An efficient cnn-based prediction for reversible data hiding," in *Proceedings of the 5th ACM International Conference on Multimedia in Asia*, 2023, pp. 1–5.

[147] Y. Luo, Y. Qiu, B. Lu, S. Qin, Q. Fu, S. Zhang, Y. Huang, and Y. Su, "Reversible data hiding based on optimized cnn predictor and prediction error expansion with lower surround background complexity," *Computers and Electrical Engineering*, vol. 119, p. 109472, 2024.

[148] B. Ma, H. Duan, R. Ma, Y. Xian, and X. Li, "High-performance optimization framework for reversible data hiding predictor," *IEEE Signal Processing Letters*, 2024.

[149] H. Rao, S. Weng, L. Yu, L. Li, and G. Cao, "High-precision reversible data hiding predictor: Ucanet," *IEEE Signal Processing Letters*, 2024.

[150] Y. Qiu, W. Peng, and X. Lin, "Improved cnn prediction based reversible data hiding for images," *Entropy*, vol. 27, no. 2, p. 159, 2025.

[151] A. Vaswani *et al.*, "Attention is all you need," *Advances in neural information processing systems*, vol. 30, 2017.

[152] J. Deng *et al.*, "Imagenet: A large-scale hierarchical image database," in *2009 IEEE conference on computer vision and pattern recognition*. Ieee, 2009, pp. 248–255.

[153] R. Franzen, "Kodak lossless true color image suite," *source: http://r0k. us/graphics/kodak*, vol. 4, no. 2, p. 9, 1999.

[154] F. S. Hassan and A. Gutub, "Efficient image reversible data hiding technique based on interpolation optimization," *Arabian Journal for Science and Engineering*, vol. 46, no. 9, pp. 8441–8456, 2021.

[155] C.-F. Lee and Y.-L. Huang, "An efficient image interpolation increasing payload in reversible data hiding," *Expert systems with applications*, vol. 39, no. 8, pp. 6712–6719, 2012.

[156] A. Malik, G. Sikka, and H. K. Verma, "Image interpolation based high capacity reversible data hiding scheme," *Multimedia Tools and Applications*, vol. 76, pp. 24 107–24 123, 2017.

[157] C. Hall, "Bicubic interpolation over triangles," *Journal of mathematics and mechanics*, vol. 19, no. 1, pp. 1–11, 1969.

[158] B. L. Hulme, "A new bicubic interpolation over right triangles," *Journal of Approximation Theory*, vol. 5, no. 1, pp. 66–73, 1972.

[159] P. Smith, "Bilinear interpolation of digital images," *Ultramicroscopy*, vol. 6, no. 2, pp. 201–204, 1981.

[160] X. Zhang, Z. Sun, Z. Tang, C. Yu, and X. Wang, "High capacity data hiding based on interpolated image," *Multimedia Tools and Applications*, vol. 76, pp. 9195–9218, 2017.

[161] F. S. Hassan and A. Gutub, "Novel embedding secrecy within images utilizing an improved interpolation-based reversible data hiding scheme," *Journal of King Saud University-Computer and Information Sciences*, vol. 34, no. 5, pp. 2017–2030, 2022.

[162] R. Hu and S. Xiang, "Reversible data hiding by using cnn prediction and adaptive embedding," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 44, no. 12, pp. 10 196–10 208, 2021.

[163] X. Yang, Y. Wang, and F. Huang, "Cnn-based reversible data hiding for jpeg images," *IEEE Transactions on Circuits and Systems for Video Technology*, 2024.

[164] D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization," *arXiv preprint arXiv:1412.6980*, 2014.

[165] H. Zhang, H. Sun, and F. Meng, "Reversible data hiding scheme based on improved interpolation and three-in-one intelligent strategy," *Journal of Information Security and Applications*, vol. 77, p. 103573, 2023.

[166] Y.-T. Chang, C.-T. Huang, C.-F. Lee, and S.-J. Wang, "Image interpolating based data hiding in conjunction with pixel-shifting of histogram," *The Journal of Supercomputing*, vol. 66, pp. 1093–1110, 2013.

[167] X. Xiong, Y. Chen, M. Fan, and S. Zhong, "Adaptive reversible data hiding algorithm for interpolated images using sorting and coding," *Journal of Information Security and Applications*, vol. 66, p. 103137, 2022.

[168] Y.-q. Chen, W.-j. Sun, L.-y. Li, C.-C. Chang, and X. Wang, "An efficient general data hiding scheme based on image interpolation," *Journal of Information Security and Applications*, vol. 54, p. 102584, 2020.

[169] X. Bai, Y. Chen, G. Duan, C. Feng, and W. Zhang, "A data hiding scheme based on the difference of image interpolation algorithms," *Journal of Information Security and Applications*, vol. 65, p. 103068, 2022.

[170] P. C. Mandal, I. Mukherjee, and B. N. Chatterji, "High capacity reversible and secured data hiding in images using interpolation and difference expansion technique," *Multimedia Tools and Applications*, vol. 80, no. 3, pp. 3623–3644, 2021.

[171] A. Malik, G. Sikka, and H. K. Verma, "A reversible data hiding scheme for interpolated images based on pixel intensity range," *Multimedia Tools and Applications*, vol. 79, pp. 18 005–18 031, 2020.

[172] S. Zhong, Y. Lu, and X. Xiong, "Reversible data hiding algorithm in encrypted domain based on image interpolation," *IEEE Access*, vol. 11, pp. 108 281–108 294, 2023.

[173] P. Bas and T. Furon, "Image database of bows-2," *Accessed: Jun*, vol. 20, pp. 2016–2017, 2017.

[174] P. Puteaux, S. Ong, K. Wong, and W. Puech, "A survey of reversible data hiding in encrypted images–the first 12 years," *Journal of Visual Communication and Image Representation*, vol. 77, p. 103085, 2021.

[175] Ankur, R. Kumar, and A. K. Sharma, "A systematic review of rdhei with consistent experimental evaluation," in *The International Conference on Recent Innovations in Computing*. Springer, 2023, pp. 15–27.

[176] K. Gupta, R. Kumar *et al.*, "Adaptive ambtc based reversible data hiding in encrypted images," in *2024 15th International Conference on Computing Communication and Networking Technologies (ICCCNT)*. IEEE, 2024, pp. 1–6.

[177] R. Kumar, P. Ranjan, K.-H. Jung *et al.*, "Leveraging rans for synchronized high capacity reversible data hiding in encrypted image," *Expert systems with applications*, vol. 267, p. 126181, 2025.

[178] K. Ma, W. Zhang, X. Zhao, N. Yu, and F. Li, "Reversible data hiding in encrypted images by reserving room before encryption," *IEEE Transactions on information forensics and security*, vol. 8, no. 3, pp. 553–562, 2013.

[179] H. Gao, X. Zhang, and T. Gao, "Hierarchical reversible data hiding in encrypted images based on multiple linear regressions and multiple bits prediction," *Multimedia Tools and Applications*, vol. 83, no. 3, pp. 8757–8783, 2024.

[180] D. M. Thodi and J. J. Rodríguez, "Expansion embedding techniques for reversible watermarking," *IEEE transactions on image processing*, vol. 16, no. 3, pp. 721–730, 2007.

[181] R. Bhardwaj and I. Tripathi, "An enhanced reversible data hiding algorithm using deep neural network for e-healthcare," *Journal of Ambient Intelligence and Humanized Computing*, vol. 14, no. 8, pp. 10 567–10 585, 2023.

[182] X. Chai, G. Cao, Z. Fu, Z. Gan, B. Wang, and Y. Zhang, "High-capacity reversible data hiding in encrypted medical images using adaptive pixel-modulation and hbp-rmc," *Biomedical Signal Processing and Control*, vol. 95, p. 106424, 2024.

[183] Y. Fu, P. Kong, H. Yao, Z. Tang, and C. Qin, "Effective reversible data hiding in encrypted image with adaptive encoding strategy," *Information Sciences*, vol. 494, pp. 21–36, 2019.

[184] S. Chen and C.-C. Chang, "Reversible data hiding in encrypted images using block-based adaptive msbs prediction," *Journal of Information Security and Applications*, vol. 69, p. 103297, 2022.

[185] X. Wang, C.-C. Chang, and C.-C. Lin, "Reversible data hiding in encrypted images with block-based adaptive msb encoding," *Information Sciences*, vol. 567, pp. 375–394, 2021.

[186] C. Yu, X. Zhang, G. Li, S. Zhan, and Z. Tang, "Reversible data hiding with adaptive difference recovery for encrypted images," *Information Sciences*, vol. 584, pp. 89–110, 2022.

[187] K.-M. Chen, "High capacity reversible data hiding based on the compression of pixel differences," *Mathematics*, vol. 8, no. 9, p. 1435, 2020.

[188] N. Ostu, “A threshold selection method from gray-level histograms.” *IEEE Trans SMC*, vol. 9, p. 62, 1979.

[189] P.-Y. Pai, C.-C. Chang, Y.-K. Chan, and M.-H. Tsai, “An adaptable threshold detector,” *Information Sciences*, vol. 181, no. 8, pp. 1463–1483, 2011.

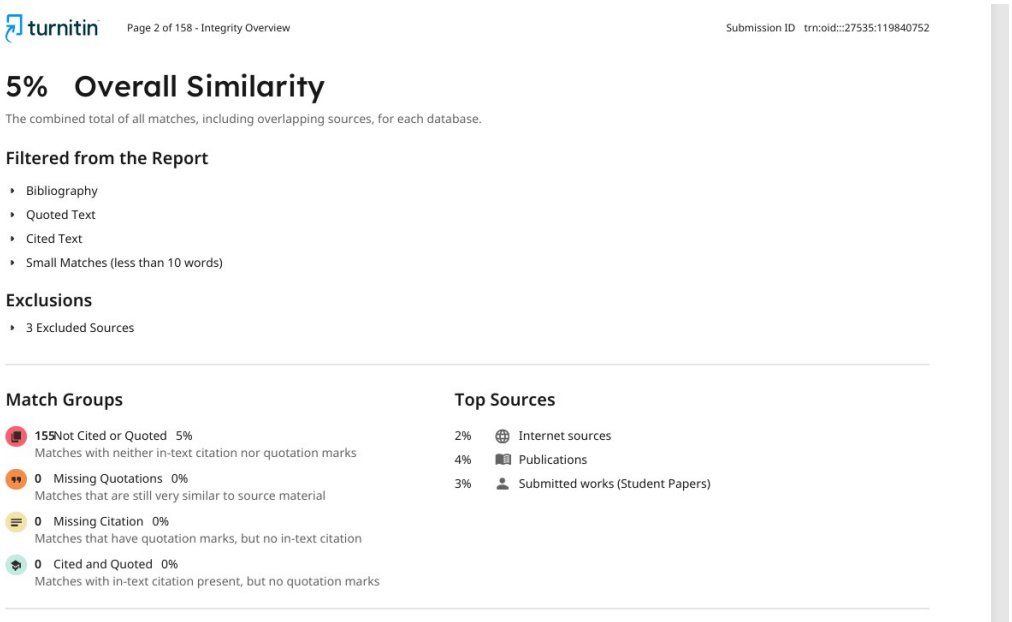
[190] J. Hou, B. Ou, H. Tian, and Z. Qin, “Reversible data hiding based on multiple histograms modification and deep neural networks,” *Signal Processing: Image Communication*, vol. 92, p. 116118, 2021.

[191] K. Dzhanashia and O. Evsutin, “Neural networks-based data hiding in digital images: overview,” *Neurocomputing*, p. 127499, 2024.

PLAGIARISM REPORT

The thesis entitled **"DEVELOPMENT OF DATA HIDING METHODS FOR CONTENT PROTECTION IN MEDICAL DOMAIN"** submitted by **Sonal Gandhi** has been checked for plagiarism using approved software. The similarity index obtained is **5%**, which is within the permissible limit as per the university guidelines.

A copy of the official plagiarism report is appended below.



turnitin Page 2 of 158 - Integrity Overview Submission ID trn:oid:::27535:119840752

5% Overall Similarity

The combined total of all matches, including overlapping sources, for each database.

Filtered from the Report

- Bibliography
- Quoted Text
- Cited Text
- Small Matches (less than 10 words)

Exclusions

- 3 Excluded Sources

Match Groups

- 155 Not Cited or Quoted 5%
  Matches with neither in-text citation nor quotation marks
- 0 Missing Quotations 0%
  Matches that are still very similar to source material
- 0 Missing Citation 0%
  Matches that have quotation marks, but no in-text citation
- 0 Cited and Quoted 0%
  Matches with in-text citation present, but no quotation marks

Top Sources

- 2% Internet sources
- 4% Publications
- 3% Submitted works (Student Papers)

DELHI TECHNOLOGICAL UNIVERSITY

(Formerly Delhi College of Engineering)

Shahbad Daulatpur, Main Bawana Road, Delhi-42

PLAGIARISM VERIFICATION

Title of the Thesis: DEVELOPMENT OF DATA HIDING METHODS FOR CONTENT PROTECTION IN MEDICAL DOMAIN

Total Pages: 150

Name of the Scholar: Sonal Gandhi

Supervisor: Dr.Rajeev Kumar

Department: Computer Science and Engineering

This is to report that the above thesis was scanned for similarity detection. The process and outcome are given below:

Software used: Turnitin **Similarity Index:** 5% **Word Count:** 44,094 Words

Date: 5th November 2025

Candidate's Signature **Signature of Supervisor**

LIST AND PROOF OF PUBLICATIONS

LIST OF PUBLICATIONS:

Journal Publications

1. Sonal Gandhi, Rajeev Kumar; Survey of reversible data hiding: Statistics, current trends, and future outlook, **Computer Standards and Interfaces** , Elsevier, Vol. 112, p. 109017, 202, DOI:10.1016/j.csi.2025.104003. (IF 3.1)

2. Sonal Gandhi and Rajeev Kumar; A high-capacity reversible data hiding with contrast enhancement and brightness preservation for medical images, **Multimedia Tools and Applications**, Springer, Vol. 84, p. 5239–526 DOI:10.1007/s11042-024-18934-1. (IF 3.0)

3. Sonal Gandhi and Rajeev Kumar; UMANeT: A Two-Stage Interpolation-Based Reversible Data Hiding Framework with Attention-Enhanced Prediction. **Journal of Information Security and Applications**, Elsevier. Vol. 94, p. 104217 10.1016/j.jisa.2025.104217 (IF 3.7)

4. Sonal Gandhi and Rajeev Kumar; Reversible Data Hiding for Color Images using a Novel Self-attention based CNN predictor and Error Adjustment, **Signal, Image and Video Processing**, Springer. Vol. 19, p. 1407 10.1007/s11760-025-04991-y (IF 2.1)

Conference Publications

1. Sonal Gandhi and Rajeev Kumar; High Fidelity Reversible Data Hiding For Color Images using CNN Predictor and Reference Error. Fifteenth International Conference on Computing, Communication and Networking Technologies (ICCCNT) 2024, SCOPUS. IEEE. Venue of the conference: IIT Mandi, Himachal Pradesh India. DOI: 10.1109/ICCCNT61001.2024.10724992. (Published November 4, 2024).

2. Sonal Gandhi and Rajeev Kumar; SCAM-Net: Spatial-Channel Attention Multi-Scale Network for Reversible Data Hiding in Encrypted Images. Third International Conference on Women Researchers in Electronics and Computing (WREC 2025), Springer, Springer. Venue of the conference- Dr. B. R. Ambedkar National Institute of Technology Jalandhar, Punjab, India. (Accepted and Presented April 18, 2025).

3. Ankur, Sonal Gandhi, and Rajeev Kumar; Exploring Technological Trend and Collaboration Analysis in Reversible Data Hiding in Encrypted Images. Fortressing Pixels: Information security for images, videos, audio and beyond', IET. (Accepted June 24, 2025).

PUBLICATION PROOF:

Journal Publications

Paper 1: Sonal Gandhi, Rajeev Kumar; Survey of reversible data hiding: Statistics, current trends, and future outlook, **Computer Standards and Interfaces** , Elsevier, Vol. 112, p. 109017, 202, DOI:10.1016/j.csi.2025.104003. ***(IF 3.1)***

Computer Standards & Interfaces 94 (2025) 104003

Contents lists available at ScienceDirect

Computer Standards & Interfaces

journal homepage: www.elsevier.com/locate/csi

Survey of reversible data hiding: Statistics, current trends, and future outlook

Sonal Gandhi, Rajeev Kumar *

Department of Computer Science and Engineering, Delhi Technological University, Delhi, 110042, India

ARTICLE INFO

Keywords:
Reversible data hiding
Prediction
Embedding
Review
Comprehensive
Deep learning
RDHEI
Contrast enhancement

ABSTRACT

In the era of increasing digital media storage and transmission over networks, reversible data hiding (RDH) has evolved as a prominent area of research mitigating information security risk. To study the evolution of research, highlight its achievements over the years, and provide future prospects, this paper presents an extensive review of RDH utilizing the dataset extracted from one of the most popular and exhaustive databases, Web of Science. The study aims to first perform quantitative analysis that includes trend analysis, citation analysis, prominent authors and organizations, and geographical coverage, along with qualitative analysis focusing on key research areas and future prospects within RDH. The study further provides a structured view of sub-technologies within RDH, along with the key contributors and their proposed techniques that have led to the evolution of RDH over the years. Next, we provide a comprehensive review of some of the prominent works in each of the sub-technologies of RDH. Finally, several key research directions, identified based on current research trends and early-stage problems and motivations, are discussed. Overall, the proposed study provides valuable insights into the evolution, key milestones, current state, and future prospects of RDH, serving as a guide for the research community.

Contents

Paper 2: Sonal Gandhi and Rajeev Kumar; A high-capacity reversible data hiding with contrast enhancement and brightness preservation for medical images, **Multimedia Tools and Applications**, Springer, Vol. 112, p. 109017, 202, DOI:10.1007/s11042-024-18934-1. ***(IF 3.0)***

Multimedia Tools and Applications
https://doi.org/10.1007/s11042-024-18934-1

Check for updates

A high-capacity reversible data hiding with contrast enhancement and brightness preservation for medical images

Sonal Gandhi[1] · Rajeev Kumar[1]

Received: 16 May 2023 / Revised: 7 November 2023 / Accepted: 13 March 2024

Abstract
The healthcare industry has witnessed an increase in the use of cloud storage, resulting in a significant demand for safeguarding medical records from potential attackers. In response to this challenge, reversible data hiding (RDH) has emerged as a lifesaver. The RDH ensures the concealment of private and confidential data with minimal loss to the cover image while offering opportunities for image enhancement. This paper introduces a new RDH method with contrast enhancement for medical images. The method aims to provide high embedding capacity (EC) while preserving brightness. To achieve this, the proposed method initially segments cover images into regions of interest (ROI) and non-regions of interest (NROI) and employs different embedding strategies based on the characteristics of each region, thereby enhancing embedding performance. Furthermore, a novel pre-processing technique is introduced for processing ROI pixels. This technique re-organizes and creates empty bins to provide enlarged EC with less distortion by capitalizing on the unique properties of medical images. Additionally, the proposed method embeds secret data in the NROI to further increase the EC. Experimental results show that the proposed method effectively preserves brightness (with 0.99 Relative Mean Brightness Error) while providing higher EC, exceeding the best-known methods by 20%. Furthermore, the method enhances contrast with minimal distortion in the output images.

Keywords Contrast enhancement · Medical images · Reversible data hiding · Embedding capacity · Image segmentation · Pre-processing

1 Introduction

As the healthcare industry progresses towards digitization, cloud storage has become a popular and convenient choice to store and analyze medical data. Electronic Patient Records

✉ Rajeev Kumar
rajeevkumar@dtu.ac.in

Sonal Gandhi
sonalgandhi26@gmail.com

[1] Department of Computer Science and Engineering, Delhi Technological University, Delhi 110042, India

Published online: 02 April 2024

Paper 3: Sonal Gandhi and Rajeev Kumar; UMANeT: A Two-Stage Interpolation-Based Reversible Data Hiding Framework with Attention-Enhanced Prediction. **Journal of Information Security and Applications**, Elsevier. Vol. 94, p. 104217. 10.1016/j.jisa.2025.104217 ***(IF 3.7)***

Contents lists available at ScienceDirect

Journal of Information Security and Applications

journal homepage: www.elsevier.com/locate/jisa

UMANeT: A two-stage interpolation-based reversible data hiding framework with attention-enhanced prediction

Sonal Gandhi, Rajeev Kumar *

Department of Computer Science and Engineering, Delhi Technological University, Delhi, India

ARTICLE INFO

Dataset link: https://gitfront.io/r/sana26/Fgppkvhe87JD/private/

Keywords:
Reversible data hiding
Unet
Interpolation
CNN
Multihead attention
Pixel prediction

ABSTRACT

Interpolation-based reversible data hiding (RDH) techniques have recently attracted significant attention due to their ability to enhance image resolution while ensuring secure data embedding. However, the effectiveness of these methods heavily depends on the quality of the interpolated cover images. Conventional interpolation techniques, typically based on linear models and limited local pixel contexts, often fail to generate high-quality cover images, thereby compromising the visual quality of the resulting stego images and limiting embedding capacity. To address these limitations, this paper introduces a novel hybrid interpolation framework that combines bicubic interpolation with a deep learning-based predictor to construct a high-fidelity two-stage interpolation mechanism. Central to this framework is a newly proposed predictor, termed UMANeT, which leverages a broader contextual region for improved pixel prediction accuracy. By effectively capturing non-linear and long-range dependencies, UMANeT enhances the overall image quality used for data embedding. Experimental results demonstrate that the proposed method not only achieves superior embedding capacity but also generates cover and stego images of significantly higher visual quality compared to existing interpolation-based RDH techniques.

1. Introduction

The rapid expansion of digital communication and cloud storage has significantly increased the volume of sensitive data being transmitted across networks. Critical fields such as medical imaging, satellite communication, and military operations rely heavily on secure data exchange [1]. However, this surge in data transmission has also led to heightened security concerns, particularly the risk of unauthorized access, data breaches, and adversarial threats. These challenges necessitate the adoption of effective techniques to safeguard sensitive information from malicious entities [2].

To address this, various encryption [3] and data hiding [4] techniques have been extensively employed to enhance data security during transmission and storage. Encryption transforms secret data (SD) into an unintelligible format (i.e., ciphertext), ensuring that only authorized users can decipher it. In contrast, data hiding embeds SD within cover media that can be text, images, audio, or video, without altering the overall perceptibility. This technique enhances security by concealing the very existence of SD, ensuring that only authorized recipients can retrieve the hidden data using specialized extraction methods. However, the conventional data-hiding methods often introduce irreversible distortions to the cover media, making them unsuitable for applications where image integrity is crucial. Reversible Data Hiding (RDH) [5–7] has emerged as a viable solution to address this limitation. Unlike traditional approaches, RDH ensures that after extracting the hidden information, the original cover media can be fully restored to its pristine state. This lossless property makes RDH particularly valuable in scenarios where maintaining data authenticity is as important as securing the hidden content.

The existing RDH techniques can be broadly categorized into transform domain [8], compression domain [9], and spatial domain methods [10]. Transform domain techniques embed SD into frequency components obtained after employing transformations such as Discrete Cosine Transform (DCT) [11], Discrete Fourier Transform (DFT) [12], and Discrete Wavelet Transform (DWT) [13,14], etc. In the compression domain, RDH techniques involve the use of various compression [15–17] methods to generate compressed codes from the cover image, thereby reducing its size. The secret data is embedded within these compressed codes in a manner that minimally affects the original content. On the other hand, spatial domain techniques directly modify pixel intensities through methods like Difference Expansion (DE) [18], Histogram Shifting (HS) [1], Pixel Value Ordering (PVO) [19,20], etc. While these methods effectively enable reversible embedding, they often come at the cost of visual distortions or limited embedding capacity.

* Corresponding author.
E-mail addresses: sonalgandhi26@gmail.com (S. Gandhi), rajeevkumar@dtu.ac.in (R. Kumar).

https://doi.org/10.1016/j.jisa.2025.104217

Available online 13 September 2025

Paper 4: Sonal Gandhi and Rajeev Kumar; Reversible Data Hiding for Color Images using a Novel Self-attention based CNN predictor and Error Adjustment, **Signal, Image and Video Processing**, Springer. Vol. 19, p. 1407 10.1007/s11760-025-04991-y ***(IF 2.1)***

Signal, Image and Video Processing (2025) 19:1407
https://doi.org/10.1007/s11760-025-04991-y

ORIGINAL PAPER

Reversible data hiding for color images using a novel self-attention based CNN predictor and error adjustment

Sonal Gandhi[1,2] · Rajeev Kumar[1]

Received: 2 December 2024 / Revised: 17 September 2025 / Accepted: 16 November 2025

Abstract
In prediction-based reversible data hiding (RDH), achieving high prediction accuracy is essential for obtaining high-fidelity marked images with increased embedding capacity. To this end, this paper proposes two novel techniques: a self-attention-based convolutional neural network predictor (SA-CNNP) and an error adjustment strategy for color images. The SA-CNNP effectively captures both local characteristics and global pixel dependencies, resulting in comprehensive coverage and improved prediction accuracy. The error adjustment strategy further enhances accuracy by refining the prediction errors of two color channels using the error distribution of a reference channel, thereby promoting inter-channel consistency. Experimental results demonstrate that these innovations lead to significant improvements. More specifically, the proposed SA-CNNP achieves a sharper prediction error histogram with approximately 8% improvement in MSE over the best known state-of-the-art predictor. Additionally, the error adjustment strategy increases prediction accuracy for 'Kodim09' color image by 58%. Consequently, the proposed RDH approach achieves an average PSNR gain of around 1.2 dB compared to existing methods for color images.

Keywords Reversible data hiding · CNN · Color image · Embedding capacity · Self-Attention

1 Introduction

In today's digital age, Reversible Data Hiding (RDH) has gained more prominence due to its capability of allowing covert communication with the lossless restoration of the original cover [10]. This has led to the development of a variety of RDH works [8, 9] such as Lossless Compression, Prediction-based, Histogram Shifting [6, 7], etc. Among them, prediction-based RDH methods usually provide the highest embedding efficiency because of their ability to efficiently exploit the prevalent image correlation. In these methods, prediction methodology (in short "Predictor") plays a pivotal role as higher prediction accuracy translates to higher embedding efficiency. In literature, various predictors, such as difference predictor (DP) [28], median edge detection [27], gradient adaptive predictor [3], rhombus predictor (RP) [25], XGBoost [17] and Comprehensive RP [16] have been introduced with an aim to improve the prediction accuracy. However, these traditional predictors have performance limitations as they only use a few neighboring pixels as context and struggle to establish non-linear relationships between the pixels.

1.1 Deep learning based predictors

To address the limitations of aforementioned traditional predictors, Hu et al. [13] introduced a convolutional neural network-based predictor (CNNP) that leverages the multi-receptive field and complex computation capabilities of CNN. The CNNP first partitions the original image into two independent sets as shown in Fig. 1 and then predicts one from the other, thus exploiting a larger set of context pixels along with the nonlinear relationship among them.

To enhance performance, Hu et al. [14] modified the image-division strategy used in CNNP training and implemented prediction error ordering (PEO) for embedding,

✉ Rajeev Kumar
rajeevkumar@dtu.ac.in

Sonal Gandhi
sonalgandhi_2k21phdco02@dtu.ac.in

[1] Department of Computer Science and Engineering, Delhi Technological University, Delhi, India

[2] Department of Computer Science and Engineering, G. L. Bajaj Institute of Technology and Management, Greater Noida, India

Published online: 04 December 2025

Conference Publications

Paper 1: Sonal Gandhi and Rajeev Kumar; High Fidelity Reversible Data Hiding For Color Images using CNN Predictor and Reference Error. Fifteenth International Conference on Computing, Communication and Networking Technologies (ICCCNT) 2024, SCOPUS. IEEE. Venue of the conference: IIT Mandi, Himachal Pradesh India. DOI: 10.1109/ICCCNT61001.2024.10724992. **(Published November 4, 2024)**

IEEE PRO COMM Professional Communication Society | IEEE ELECTRONICS PACKAGING SOCIETY | IAS IEEE INDUSTRY APPLICATIONS SOCIETY | IEEE

THE FIFTEENTH INTERNATIONAL IEEE CONFERENCE ON COMPUTING, COMMUNICATION AND NETWORKING TECHNOLOGIES

June 24th – 28th, 2024, IIT – Mandi, Himachal Pradesh, India.

This is to certify that Prof./Dr./Mr./Ms. **Sonal Gandhi**

has presented a paper on

Paper ID-268: High Fidelity Reversible Data Hiding For Color Images using CNN Predictor and Reference Error

in the Fifteenth International Conference on Computing, Communication and Networking Technologies (ICCCNT), held at IIT-Mandi, in association with IEEE Electronics Packaging Society, during June 24th - 28th, 2024.

Conference Chair / Co-Chair

2024 15th International Conference on Computing Communication and Networking Technologies (ICCCNT) | 979-8-3503-7024-9/24/$31.00 ©2024 IEEE | DOI: 10.1109/ICCCNT61001.2024.10724992

High Fidelity Reversible Data Hiding For Color Images using CNN Predictor and Reference Error

Sonal Gandhi
Dept. of Computer Science and Engg.
Delhi Technological University
Delhi, India
sonalgandhi26@gmail.com

Rajeev Kumar
Dept. of Computer Science and Engg.
Delhi Technological University
Delhi, India
rajeevkumar@dtu.ac.in

***Abstract*—In the field of reversible data hiding, both prediction and embedding techniques play an important role in defining the overall effectiveness of a data hiding scheme. A lot of efforts have been invested by the research community to improve the embedding performance. However, the prediction area is still evolving, wherein the Convolutional neural network (CNN), with its capability to consider long-range dependencies, has shown a new direction to the researchers. This paper proposes an error adjustment strategy with a CNN-based predictor for color images. The proposed strategy utilizes the inter-channel correlation of the three channels to enhance the prediction performance. Experimental results show that the prediction accuracy achieved by the proposed method is significantly higher than the state-of-the-art methods. Furthermore, using the proposed predictor, the Peak signal-to-noise ratio achieved for the embedded images for most of the standard test images is better than the known state-of-the-art methods.**

***Index Terms*—Reversible data hiding, Color image, Channel correlation, CNN, Reference error.**

I. Introduction

With the increase of internet technology in recent years, the importance of content protection and information security has increased. Data hiding entails an important role in the realm of information security, especially in content authentication and copyright protection [1]. It involves concealing confidential/secret information within a cover medium in such a manner that it remains imperceptible to the human eye. However, conventional data-hiding techniques [2] often result in permanent distortions to the original cover when secret data are extracted. This poses a significant challenge, especially in the military and medical fields, where along with the stored sensitive information original cover image is equally important. To address this concern, reversible data hiding (RDH) [3] has evolved as a savior that allows the reconstruction of the original cover without any distortion.

Recent breakthroughs in RDH have led to promising techniques such as difference expansion (DE) [4], histogram shifting (HS) [5], and prediction-error expansion (PEE) [6] for minimizing distortion. Other techniques rely on designing accurate predictors to predict pixel values, such as the difference predictor (DP), median edge direction predictor (MEDP) [7], gradient adaptive predictor (GAP) [8], and bilinear interpolation predictor (BIP) [9]. RDH techniques have shown significant improvements in grayscale images. However, not much attention has been paid to color images.

With the increase in the trend of color images in our daily lives, it is crucial to explore RDH methods specifically tailored to color images. Most of the existing methods designed for color images concentrate on designing an efficient predictor to leverage inter-channel correlation. However, because of the limited performance of the predictors, these models fail to utilize channel correlation in the overall process. To address this concern, this paper takes into account inter-channel correlations and proposes a CNN-based predictor to accurately predict pixel values.

The rest of this paper is organized as follows: Section II provides a comprehensive review of the several existing RDH schemes for color images. Section III describes the execution of the proposed method and outlines the detailed embedding procedure. Additionally, in Section IV, we present the experimental analysis. Finally, we conclude the paper in Section V.

II. Related Work

In this section, we will discuss the existing and related RDH schemes for color images [10]–[17]. Nowadays, RDH for color images has become a new direction for researchers with two primary aspects. One aspect involves leveraging inter-channel correlation to enhance prediction accuracy [18] whereas, the second aspect aims to improve the embedding performance [19]–[23] by utilizing the inter-channel correlation.

Most conventional prediction methods focus on improving prediction accuracy by utilizing pixel similarities. However, these approaches fall short of the target due to their local and linear nature, making them inadequate for complex distributed images. Hence, immediate action is required to investigate global and non-linear predictors to improve accuracy.

Deep learning techniques like CNN have already accomplished new heights in the area of image classification and object recognition. These achievements have also attracted researchers to leverage CNN in the RDH domain. Various CNN-based RDH methods for grayscale images have been introduced by the research community. In [24], Luo et al. proposed a CNN-based RDH approach for stereo images. Each

15th ICCCNT IEEE Conference,
June 24-28, 2024,
IIT - Mandi, Kamand, India

Paper 2: Sonal Gandhi and Rajeev Kumar; SCAM-Net: Spatial-Channel Attention Multi-Scale Network for Reversible Data Hiding in Encrypted Images. Third International Conference on Women Researchers in Electronics and Computing (WREC 2025), Springer, Springer. Venue of the conference- Dr. B. R. Ambedkar National Institute of Technology Jalandhar, Punjab, India. ***(Accepted and Presented April 18, 2025)***

Certificate No. NITJ/WREC_2025/OL_021

3rd International Conference

Women Researchers in Electronics and Computing

WREC'25

Theme: Sustainable Development Goals

(18-19 April, 2025)

Springer

Certificate of Appreciation

This is to certify that paper titled

SCAM-Net: Spatial-Channel Attention Multi-Scale Network for Reversible Data Hiding in Encrypted Images

authored by

Sonal Gandhi,Rajeev Kumar

has been successfully presented by

Sonal Gandhi from Delhi Technological University(DTU)

during 3rd International Conference "Women Researchers in Electronics and Computing" on Sustainable Development Goals held on 18-19 April 2025 organized by **Dr B R Ambedkar National Institute of Technology Jalandhar, Punjab, India.**

Prof. Mamta Khosla
Executive General Chair

Dr. Indu Saini
Executive General Chair

AUTHOR'S BIOGRAPHY

Sonal Gandhi earned her B.Tech degree in Information Technology in 2011, and M.Tech degree in Computer Science in 2014, both from Kurukshetra University, India. Following her postgraduate studies, she gained valuable academic experience, serving as a faculty member in the field of engineering and technology in prestigious institutes.

She is currently pursuing her Doctor of Philosophy (Ph.D.) in Computer Science and Engineering at Delhi Technological University (DTU), Delhi, India. Her ongoing research focuses on Machine Learning, Information Security, and Reversible Data Hiding with applications in secure image processing and medical data protection. She has made significant contributions to the field through multiple research publications in reputable journals and conferences.