

DEVELOPMENT OF HIGH-CAPACITY REVERSIBLE DATA HIDING METHODS FOR ENCRYPTED IMAGES

A Thesis Submitted

in Partial Fulfillment of the Requirements for the

Degree Of

MASTER OF TECHNOLOGY

BY RESEARCH

in

Computer Science and Engineering

by

AJIT KUMAR YADAV

Roll No.: 2K24/RCO/05

Under the supervision of

Dr. RAJEEV KUMAR

Assistant Professor



Department Of Computer Science and Engineering

DELHI TECHNOLOGICAL UNIVERSITY

(Formerly Delhi College of Engineering)

Bawana Road, Delhi – 110042

MAY 2026

CANDIDATE'S DECLARATION

I, **Ajit Kumar Yadav**, Roll No. 2K24/RCO/05, hereby declare that the thesis entitled “**Development of High-Capacity Reversible Data Hiding Methods for Encrypted Images**” submitted to Delhi Technological University, Delhi, in partial fulfillment of the requirements for the award of the degree of Master of Technology, is an original work carried out by me under the supervision of Dr. Rajeev Kumar.

I further declare that this thesis has not been submitted, either in full or in part, for the award of any degree, diploma, associateship, fellowship, or any other academic distinction at this or any other institution. The research work reported in the thesis has been prepared with due acknowledgement of the sources used.

Place: Delhi

Date: _____

Ajit Kumar Yadav

This is to certify that the student has incorporated all the corrections suggested by the examiners in the thesis and the statement made by the candidate is correct to the best of our knowledge

Signature of Supervisor

Signature of External Examiner

CERTIFICATE

DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING

DELHI TECHNOLOGICAL UNIVERSITY

(Formerly Delhi College of Engineering)

Bawana Road, Delhi – 110042

This is to certify that the thesis entitled “**Development of High-Capacity Reversible Data Hiding Methods for Encrypted Images**” submitted by Ajit Kumar Yadav, Roll No. 2K24/RCO/05, Department of Computer Science and Engineering, Delhi Technological University, Delhi, in partial fulfillment of the requirements for the award of the degree of Master of Technology, is a record of research work carried out under my supervision.

To the best of my knowledge, the work presented in this thesis has not been submitted previously for any degree or diploma at this university or any other institution.

Place:

Dr. Rajeev Kumar

Date:

Supervisor

ACKNOWLEDGEMENT

I express my sincere gratitude to my supervisor, Dr. Rajeev Kumar, for his valuable guidance, technical suggestions, and continuous encouragement throughout this research work. His support helped me refine the direction of the thesis and complete the work in a disciplined academic manner.

I am also thankful to the Department of Computer Science and Engineering, Delhi Technological University, for providing the academic environment, resources, and institutional support required for this dissertation. I gratefully acknowledge the assistance of the faculty members, laboratory staff, and my fellow researchers for their constructive discussions and cooperation.

Finally, I thank my family and friends for their constant encouragement, patience, and moral support during the completion of this thesis.

Place: Delhi

Ajit Kumar Yadav

Date: _____

ABSTRACT

Security issues in image handling have gained significant importance because of the increasing use of outsourced storage facilities for handling images through operations such as storage, transmission, indexing, and manipulation. Confidentiality and data association are two aspects that should be taken into account when dealing with images in such an environment since one aspect of confidentiality is that the information embedded within the images should not be disclosed, and at the same time, auxiliary information, such as authentication, ownership, indexing, or medical data, should also be associated with the images. In view of these requirements, RDH-EI technology has been developed that enables the embedding of auxiliary information in encrypted images with the ability to reconstruct the original image without any error. This work considers two different RDH-EI approaches. The first method is related to room reservation prior to encryption in order to keep the embedding space intact, whereas the second approach relates to embedding a higher amount of information after encryption using block index modulation.

The first contribution develops an Adaptive Weighted Predictor for a checkerboard-based RDH-EI framework. The image is separated into dot and cross pixel sets. Dot pixels are predicted first with a Rhombus Mean Predictor, and the error map obtained from this stage is reused as a local reliability indicator for cross-pixel prediction. Instead of treating all contextual predictors equally, the proposed predictor assigns lower influence to neighborhoods with larger first-stage prediction errors and higher influence to more reliable neighborhoods. A damping parameter is incorporated to avoid excessive compensation in smooth regions. The reported evaluation demonstrates improvement in MAE, MSE, and prediction-error variance with respect to the RMP plus AMP baseline.

The second contribution introduces an index-modulated high-capacity RDH-EI method for distortion-free data embedding in encrypted images. The method employs dual-stage encryption, combining block permutation with intra-block pixel scrambling to protect visual content while enabling structured encrypted-domain embedding. Through adaptive block regulation, encrypted blocks are arranged in a

deterministic order, and secret bits are represented by selecting valid permutation indices rather than by directly modifying encrypted pixel values. This index-based embedding strategy preserves reversibility, allowing complete recovery of both the embedded auxiliary information and the original image. Experimental results demonstrate stable high-capacity embedding performance across standard benchmark images, while perfect reconstruction is consistently achieved for the recovered images.

Together, the two contributions investigate separate but related bottlenecks in RDH-EI design. The first improves the compactness of prediction errors in an RRBE-oriented workflow, whereas the second increases encrypted-domain embedding capability within a VRAE-style setting.

LIST OF PUBLICATIONS

Conference Publications

- [1] **A. K. Yadav** and R. Kumar, “Adaptive weighted predictor for checkerboard-based reversible data hiding in encrypted images,” in Proc. Fourth Int. Conf. Secure Cyber Computing and Communications (ICSCCC), Jalandhar, Punjab, India, 2026. (Accepted)
- [2] **A. K. Yadav** and R. Kumar, “Index-modulated high-capacity reversible data hiding in encrypted images,” in Proc. Fourth Int. Conf. Secure Cyber Computing and Communications (ICSCCC), Jalandhar, Punjab, India, 2026. (Accepted)
- [3] M. Prasad, **A. K. Yadav**, R. Kumar, and V. Fedoseev, “ShortcutProbe: Identifying backdoor samples via internal stability analysis,” in Proc. 2026 14th Int. Symp. Digital Forensics and Security (ISDFS), 2026, pp. 1–6.

Book Chapter

- [1] Ankur, S. Gandhi, **A. K. Yadav**, and R. Kumar, “Reversible data hiding in encrypted images: Qualitative deep-dive,” in *Crypto-Forensics: A Useful Direction to Digital Evidence Recovery*. Cham, Switzerland: Springer, 2026. (Accepted)

Patents

- [1] **A. K. Yadav** and R. Kumar, “A system and method for prioritized data delivery,” Indian Patent Application 202511070240, filed on July 23, 2025. (Published)
- [2] **A. K. Yadav**, A. Kumar, D. Joshi and R. Kumar, “System and method for securely charging electric vehicle,” Indian Patent Application 202511054094, Patent No. 588448, granted May 6, 2026.

TABLE OF CONTENTS

CANDIDATE’S DECLARATION	ii
CERTIFICATE	iii
ACKNOWLEDGEMENT	iv
ABSTRACT.....	v
LIST OF PUBLICATIONS	vii
Conference Publications.....	vii
Book Chapter	vii
Patents	vii
TABLE OF CONTENTS.....	viii
LIST OF FIGURES	x
LIST OF TABLES	xi
LIST OF ABBREVIATIONS	xii
CHAPTER 1 INTRODUCTION	1
1.1 Background and Motivation.....	1
1.2 Problem Domain Overview	3
1.3 Existing Challenges.....	4
1.4 Problem Statement	5
1.5 Research Objectives	5
1.6 Scope of the Research	5
1.7 Applications	6
1.8 Research Contributions	6
1.9 Thesis Organization	7
CHAPTER 2 LITERATURE REVIEW	8
2.1 Evolution of Reversible Data Hiding.....	8
2.2 RDH-EI Frameworks: VRAE and RRBE	9
2.3 Prediction-based RRBE Methods	10
2.4 Encrypted-domain VRAE and Block-structure Methods	11
2.5 Deep Learning and Attention-based Prediction	12
2.6 Comparative Analysis of Existing Methods	13
2.7 Research Gaps.....	14
2.8 Summary	14

CHAPTER 3 PROPOSED WORK – PART I: ADAPTIVE WEIGHTED PREDICTOR FOR CHECKERBOARD-BASED RDH-EI	15
3.1 Introduction	15
3.2 Related Background Specific to the Method	16
3.3 System Model	17
3.4 Proposed Methodology	17
3.4.1 Image Partitioning	17
3.4.2 Stage 1: Dot Prediction and Error Map Generation	18
3.4.3 Stage 2: Adaptive Weighted Cross-set Prediction	18
3.5 Mathematical Formulation	20
3.6 Algorithm Explanation	21
3.7 Data Embedding and Image Recovery	22
3.8 Experimental Setup	24
3.9 Performance Evaluation and Discussion	26
3.10 Applicability, and Limitations	27
3.11 Chapter Summary	28
CHAPTER 4 PROPOSED WORK – PART II: INDEX-MODULATED HIGH-CAPACITY RDH-EI	29
4.1 Introduction	29
4.2 Related Background Specific to the Method	30
4.3 System Model	31
4.4 Proposed Methodology	31
4.4.1 Two-stage Encryption	31
4.4.2 Block Regulation	32
4.4.3 Permutation-index Embedding	33
4.5 Data Extraction and Image Recovery	34
4.7 Algorithm Explanation	35
4.8 Experimental Setup	37
4.9 Performance Evaluation and Discussion	38
4.10 Applicability, and Limitations	39
4.11 Chapter Summary	40
CHAPTER 5 CONCLUSION AND FUTURE SCOPE	41
5.1 CONCLUSION	41
5.2 FUTURE SCOPE	42
REFERENCES	43

LIST OF FIGURES

Figure 1:: General Architecture of Reversible Data Hiding in Encrypted Images	2
Figure 2:Block diagram of different RDHEI categories	9
Figure 3:Overall Framework of the Proposed Adaptive Weighted Predictor Method	15
Figure 4:Checkerboard Partitioning of Dot and Cross Pixel Sets.....	17
Figure 5:Prediction Context for Cross-set Pixel Using Error-guided Weights.....	19
Figure 6:Complete Workflow of the Proposed AWP-based RDH-EI Method.....	22
Figure 7:Standard 512×512 Grayscale Test Images Used for Evaluation	25
Figure 8:MAE, MSE, and Variance Comparison between Baseline and Proposed Method	26
Figure 9:: Block Diagram of the Proposed Index-Modulated RDH-EI Method.....	30
Figure 10:Illustration of Two-stage Encryption Process	31
Figure 11:Data Embedding Workflow Using Permutation-index Modulation.....	34
Figure 12:Data Extraction and Image Recovery Workflow	35
Figure 13:Complete Proposed Index-Modulated RDH-EI Workflow	36
Figure 14:Test Images Used for Evaluation - Baboon, Barbara, Boat, and Lena.....	37
Figure 15:Embedding Rate Comparison on Standard Test Images	39

LIST OF TABLES

Table 1:Comparative view of VRAE and RRBE frameworks.....	10
Table 2:Comparative analysis of representative RDH-EI approaches.....	13
Table 3:Proposed adaptive weighted predictor-based room-reservation workflow ..	21
Table 4:Prediction performance comparison between RMP+AMP baseline and proposed method	25
Table 5:Embedding rate for different test images under different values of d	37
Table 6:Performance comparison for different test images	38
Table 7:Dataset-level performance comparison of the proposed method.....	38

LIST OF ABBREVIATIONS

AMP	Adaptive Mean Predictor
AWP	Adaptive Weighted Predictor
BOSSBase	Break Our Steganographic System Base Dataset
BOWS-2	Break Our Watermarking System 2 Dataset
CNN	Convolutional Neural Network
DH	Data Hiding
EC	Embedding Capacity
ER	Embedding Rate
GAP	Gradient-Adjusted Predictor
LSB	Least Significant Bit
MAE	Mean Absolute Error
MED	Median Edge Detector
MSE	Mean Squared Error
MSB	Most Significant Bit
PEE	Prediction Error Expansion
PSNR	Peak Signal-to-Noise Ratio
RDH	Reversible Data Hiding
RDH-EI	Reversible Data Hiding in Encrypted Images

RMP	Rhombus Mean Predictor
RRBE	Reserving Room Before Encryption
SSIM	Structural Similarity Index
VRAE	Vacating Room After Encryption

CHAPTER 1

INTRODUCTION

1.1 Background and Motivation

Digital images are a primary form of information exchange in modern computing systems. They are generated and used in healthcare, surveillance, e-governance, education, remote sensing, intelligent transportation, social networking, and industrial inspection. In many of these applications, the image is not only a visual record but also a carrier of contextual information. Information such as patient identifiers, access labels, authentication tags, ownership marks, case numbers, or audit records may have to remain bound to the image during storage and transmission [1]. If these details are stored as independent metadata, they can be deleted, modified, separated from the image, or lost during conversion. Data hiding provides a direct way to bind auxiliary information with the cover image and is therefore an important area within secure multimedia processing [2].

Conventional data hiding algorithms modify the cover image to include the embedded data. Even though the distortion caused due to such an operation may be invisible to the naked eye, the original pixel values cannot be retrieved accurately. Such a disadvantage would pose problems in applications such as diagnostic medical imaging, forensics, military, and scientific measurements, since a slight variation in the pixel values would be sufficient to influence the final results. To circumvent such problems, Reversible Data Hiding (RDH) was devised for the purpose of extracting the payload, as well as reconstructing the original image [1]-[4].

The problem with ordinary RDH is the requirement for plaintext images, which cannot be guaranteed in some cases, particularly when outsourcing image data to a cloud server, or using a third-party data managing entity. If the image is outsourced to a cloud server, the data hider will not have access to its visual contents due to security concerns. Therefore, RDH-EI involves encrypting, reversible embedding, and recovering the image in lossless manner with extra data embedded into it.

The performance of RDH-EI depends strongly on the way embeddable space is produced. In the Reserving Room Before Encryption (RRBE) framework, the content owner analyzes the original image and creates room before encryption. This preprocessing can exploit spatial correlation, prediction errors, and bit-plane redundancy, and therefore RRBE methods often achieve comparatively high capacity. In the Vacating Room After Encryption (VRAE) framework, the image is encrypted first and the data hider works in the encrypted domain. VRAE is useful when the data hider must not access plaintext pixels, but it is more challenging because encryption suppresses ordinary image redundancy [7], [22].

This thesis brings together two RDH-EI contributions that address these two design directions. The first contribution is an Adaptive Weighted Predictor (AWP) for checkerboard-based RRBE. It improves second-stage prediction by using first-stage prediction errors as reliability information. The second contribution is an index-modulated VRAE-style method in which secret bits are represented by permutation states of regulated encrypted blocks. These two approaches examine different mechanisms for improving embedding efficiency while maintaining image confidentiality and exact recovery.

Figure 1. presents the general RDH-EI processing architecture considered in this thesis. The model separates the roles of content owner, data hider, and receiver so that privacy-preserving outsourced image management can be discussed clearly.

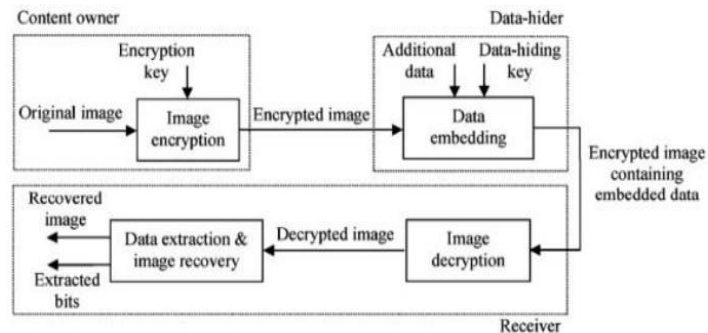


Figure 1:: General Architecture of Reversible Data Hiding in Encrypted Images

1.2 Problem Domain Overview

An RDH-EI system is generally described using three functional entities. The content owner possesses the original image and applies encryption using an image-encryption key. The data hider inserts auxiliary data and may use a separate data-hiding key. The recipient is given one or both keys according to the required access privilege. With the data-hiding key, the receiver extracts the embedded payload. With the image decryption key and the required recovery information, the receiver reconstructs the original image. The separation of these privileges is important because it allows a third party to manage or annotate encrypted images without learning their visual content.

Two processing paradigms dominate this area. In RRBE, the owner reserves room while the image is still available in plaintext form. Prediction-error coding, MSB compression, histogram modification, and bit-plane analysis are commonly used to create embeddable space. In VRAE, encryption precedes embedding. Since the encrypted image appears random, the embedding algorithm must rely on preserved structure, block organization, or carefully designed reversible operations instead of ordinary spatial correlation.

The proposed work considers both major RDH-EI paradigms. In the RRBE setting, room is prepared before encryption by exploiting plaintext image correlation. In the VRAE setting, the image is encrypted first and the embedding operation is performed later in the encrypted domain. The first proposed method improves checkerboard prediction by using dot-set errors as reliability information for cross-set prediction. The second method uses block regulation and permutation-index mapping so that secret bits can be represented by valid encrypted-block states rather than by direct plaintext prediction.

Accordingly, the thesis treats the two methods as complementary responses to a common problem: increasing useful embedding capability while preserving confidentiality and perfect reversibility

1.3 Existing Challenges

Designing an RDH-EI method turns out to be quite challenging, because a bunch of requirements have to match each other, more or less, at the same time. Like, the technique should still offer enough hiding capacity, keep the image confidentiality, allow exact reconstruction, manage side information, and at the end still be computationally practical. In reality these requirements clash with each other. For instance, if you push the payload higher, you usually need more reserved room, but making that room available can also introduce additional auxiliary data. Likewise, if you strengthen encryption to boost privacy, you often lose the spatial redundancy that many reversible embedding methods usually rely on.

In RRBE- based schemes, prediction accuracy matters a lot, since the distribution of prediction errors basically decides how efficiently you can carve out that reserved room. If most residuals stay near zero, then the error map can be stored in a more compact way. Classical predictors such as MED, GAP, and RMP are appealing because they're simple and fully deterministic, but they can stumble near edges and in textured regions. Adaptive predictors do better with directionality, however they may still assume all nearby samples are equally trustworthy even when some neighboring values are linked to larger local errors [8]–[11].

In VRAE-based schemes, the difficulty is different. After encryption, pixels no longer exhibit the visible smoothness or local correlation found in the original image. Direct modification of encrypted pixels can create ambiguity during recovery, especially when the receiver must reconstruct the original image without loss. Block-regulation and permutation-based methods attempt to introduce a reversible structure inside the encrypted domain, but the embedding rule must not leak image content or create irreversible changes [5], [6], [22].

Side information also affects practical capacity. Recovery may require label maps, overflow records, block states, predictor indicators, location maps, or compressed auxiliary streams. If these data become too large, the net embedding gain decreases.

For this reason, an effective RDH-EI method should reuse information already produced during processing and should avoid unnecessary side-information growth.

1.4 Problem Statement

The problem addressed in this thesis is the development and analysis of reversible data hiding methods for encrypted images that improve embedding capacity while preserving confidentiality, deterministic extraction, and exact recovery.

1.5 Research Objectives

- To examine the principles of RDH, RDH-EI that includes both RRBE, and VRAE with respect to privacy, capacity, reversibility, and computational feasibility.
- To study the influence of prediction accuracy on RRBE-based room reservation and identify the limitations of fixed-context or uniformly weighted predictors.
- To design a two-stage adaptive weighted predictor in the RDH-EI domain that exploits first-stage errors for enhanced accuracy.
- To design a new high capacity VRAE based RDH-EI method.
- To compare the proposed methods with representative RDH-EI approaches using prediction accuracy, embedding rate, reversibility, security, complexity, and implementation practicality.

1.6 Scope of the Research

The work reported in this thesis is limited to reversible data hiding in encrypted grayscale images. The first proposed method is studied in an RRBE workflow, where prediction and room reservation are performed before encryption by the content owner. The second proposed method is studied in a VRAE-style workflow, where embedding is carried out after encryption by using block states and permutation indices. Both methods are discussed through their system models, mathematical operations, algorithmic steps, and reported experimental results. No additional

experimental claims are introduced beyond the values available in the corresponding conference-paper work.

1.7 Applications

RDH-EI is useful in applications where image privacy and attached information must be maintained together. In medical image storage, diagnostic images may be encrypted before being uploaded, while patient identifiers or authentication codes can still be inserted and later recovered. In forensic and legal applications, exact image reconstruction is important because even a minor pixel change may affect the reliability of evidence. Similar requirements also appear in surveillance, defense image handling, digital archiving, and cloud-based image management, where data must remain confidential but still traceable.

Cloud storage, digital archiving, and Internet of Things deployments also benefit from RDH-EI. An edge device can encrypt an image before upload, and a server can insert indexing or management information without accessing the plaintext image. The first proposed method is suitable when the content owner can reserve room before encryption and wants stronger prediction compactness. The second proposed method is suitable when embedding must be performed on encrypted data and plaintext access is not available to the data hider.

1.8 Research Contributions

- The thesis contains two primary contributions derived from the accepted conference-paper work.
- **Contribution 1:** An Adaptive Weighted Predictor is developed for checkerboard-based RDH-EI. The dot set is predicted using RMP, and the resulting error map is reused to assign reliability-based weights during cross-set prediction. The weighting reduces the influence of locally unreliable predictors and improves the compactness of prediction errors.
- **Contribution 2:** An index-modulated RDH-EI method is developed for encrypted-domain embedding. The method applies two-stage encryption,

block regulation, and permutation-index mapping so that secret bits are carried by valid encrypted-block states.

1.9 Thesis Organization

Chapter 2 reviews RDH and RDH-EI literature, including RRBE, VRAE, prediction-based reservation, block-regulation methods, and learning-based prediction. Chapter 3 presents the Adaptive Weighted Predictor for checkerboard-based RDH-EI. Chapter 4 presents the index-modulated high-capacity RDH-EI method. Chapter 5 concludes the thesis and identifies future research directions.

CHAPTER 2

LITERATURE REVIEW

2.1 Evolution of Reversible Data Hiding

Reversible data hiding developed from the requirement that a cover image should carry auxiliary data and still be reconstructed without loss. In early RDH research, Tian's difference expansion method showed that differences between neighboring pixels could be expanded to insert payload bits while retaining an inverse mapping for recovery [1]. Later studies improved the concept by using prediction-error expansion, histogram shifting, pixel-value ordering, and related statistical methods [2], [3]. These approaches make use of spatial correlation in natural images, where neighboring pixels are often similar and prediction errors are frequently small.

Plaintext RDH methods have the advantage that the embedding algorithm can examine the cover image directly. Local statistics, smoothness, edge direction, histogram peaks, overflow conditions, and predictor behavior can be measured before data insertion. However, this also means that the data hider or storage provider may see the original image. As image outsourcing became common, the need for reversible hiding under encryption led to RDH-EI [5], [6].

The transition from RDH to RDH-EI changes the design problem. A method must now support three functions together: preserving confidentiality, carrying the hidden message, and recovering the cover image exactly. If encryption is applied in a conventional strong form, the encrypted image loses the visible redundancy on which standard RDH depends. Consequently, RDH-EI techniques either create embeddable space before encryption or deliberately preserve a reversible structure that can be used after encryption.

Figure 2.1 summarizes the conceptual movement from plaintext RDH to the two major RDH-EI frameworks. This positioning is important for understanding why the two proposed contributions in this thesis are designed under different operating assumptions.

2.2 RDH-EI Frameworks: VRAE and RRBE

RDH-EI methods are commonly grouped according to the time at which embedding room is created depicted in fig 2. In Vacating Room After Encryption (VRAE), the content owner encrypts the image first. The data hider then inserts the payload into the encrypted image. This model is attractive because the embedding party does not require plaintext access. Its difficulty is that encryption makes the image appear random, and therefore ordinary spatial redundancy is not readily available for reversible embedding [5], [6], [22].

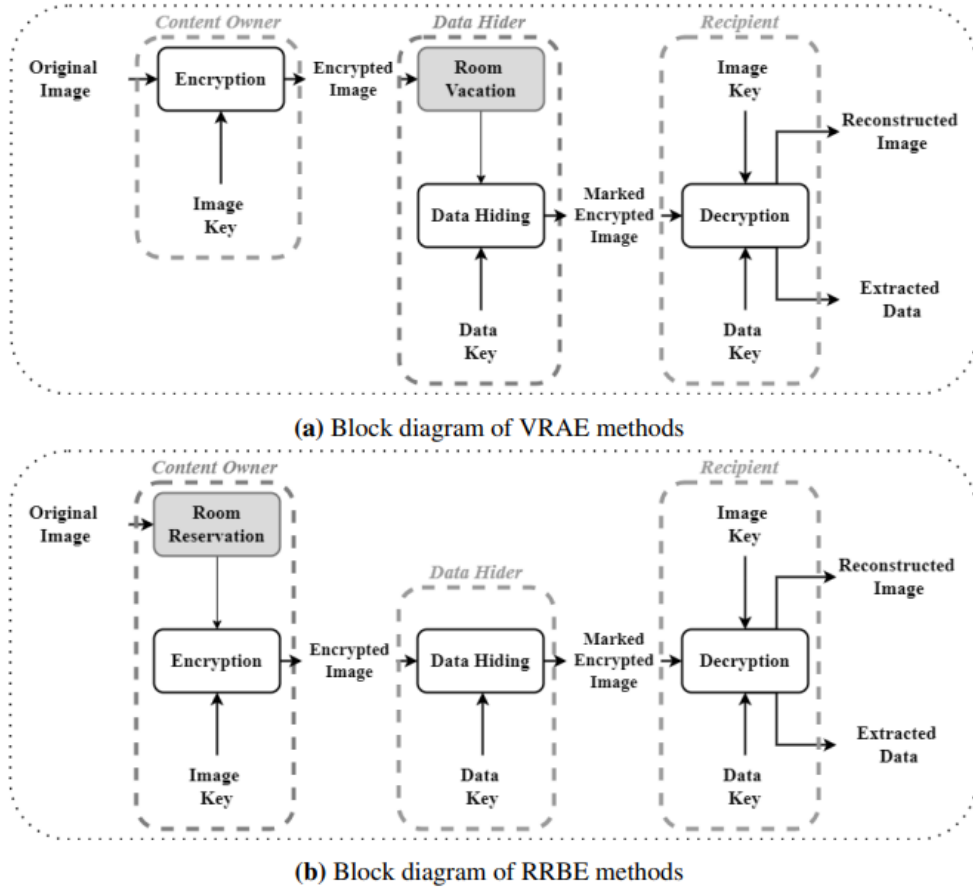


Figure 2: Block diagram of different RDHEI categories

In Reserving Room Before Encryption (RRBE), the content owner performs preprocessing before encryption. Ma et al. established this direction by showing that the owner can create embeddable room in the plaintext image and then encrypt the processed image for later data insertion [7]. Since prediction and compression can be

performed while the image remains in the owner’s secure environment, RRBE methods generally have a stronger opportunity to achieve high capacity.

Both paradigms are important in practice. RRBE is effective when the content owner can actively prepare the image and wants higher embedding efficiency. VRAE is appropriate when encrypted content has already been outsourced or when the data hider must remain completely separated from plaintext image content. The present thesis includes one contribution from each side: an RRBE-oriented prediction method and a VRAE-oriented index-modulation method.

Table 1: Comparative view of VRAE and RRBE frameworks

Aspect	VRAE	RRBE
Room creation	After encryption	Before encryption
Use of plaintext redundancy	Limited or specially preserved	Directly available during preprocessing
Typical capacity	Often lower unless special structure is used	Generally higher due to prediction/compression
Main challenge	Embedding in random-looking encrypted data	Efficient preprocessing and side-information control
Representative relevance in this thesis	Index-modulated encrypted-domain embedding	Adaptive weighted prediction for room reservation

2.3 Prediction-based RRBE Methods

Prediction-based RRBE methods are built on the correlation of nearby pixels. If a target pixel can be estimated accurately from its neighborhood, the residual between the actual value and predicted value becomes small. A compact prediction-error distribution is valuable because it can be encoded more efficiently, leaving more capacity for payload bits [7], [8].

Several deterministic predictors have been used in RDH and RDH-EI systems. The Median Edge Detector, Gradient-Adjusted Predictor, local mean predictors, and Rhombus Mean Predictor are popular because they are simple, fast, and reproducible. However, fixed contexts may behave poorly near edges, strong texture, or irregular intensity transitions [8], [14], [15]. Adaptive Mean Prediction improves the situation by selecting direction-aware estimates from local context [10].

Despite its usefulness, AMP does not directly measure whether neighboring pixels are reliable. A neighboring value may itself have been reconstructed with a high error, and such a value should not contribute as strongly as a stable neighbor. This observation motivates the first proposed method. The method uses the error map already produced during dot-set prediction and converts it into reliability weights for the later cross-set prediction stage.

Multi-predictor methods provide another way of improving estimation accuracy by testing several predictors and storing the selected predictor index. Such methods can reduce local error, but the index map becomes auxiliary information that must be protected and transmitted. If the number of blocks or contexts is large, this overhead can reduce the net embedding gain [18], [19]. A reliability-aware predictor that avoids explicit predictor-index storage is therefore attractive for reversible systems.

2.4 Encrypted-domain VRAE and Block-structure Methods

VRAE methods run under a stricter condition than RRBE methods, since in VRAE the data hider gets the encrypted image not the actual original one. So it looks great for outsourced storage, yet it also cuts off the direct access to pixel correlation, at least in practice. Early RDH-EI approaches already showed that reversible embedding plus separable extraction can work on encrypted images, but the capacity was often kinda limited, because encrypted pixels don't really keep the same smoothness, or the histogram concentration, that plaintext RDH methods rely on [5], [6].

Subsequent studies attempted to overcome this limitation by introducing block-level structure. Instead of treating the encrypted image as a completely random sequence, these approaches organize the encrypted image into blocks and apply reversible

operations such as block regulation, classification, scrambling, or pixel-value ordering. Such operations can provide controlled embedding opportunities while still keeping the image protected. Block regulation, in particular, is useful because it creates a predictable state space from which reversible embedding choices can be made [22].

Secondly, the method described here is based on this principle. Instead of embedding through intensity modulation, the secret message is embedded by representing it using legitimate permutations of the encrypted block values. This means that the embedding process will be related to the value ordering in a block, allowing the restoration of the initial block state at recovery time.

2.5 Deep Learning and Attention-based Prediction

Some recent RDH-EI investigations have considered deep learning since CNNs and attention mechanisms are capable of modeling complicated spatial relationships, which cannot easily be handled by the simple linear prediction method [10], [11], [20]. Simple linear predictors might work well on textured areas, repetitive structures, and edge parts due to the large residuals caused by local averaging techniques. The importance of the enhancement is also clear from the viewpoint of the prediction-based room reservation since a narrower residual distribution increases the embeddable area.

Even if there are some advantages in using deep models, they are not always convenient for the reversible systems since they need training data, parameter storage, and GPU processing with consistent floating-point calculations. To reconstruct the original image accurately in the RDH-EI method, the prediction process should be exactly duplicated in both the embedding phase and recovery phase. If there is a difference between the two processes, reconstruction becomes impossible.

The Adaptive Weighted Predictor proposed in this thesis follows this deterministic direction. It does not introduce a trained model or additional predictor-index map.

Instead, it uses the dot-set error map already obtained during the first prediction stage and converts that information into reliability-aware weights for cross-set prediction.

2.6 Comparative Analysis of Existing Methods

Table 2: Comparative analysis of representative RDH-EI approaches

Category	Representative idea	Strength	Limitation	Research gap addressed
Difference expansion / PEE	Embed data by expanding pixel differences or prediction errors	High reversibility and strong theoretical foundation	Primarily plaintext and vulnerable to privacy exposure	Need encrypted-image compatibility
Histogram shifting	Shift histogram bins to create reversible space	Simple and effective in smooth images	Capacity depends on histogram distribution	Limited encrypted-domain applicability
RRBE prediction methods	Reserve room before encryption using prediction errors	High embedding capacity	Prediction error can be large in complex regions	Need reliability-aware prediction
Multi-predictor methods	Choose best predictor from candidate set	Improved accuracy	Requires predictor index side information	Need adaptation without overhead
Deep learning predictors	Use CNN or attention models	Strong nonlinear prediction	High computation and hardware dependency	Need lightweight alternatives
VRAE block methods	Use encrypted-domain block structure	Data hider need not access plaintext	Capacity and recovery constraints are difficult	Need efficient index-based embedding

2.7 Research Gaps

- Many adaptive predictors use local context but do not distinguish between reliable and unreliable neighbors.
- Prediction errors generated in an earlier stage are often discarded even though they contain useful local reliability information.
- Multi-predictor selection can improve accuracy but may require an index map that reduces net capacity.
- Learning-based prediction may improve residual compactness but can be costly and difficult to reproduce deterministically in practical systems.
- VRAE methods still require encrypted-domain embedding mechanisms that provide capacity without exposing visual information or increasing side information excessively.
- A unified discussion is needed to relate RRBE prediction improvement and VRAE index modulation under the broader RDH-EI objective.

2.8 Summary

This chapter reviewed the development of RDH and RDH-EI with emphasis on the distinction between RRBE and VRAE. The literature indicates that prediction compactness is essential for RRBE capacity, while reversible encrypted-domain structure is essential for VRAE capacity. The following chapter presents the first proposed method, which improves checkerboard-based room reservation through adaptive weighted prediction.

CHAPTER 3

PROPOSED WORK – PART I: ADAPTIVE WEIGHTED PREDICTOR FOR CHECKERBOARD-BASED RDH-EI

3.1 Introduction

This chapter presents the first proposed contribution: an Adaptive Weighted Predictor (AWP) for checkerboard-based reversible data hiding in encrypted images. The method is designed for an RRBE setting, where the content owner performs preprocessing on the original image before encryption. The proposed idea is based on a simple observation: the error map produced in the first prediction stage carries information about local predictability. A region that produces small first-stage errors is likely to provide reliable context for a later prediction, whereas a region with large errors should be treated with caution.

A conventional two-stage checkerboard predictor divides the image into dot and cross sets. One set is predicted first and the recovered values are then used to support prediction of the other set. In the RMP plus AMP baseline, rhombus mean prediction is combined with adaptive mean prediction. Although this baseline is efficient, it does not assign different reliability levels to contextual pixels based on their previous prediction behavior. The proposed AWP introduces reliability-dependent weighting so that neighbors associated with lower first-stage error contribute more strongly to the cross-set estimate.

Figure 3. shows the complete processing chain. It connects checkerboard partitioning, dot-set prediction, error-map generation, adaptive cross-set prediction, encryption, data embedding, and marked encrypted image generation.

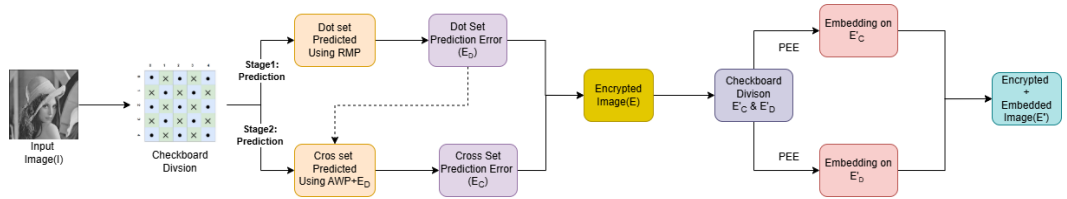


Figure 3: Overall Framework of the Proposed Adaptive Weighted Predictor Method

3.2 Related Background Specific to the Method

Room reservation based on prediction largely relies on the capability to represent prediction errors. The smaller the difference between the predicted and the original pixels, the greater will be the concentration of residuals around zero, and the less will be the amount of bits required to store the recovery information. Room reservation can be accomplished easily using checkerboard partitioning of the picture, since such a partitioning provides two different parity-based subsets. This partitioning scheme also provides an ordered prediction and reconstruction procedure, which is essential in avoiding circular dependence.

The Rhombus Mean Predictor is well-suited for the reversible preprocessing since the target pixel value is predicted using a deterministic neighbourhood in its immediate vicinity. It suffers from the disadvantage of treating all selected pixels identically. Adaptive Mean Predictor enhances the estimate by introducing the idea of direction, but it fails to evaluate whether a particular context element is relevant or not. Dot-set error map provides the means to accomplish this objective. An adjacent pixel with a smaller previous estimation error is assumed to be more relevant than a neighbour with a greater error.

The approach is still very feasible in terms of encryption of images since it does not need any kind of neural network training, additional indices to predict the values, nor any form of heavy computations. The owner of the content will only do simple preprocessing, allocate space, encrypt the image, and finally send it for data embedding.

Figure 4. shows the dot and cross pixel arrangement used by the method. This structure fixes the prediction order and supports exact reconstruction during recovery.

	0	1	2	3	4
0	×	•	×	•	×
1	•	×	•	×	•
2	×	•	×	•	×
3	•	×	•	×	•
4	×	•	×	•	×

Figure 4: Checkerboard Partitioning of Dot and Cross Pixel Sets

3.3 System Model

The architecture of this system is based on the following three roles: the content owner, the data hider, and the receiver. The process starts with the content owner, who takes an 8-bit grayscale image of size $H \times W$ and partitions it in the form of a checkerboard pattern followed by encryption. Two different sets of dots and crosses are considered, and the payload generation involves two steps that result in obtaining a single error map, which identifies areas suitable for room reservation. After the processing phase, the image undergoes encryption and becomes accessible to the data hider.

The latter embeds payload bits in reserved encrypted positions, using the data-hiding key to do so. The permitted operation at the receiver site is controlled by the existing keys. The operation that makes it possible to obtain the payload bits is performed via the use of the data-hiding key. In addition, if the image key and recovery information are present, it is possible to restore the image.

3.4 Proposed Methodology

3.4.1 Image Partitioning

Let I be an 8-bit grayscale image with dimensions $H \times W$. The image coordinates are separated into two parity-based sets. Pixels for which the coordinate sum is odd are

assigned to the dot set, while pixels for which the coordinate sum is even are assigned to the cross set. Equations (3.1) and (3.2) define these two sets. This separation is important because the dot-set residuals generated in the first stage are later used as reliability information for cross-set prediction.

$$S_D = \{(i, j) \mid (i + j) \bmod 2 \neq 0\} \quad (3.1)$$

$$S_C = \{(i, j) \mid (i + j) \bmod 2 = 0\} \quad (3.2)$$

3.4.2 Stage 1: Dot Prediction and Error Map Generation

For every dot-set location, the Rhombus Mean Predictor estimates the target value from four cardinal neighbours positioned at a distance of two pixels. Equation (3.3) gives the prediction rule. Since image pixels are discrete values, the floor operation is used so that the predicted value remains compatible with integer-valued grayscale reconstruction.

$$\hat{I}_D(i, j) = \text{floor} \left(\frac{I(i-2, j) + I(i+2, j) + I(i, j-2) + I(i, j+2)}{4} \right) \quad (3.3)$$

After the dot-set prediction is computed, the residual is obtained by subtracting the predicted value from the actual pixel value, as shown in Equation (3.4). This residual is preserved because it describes the local prediction behaviour of the region. In the second stage, the same residual information is used to decide how much influence neighbouring contexts should have during cross-set prediction.

$$E_D(i, j) = I(i, j) - \hat{I}_D(i, j), \quad \forall (i, j) \in S_D \quad (3.4)$$

3.4.3 Stage 2: Adaptive Weighted Cross-set Prediction

In the second stage, cross-set pixels are predicted after the dot-set residuals have already been calculated. For the target cross-set pixel at location (2, 2), the diagonal cross-set pixels at (1, 1) and (1, 3) are used as spatial predictors. This diagonal selection maintains the checkerboard dependency structure and allows the same prediction context to be reconstructed during the inverse process.

The prediction is guided by three dot-set residuals obtained from the first stage. Two residuals are associated with the selected diagonal predictors, and one residual lies between them as a shared local indicator. These residuals are treated as evidence of local reliability. If a surrounding dot pixel was predicted with a large error, the corresponding neighbourhood is considered less dependable for the cross-set estimate.

The weighting rule is designed to reduce the effect of predictors located in less reliable neighbourhoods. A smaller dot-set residual produces a larger reliability weight, whereas a larger residual produces a smaller weight. The scaling parameter k controls the sensitivity of this reliability mapping and is fixed at 0.03 in the experiments. The damping parameter λ controls how strongly the residual information is used for compensation and is fixed at 0.50 to avoid excessive correction in smooth regions.

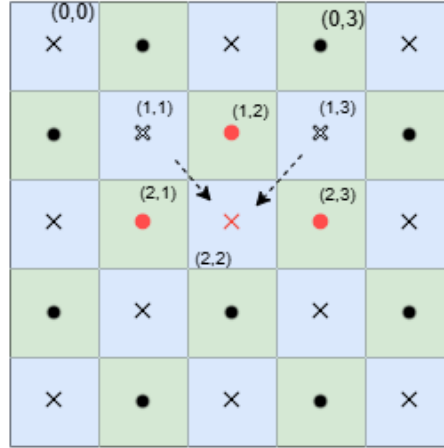


Figure 5: Prediction Context for Cross-set Pixel Using Error-guided Weights

Equation (3.5) converts the three associated first-stage errors into reliability weights. Because each denominator increases with $|E|$, a context with a larger error receives a smaller weight and contributes less to the final prediction.

$$w_1 = \frac{1}{1+k|E_1|}, \quad w_2 = \frac{1}{1+k|E_2|}, \quad w_s = \frac{1}{1+k|E_s|} \quad (3.5)$$

Equation (3.6) gives the final adaptive weighted prediction for the target cross-set pixel. The first two numerator terms use the two diagonal predictors after residual-

based compensation. The third numerator term uses their local average together with the shared residual. The denominator normalizes the weighted combination, ensuring that the estimate remains on the original intensity scale.

$$\hat{I}_C(2,2) = \frac{w_1(P_1 + \lambda E_1) + w_2(P_2 + \lambda E_2) + w_s\left(\frac{P_1 + P_2}{2} + \lambda E_s\right)}{w_1 + w_2 + w_s} \quad (3.6)$$

After the cross-set prediction is obtained, Equation (3.7) gives the second-stage prediction residual. These residual measures the difference between the actual cross-set value and the adaptive weighted estimate. Smaller residual magnitudes indicate a more compact prediction-error distribution and are advantageous for room reservation in RDH-EI.

$$E_C(2,2) = I_C(2,2) - \hat{I}_C(2,2) \quad (3.7)$$

3.5 Mathematical Formulation

The final prediction-error map is obtained by combining the residuals of both checkerboard sets. Equation (3.8) assigns the dot-set residual to locations belonging to SD and the cross-set residual to locations belonging to SC. This unified map is then used to identify pixels that are suitable for room reservation. Pixels with smaller residual magnitudes are preferred because they are easier to encode and more useful for reversible embedding.

$$E(i,j) = \begin{cases} E_D(i,j), & (i,j) \in S_D, \\ E_C(i,j), & (i,j) \in S_C. \end{cases} \quad (3.8)$$

The formulation remains deterministic because the weights used in the second stage are derived only from first-stage residuals that are already available during preprocessing. No additional predictor-selection map is required. This is important in RDH-EI because every extra recovery parameter reduces the net payload capacity.

$$I_e(i,j) = I(i,j) \oplus K_e(i,j) \quad (3.9)$$

Once the error map and embeddable positions are determined, the image is encrypted using a key-dependent stream cipher. Equation (3.9) represents the encryption operation, where each original pixel is combined with the corresponding key-stream

value using bitwise XOR. The resulting encrypted image can then be handled by the data hider without exposing the original visual content.

3.6 Algorithm Explanation

Table 3: Proposed adaptive weighted predictor-based room-reservation workflow

Step	Operation
1	Input grayscale image I and initialize dot and cross sets using checkerboard partitioning.
2	Predict dot-set pixels using rhombus mean prediction and compute dot-set error map ED .
3	For every cross-set target pixel, select diagonal predictors and associated dot-set errors.
4	Compute reliability weights from the magnitude of associated errors.
5	Apply damping-controlled weighted prediction to estimate cross-set pixels.
6	Combine dot-set and cross-set errors into unified error map E .
7	Select reliable embedding positions based on prediction error magnitude.
8	Encrypt the image and perform data embedding in the reserved positions.
9	At the receiver side, extract data and reconstruct the original image depending on key availability.

Table 3 lists the processing sequence of the proposed approach. Figure 6 presents the same process as a flow diagram, emphasizing the dependency from error-map generation to payload embedding and recovery.

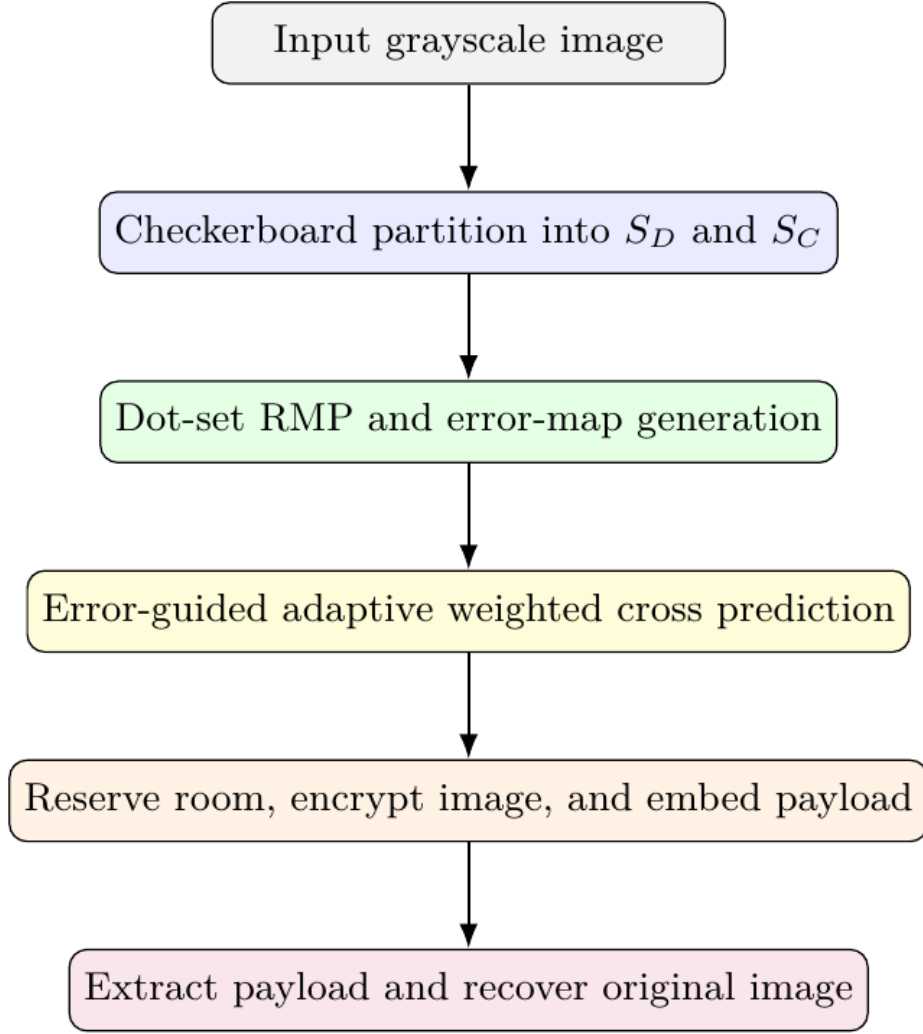


Figure 6: Complete Workflow of the Proposed AWP-based RDH-EI Method

3.7 Data Embedding and Image Recovery

The embedding stage uses the reserved positions identified from the unified prediction-error map. Before payload insertion, the original LSBs of the selected pixels are stored as recovery information. The image is then encrypted, and the payload bits are inserted into the reserved encrypted positions according to the data-hiding key. The same dot-set and cross-set processing order is maintained so that the recovery procedure can reproduce the required dependencies.

Let Ω represent the set of embeddable pixel positions. Equation (3.10) stores the original LSBs of these pixels, and Equation (3.11) describes the LSB substitution used to insert the payload into the encrypted image.

$$R = \{\text{LSB}(I(i, j)) \mid (i, j) \in \Omega\} \quad (3.10)$$

$$I'_e(i, j) = \begin{cases} I_e(i, j) - \text{LSB}(I_e(i, j)) + b_k, & (i, j) \in \Omega, \\ I_e(i, j), & \text{otherwise.} \end{cases} \quad (3.11)$$

At the receiver side, operations are performed according to available keys. With the data-hiding key, the embedded message can be read from the marked encrypted image. With the image decryption key and auxiliary recovery information, the original image can be reconstructed. If both keys are provided, the receiver can carry out payload extraction and exact image recovery. The inverse procedure restores substituted positions and then reverses the prediction dependency.

During extraction, Equation (3.12) reads the payload bit from the marked encrypted pixel, Equation (3.13) restores the corresponding encrypted pixel value, and Equation (3.14) recovers the original image by applying the image-decryption key.

$$b_k = \text{LSB}(I'_e(i, j)) \quad (3.12)$$

$$I_e(i, j) = I'_e(i, j) - b_k \quad (3.13)$$

$$I_d(i, j) = I_e(i, j) \oplus K_e(i, j) \quad (3.14)$$

Data embedding procedure.

- Select embeddable positions according to the magnitude of the unified prediction-error map.
- Store the original LSBs of selected pixels as auxiliary information required for perfect recovery.
- Encrypt the image using the image encryption key.
- Substitute payload bits into the reserved encrypted positions following the data-hiding key.
- Transmit the marked encrypted image and protected auxiliary information to the receiver.
- Data extraction and image recovery procedure.

- Use the data-hiding key to identify marked encrypted positions and extract the embedded payload.
- Restore the substituted positions by using the stored auxiliary information.
- Decrypt the restored encrypted image using the image encryption key.
- Reconstruct dot and cross pixels in the reverse prediction order to obtain the original image exactly.

3.8 Experimental Setup

The experiments for the first contribution are reported on standard 512×512 grayscale images, including CityHall, Pepper, Parthenon, Baboon, Barbara, Lena, Sailboat, Boat, Goldhill, and Airplane. These images contain different visual characteristics, ranging from smooth regions to strong textures and repeated patterns. This variation is useful for examining whether the predictor behaves consistently across different image structures.

Prediction results will be assessed based on MAE, MSE, and prediction error variance. The smaller their values are, the denser the distribution of residuals becomes. In case of RRBE-aided RDH-EI, this phenomenon is favorable since smaller residuals may need less bits to be encoded, leaving some space for auxiliary data embedding.

Figure 7 illustrates the test images utilized for the first contribution. These images have various spatial features, which enable testing the predictor on smooth and complex image areas.



Figure 7: Standard 512×512 Grayscale Test Images Used for Evaluation

Table 4: Prediction performance comparison between RMP+AMP baseline and proposed method

Image	Base MAE	Base MSE	Base VAR	Prop. MAE	Prop. MSE	Prop. VAR
CityHall	9.814	254.036	253.966	9.679	244.034	243.834
Pepper	5.388	90.227	90.061	5.316	85.313	85.168
Parthenon	6.988	120.902	120.847	6.844	115.034	114.941
Baboon	17.503	631.599	631.421	17.241	613.566	613.371
Barbara	11.622	403.855	403.744	11.467	394.946	394.809
Lena	5.413	87.843	87.757	5.323	84.941	84.813
Sailboat	8.681	200.024	199.891	8.560	191.894	191.730
Boat	8.346	183.304	183.205	8.248	174.848	174.702
Goldhill	7.618	128.837	128.801	7.532	125.002	124.936
Airplane	5.900	130.920	130.758	5.772	124.203	124.052
Average	8.727	223.155	223.051	8.598	215.378	215.239

3.9 Performance Evaluation and Discussion

As indicated in the experimental results, it can be observed that the presented AWP outperforms the baseline in terms of reducing MAE, MSE and variance. Such performance enhancement is essential since the efficient representation of prediction error is vital for effective room reservation. A more concentrated distribution in terms of error value will be favourable in terms of reversible embedding. Though the textured images like Baboon and Barbara pose a great challenge due to their fast change in pixel intensities, there is a visible improvement with the proposed algorithm.

In cases where the images have smooth transitions such as Lena, Pepper and Airplane, the effect of λ ensures that there is no over-compensation by the predictor. Otherwise, the process may end up correcting pixels that are already accurate.

Figure 8 provides a visual comparison of the average prediction metrics. The figure supports the tabulated result that the proposed approach produces a smaller average residual than the RMP+AMP baseline.

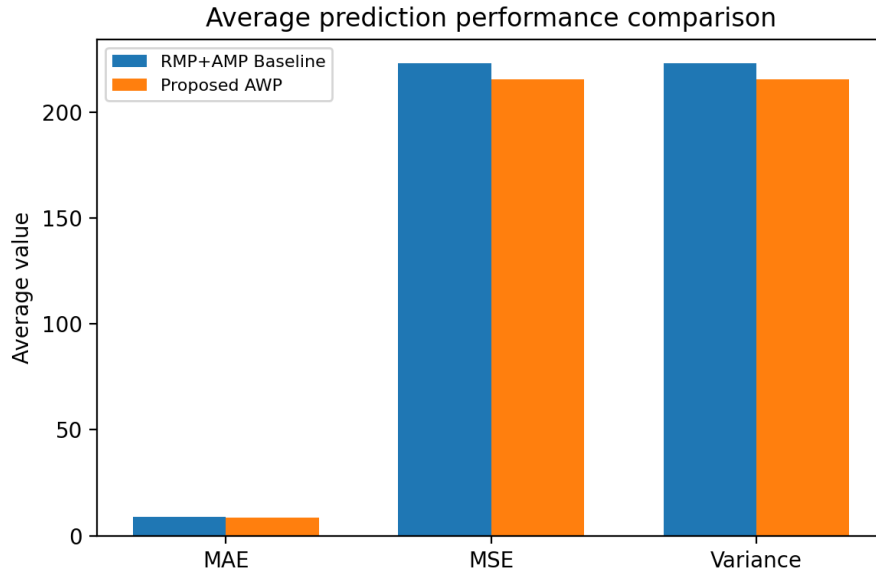


Figure 8: MAE, MSE, and Variance Comparison between Baseline and Proposed Method

The parameter sensitivity results indicate that $k = 0.03$ and $\lambda = 0.50$ provide the best reported average behaviour. Higher damping values increase the influence of residual compensation and may amplify noise. Larger values of k reduce weights more sharply and may suppress useful predictors. The chosen parameter setting therefore provides a practical balance between correction strength and prediction stability.

3.10 Applicability, and Limitations

As mentioned above, the suggested AWP algorithm is quite light since it makes use of a constant local neighborhood and a few arithmetic operations. It does not need any neural models or data training, GPU computations, and predictor-index map storage. Given that the reliability weights are computed based on an already calculated dot-set error map, the performance gains are achieved without the use of additional side information.

The increase in the values of MAE, MSE, and the variance of prediction errors demonstrates that the residual map becomes more concise than the baseline example. This is beneficial when using RRBE for RDH-EI since concise residuals mean better storage efficiency and enhanced possibility of performing a correct room reservation. The only drawback of the method lies in its dependence on local spatial correlation, implying that complicated textures might still cause large residual magnitudes.

The method is most suitable when the content owner can preprocess the plaintext image before encryption and when an efficient deterministic predictor is preferred over a heavy neural model. Its limitations are also clear: the reported evaluation is restricted to grayscale benchmark images, the parameters k and λ are fixed for the reported experiments, and the work does not yet examine color-image correlation or large-scale deployment under varying compression and transmission conditions.

3.11 Chapter Summary

This chapter described the Adaptive Weighted Predictor for checkerboard-based RDH-EI. The proposed method uses dot-set prediction errors to guide the cross-set prediction stage through reliability-based weights. The reported evaluation shows that the proposed predictor reduces MAE, MSE, and variance compared with the RMP plus AMP baseline. The next chapter presents the second contribution, which focuses on high-capacity encrypted-domain embedding through permutation-index modulation.

CHAPTER 4

PROPOSED WORK – PART II: INDEX-MODULATED HIGH-CAPACITY RDH-EI

4.1 Introduction

This chapter presents the second proposed contribution, namely an index-modulated high-capacity reversible data hiding method for encrypted images. In contrast to Chapter 3, where the image is analyzed before encryption for room reservation, this chapter focuses on an encrypted-domain embedding model. The method is intended to improve the payload limitation of conventional VRAE schemes by carrying secret data through permutation states of regulated encrypted blocks.

The proposed method avoids uncontrolled modification of image pixels. The image is first protected through two-stage encryption involving block permutation and intra-block pixel scrambling. The encrypted image is then divided into non-overlapping blocks. Eligible blocks are regulated into monotonic arithmetic sequences, and secret bits are represented by selecting valid permutation states. During recovery, the selected state is decoded, the regulated sequence is reversed, and the two-stage encryption is inverted to reconstruct the original image.

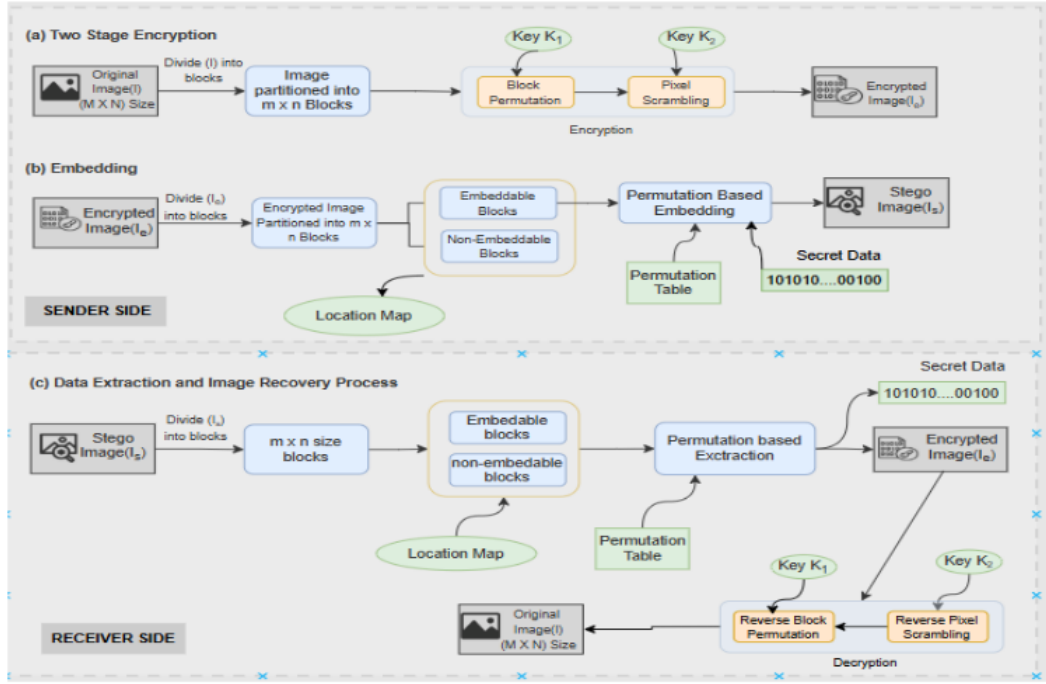


Figure 9:: Block Diagram of the Proposed Index-Modulated RDH-EI Method

4.2 Related Background Specific to the Method

In a VRAE setting, the data hider receives the encrypted image rather than the plaintext cover. This model is useful when the content owner wants to outsource encrypted images but still allow later metadata insertion. The difficulty is that encryption removes the visible correlation used by many reversible embedding methods. A VRAE method must therefore create a reversible embedding rule that works on encrypted blocks without disclosing image content.

Block-based methods address this problem by treating the encrypted image as a collection of structured units. A block may be classified, regulated, permuted, or accompanied by auxiliary information so that embedding can be reversed. The main requirement is that the marked encrypted block must still contain enough information for both payload extraction and recovery of the pre-embedding block state.

The proposed method is related to block-regulation-based covert communication. Once a block is transformed into a regulated sequence, different valid permutations of that sequence can represent different message values. If the data hider and receiver

share the mapping rule, the message can be extracted from the permutation index while the original block can still be recovered through inverse operations.

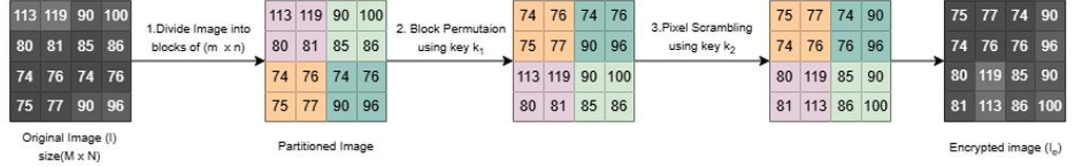


Figure 10: Illustration of Two-stage Encryption Process

4.3 System Model

The system model contains a content owner, a data hider, and a receiver. The content owner encrypts the image in two stages. First, non-overlapping image blocks are rearranged using a block permutation key. Second, pixels within each block are scrambled using a pixel scrambling key. This produces an encrypted image that hides both global layout and local pixel order.

The data hider does not require the original image. The payload is inserted by selecting permutation states of eligible regulated encrypted blocks. At the receiver end, the data-hiding key supports payload extraction, while the encryption keys support inverse scrambling and inverse block permutation. Because embedding is represented by reversible state selection, the original encrypted blocks can be restored before final decryption.

4.4 Proposed Methodology

4.4.1 Two-stage Encryption

The encryption process begins by dividing the original image into fixed-size, non-overlapping blocks. A block permutation sequence generated from key K_1 rearranges the blocks and removes the visible global structure. A second key K_2 is then used to scramble pixels inside each permuted block. This second stage reduces local spatial correlation.

The purpose of using two stages is to protect both the arrangement of image regions and the local ordering of pixels. At the same time, the block structure remains

available for later regulation and permutation-index embedding. The process is summarized below.

Two-stage image encryption process.

- Divide the original image into non-overlapping blocks of the selected size.
- Generate a block permutation sequence using the block permutation key K_1 .
- Rearrange image blocks according to the generated sequence.
- Generate intra-block scrambling sequences using the pixel scrambling key K_2 .
- Scramble pixels inside each permuted block to produce the encrypted image.

4.4.2 Block Regulation

For data embedding, each encrypted block is first represented as a one-dimensional vector V containing $N = m \times n$ pixel values. Equation (4.1) defines this block vector, where $t(i)$ represents the encrypted pixel value at position i

To create predictable embedding structures, an arithmetic progression vector is generated as:

$$R = \{0, d, 2d, \dots, (m \times n - 1)d\}$$

where d denotes the regulation parameter. The candidate regulated sequence A in Equation (4.2) arranges the block values into an ordered form suitable for reversible permutation-index embedding.

$$V = \{t(1), t(2), \dots, t(m \times n)\} \quad (4.1)$$

$$A = [a_1, a_2, \dots, a_N] \quad (4.2)$$

Only blocks satisfying the regulation rule are used for embedding. Equation (4.3) imposes the arithmetic step condition: adjacent regulated values must be separated by at least d and must remain inside the valid 8-bit intensity range. Non-embeddable blocks are preserved and recorded in the location map so that recovery remains exact.

$$a_{t+1} \geq a_t + d, \quad 0 \leq a_t \leq 255, \quad t = 1, 2, \dots, N - 1 \quad (4.3)$$

4.4.3 Permutation-index Embedding

After regulation, the distinct values within an embeddable block can be rearranged into valid permutation states, fig 11 depicts the complete Data Embedding Workflow Using Permutation-index Modulation. Equation (4.4) defines the number of available states P as the cardinality of the valid permutation set $\Pi(A)$. Equation (4.5) then gives the number of payload bits represented by one regulated block.

$$P = (m \times n)! / \prod(m_i!) \quad (4.4)$$

where m_i represents the repetition count of the i th unique pixel value.

$$b = \text{floor}(\log_2 P) \quad (4.5)$$

A group of b secret bits is converted into a decimal index as shown in Equation (4.6). The permutation state corresponding to this index is selected as the marked encrypted block, so the payload is carried by block ordering rather than by irreversible pixel modification.

$$idx = \text{bin2dec}(s_1 s_2 \dots s_b), \quad 0 \leq idx < 2^b \quad (4.6)$$

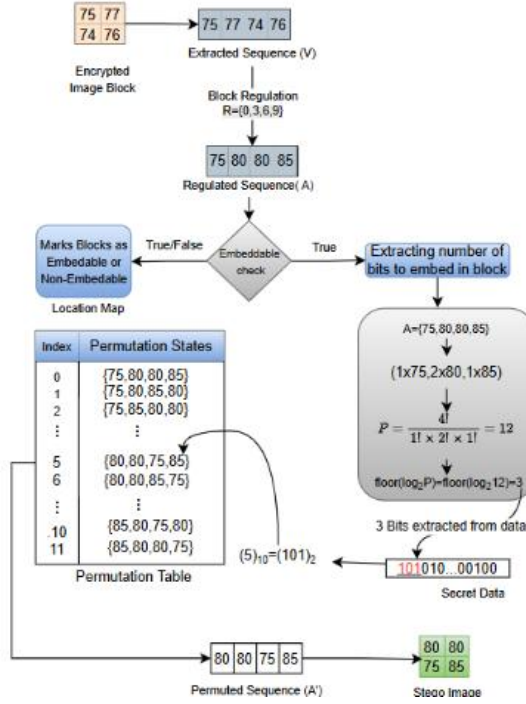


Figure 11: Data Embedding Workflow Using Permutation-index Modulation

4.5 Data Extraction and Image Recovery

During extraction, the marked encrypted image is processed block by block. The receiver first uses the location map to determine which blocks are embeddable. For each embeddable block, the present permutation state is identified and mapped back to its corresponding index. The index is then converted into the original bit sequence.

After the secret bits have been extracted, the marked block is restored to its regulated pre-embedding state. Inverse regulation recovers the encrypted pixel sequence. The receiver then applies inverse pixel scrambling with key K_2 and inverse block permutation with key K_1 . When the required keys and auxiliary data are correct, the recovered image is identical to the original image. the entire extraction and recovery process is depicted in fig 12.

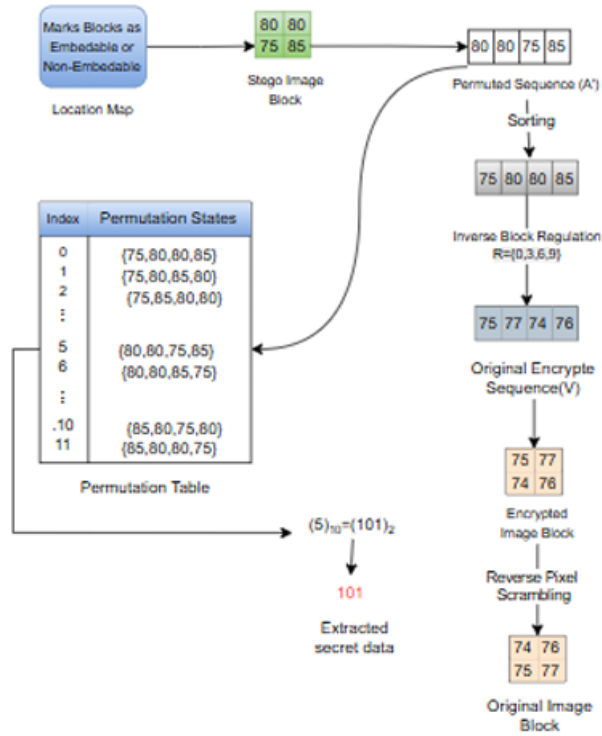


Figure 12: Data Extraction and Image Recovery Workflow

4.7 Algorithm Explanation

Step	Operation
1	Input original image and encryption keys.
2	Divide image into non-overlapping blocks.
3	Apply block permutation using key K1.
4	Apply intra-block pixel scrambling using key K2.
5	For each encrypted block, test regulation conditions and generate location map.
6	For embeddable blocks, determine valid permutation states and map secret bits to permutation index.

Step	Operation
7	Replace the regulated block arrangement with the selected permutation state.
8	At the receiver, identify embeddable blocks and decode permutation index to secret bits.
9	Restore regulated sequence, reverse pixel scrambling, reverse block permutation, and reconstruct the original image.

Figure 13. Depicts complete workflow, The diagram emphasizes that encryption, regulation, permutation-index embedding, extraction, and inverse recovery are ordered so that the original image can be restored.

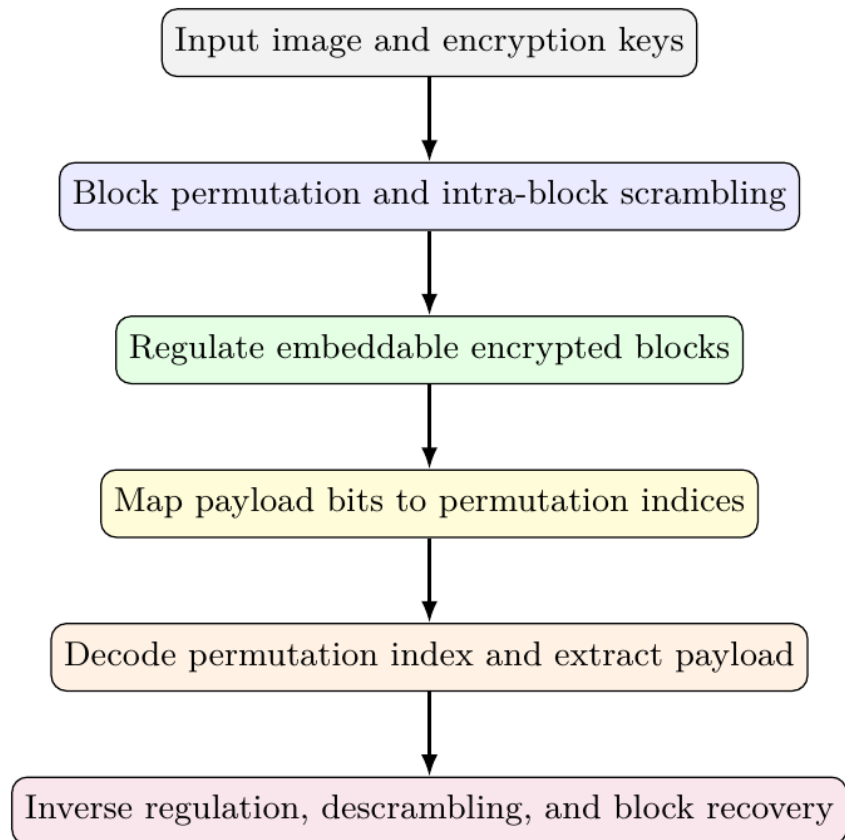


Figure 13: Complete Proposed Index-Modulated RDH-EI Workflow

4.8 Experimental Setup

The reported experiments use four standard 512×512 grayscale benchmark images: Baboon, Barbara, Boat, and Lena. These images have different texture and smoothness properties, making them useful for observing how the regulation parameter d influences embedding rate. Embedding rate is measured in bits per pixel.

Dataset-level evaluation is also reported for BOSSBase and BOWS-2. The reported reconstruction quality is perfect, with PSNR equal to infinity and SSIM equal to one. This indicates that the recovered image is identical to the original image under the reported experimental conditions.

Figure 14. shows the test images used in the evaluation of the second contribution.



Figure 14: Test Images Used for Evaluation - Baboon, Barbara, Boat, and Lena

Table 5: Embedding rate for different test images under different values of d

Image	$d = 1$	$d = 2$	$d = 3$	$d = 4$
Baboon	0.070	0.096	0.127	0.162
Barbara	0.153	0.258	0.360	0.450
Boat	0.120	0.188	0.265	0.346
Lena	0.161	0.278	0.402	0.512

4.9 Performance Evaluation and Discussion

The results show that the embedding rate increases as the step value d increases. A larger step can make more valid permutation states available for some blocks, which allows the block to represent more payload bits. The increase is image-dependent. Smooth images such as Lena provide higher reported embedding rates, while highly textured images such as Baboon provide lower rates.

The comparison with representative methods shows that the proposed scheme achieves higher embedding rates for the evaluated images while preserving perfect reconstruction. This is significant because the data hider operates in the encrypted domain and does not depend on plaintext spatial redundancy.

Table 6: Performance comparison for different test images

Image	Proposed	[20]	[17]	[21]	[8]	[9]
Baboon	0.162	0.055	0.045	0.057	0.005	0.001
Barbara	0.450	0.116	0.065	0.077	0.012	0.001
Boat	0.346	0.114	0.094	0.099	0.012	0.005
Lena	0.512	0.173	0.113	0.091	0.015	0.006

Table 7: Dataset-level performance comparison of the proposed method

Dataset	Average ER (bpp)	PSNR (dB)	SSIM
BOSSBase	0.5575	Infinity	1
BOWS-2	0.5247	Infinity	1

Figure 15. visualizes the embedding-rate values for the standard test images. The plot shows the difference between images and supports the observation that image texture and regulation parameters affect capacity.

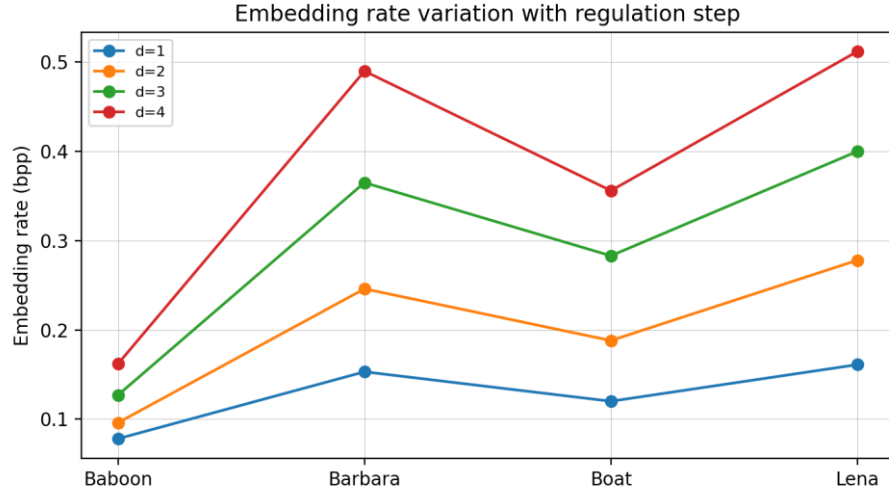


Figure 15: Embedding Rate Comparison on Standard Test Images

4.10 Applicability, and Limitations

The index-modulated method is suited to encrypted-domain data hiding because the data hider does not require access to the plaintext image. After the content owner applies block permutation and intra-block scrambling, the data hider works only with encrypted blocks. The message is inserted by selecting valid permutation states, so the payload is linked to reversible block ordering rather than direct pixel-value modification.

The achievable embedding rate depends on how many encrypted blocks satisfy the regulation rule and on the selected regulation parameter d . Blocks that cannot be regulated safely are not used for embedding and must be tracked through auxiliary information. Therefore, the practical capacity is influenced by the image content, block size, regulation step, and location-map overhead.

The main limitation of this method is that not every encrypted block can support permutation-index embedding. Non-embeddable blocks and recovery information must be managed carefully to preserve reversibility. The method is most effective

when a sufficient number of blocks can be regulated without producing excessive auxiliary data.

4.11 Chapter Summary

This chapter presented an index-modulated high-capacity RDH-EI method. The method combines block permutation, pixel scrambling, block regulation, and permutation-index embedding. The reported results indicate improved embedding rate and perfect image recovery under the evaluated conditions. The next chapter discusses both proposed methods together and interprets their comparative significance.

CHAPTER 5

CONCLUSION AND FUTURE SCOPE

5.1 CONCLUSION

This thesis presented two complementary approaches for improving reversible data hiding in encrypted images while preserving confidentiality, embedding capability, and exact image recovery. The first contribution focused on a Reserving Room Before Encryption framework through an Adaptive Weighted Predictor for checkerboard-based RDH-EI. By dividing the image into dot and cross pixel sets, predicting the dot set first, and reusing the resulting error map as reliability information for cross-set prediction, the method improved the compactness of the prediction-error distribution. The use of adaptive weights allowed more reliable neighboring contexts to contribute more strongly, while the damping factor prevented excessive correction in smooth regions. As a result, the proposed predictor achieved improved MAE, MSE, and prediction-error variance compared with the RMP plus AMP baseline.

The second contribution addressed high-capacity encrypted-domain embedding through an index-modulated RDH-EI method. In this approach, image confidentiality was maintained using two-stage encryption based on block permutation and intra-block pixel scrambling. Instead of embedding data by directly modifying encrypted pixel values, the method represented secret information through valid permutation states of regulated encrypted blocks. This design enabled reversible data insertion while maintaining deterministic extraction and perfect image recovery. The experimental results indicated that the method provides stable embedding performance across benchmark images and supports separable access, where data extraction and image recovery can be performed according to the available keys.

Overall, the thesis concludes that RDH-EI performance can be improved by addressing two important design bottlenecks: prediction accuracy in RRBE-based room reservation and structured embedding capacity in VRAE-style encrypted-domain processing. The Adaptive Weighted Predictor improves the reliability of prediction-driven room reservation, while the index-modulated method strengthens

high-capacity embedding without compromising reversibility. Together, the two methods demonstrate that carefully designed deterministic models can provide practical, secure, and reversible image data hiding mechanisms suitable for cloud-based and privacy-sensitive image management systems.

5.2 FUTURE SCOPE

Future work on the adaptive weighted predictor may extend the method from grayscale images to color images. In such an extension, the correlation among RGB channels can be used along with spatial context. The reliability function may also be improved by considering edge strength, local variance, or texture information in addition to first-stage prediction errors.

For the index-modulated method, further work may study adaptive block sizes, dynamic selection of the regulation step, and improved compression of the location map. A deeper security analysis may also be carried out under statistical, chosen-plaintext, and key-management attack models. These extensions would help evaluate the method under more practical deployment conditions.

Another possible direction is to combine the strengths of both proposed methods. Room may be reserved before encryption for a primary payload, while later updates may be inserted in the encrypted domain through permutation-index modulation. Such a hybrid model would be useful for cloud-based image records that require both initial secure storage and later encrypted-domain annotation.

REFERENCES

- [1] 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.
- [2] X. Li, J. Li, B. Li, and B. Yang, "High-fidelity reversible data hiding scheme based on pixel-value-ordering and prediction-error expansion," *Signal Processing*, vol. 93, no. 1, pp. 198-205, 2013.
- [3] 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. 13445-13457, 2018.
- [4] H. Honsinger, P. Jones, M. Rabbani, and J. Stoffel, "Lossless recovery of an original image containing embedded data," U.S. Patent 6 278 791, Aug. 2001.
- [5] X. Zhang, "Reversible data hiding in encrypted image," *IEEE Signal Processing Letters*, vol. 18, no. 4, pp. 255-258, 2011.
- [6] X. Zhang, "Separable reversible data hiding in encrypted image," *IEEE Transactions on Information Forensics and Security*, vol. 7, no. 2, pp. 826-832, 2012.
- [7] 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.
- [8] 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.
- [9] Y. Qiu, K. Chen, X. Lin, G. Li, and H. Zeng, "High-capacity reversible data hiding in encrypted images with edge-directed prediction and adaptive entropy coding," *IEEE Signal Processing Letters*, vol. 32, pp. 3012-3016, 2025.

- [10] 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.
- [11] S. Gandhi and R. Kumar, "SCAM-Net: Spatial-channel attention multi-scale network for reversible data hiding in encrypted images," in *Proc. IEEE International Conference*, 2024.
- [12] S. Gandhi and R. Kumar, "Survey of reversible data hiding: Statistics, current trends, and future outlook," *Computer Standards & Interfaces*, vol. 94, p. 104003, 2025.
- [13] R. Kumar and A. K. Sharma, "A review of different prediction methods for reversible data hiding," *Journal of Information Security and Applications*, vol. 78, 2023.
- [14] H. Ren, G. Bai, T. Chen, Z. Yue, and R. Ren, "Reversible data hiding in encrypted images with multi-prediction and adaptive Huffman encoding," *Scientific Reports*, vol. 13, p. 23104, 2023.
- [15] F. Ren, Z. Wu, Y. Xue, and Y. Hao, "Reversible data hiding in encrypted image based on bit-plane redundancy of prediction error," *Mathematics*, vol. 11, 2023.
- [16] S. Yi and Y. Zhou, "Reversible data hiding in encrypted images using prediction-error encoding," in *Proc. IEEE SMC*, 2018.
- [17] P. S. Sandhu, A. Malik, S. Singh, and V. Tyagi, "An innovative interpolation based reversible steganographic technique," *Procedia Computer Science*, vol. 259, pp. 1855-1862, 2025.
- [18] I. C. Dragoi and D. Coltuc, "Reversible data hiding in encrypted images based on reserving room after encryption and multiple predictors," in *Proc. IEEE ICASSP*, 2018.
- [19] H. Zhang, L. Li, and Q. Li, "Reversible data hiding in encrypted images based on block-wise multi-predictor," *IEEE Access*, vol. 9, 2021.

- [20] A. K. Saini and S. Singh, "Reversible data hiding using pixel error value ordering and new CNN predictor," *Signal, Image and Video Processing*, vol. 19, 2025.
- [21] Ankur, R. Kumar, A. K. Sharma, and P. Ranjan, "High-capacity reversible data hiding in encrypted images based on difference image transfiguration," *Signal, Image and Video Processing*, vol. 19, no. 5, p. 365, 2025.
- [22] S. Xu, C.-C. Chang, and J.-H. Horng, "Image covert communication with block regulation," *IEEE Signal Processing Letters*, vol. 30, pp. 1255-1258, 2023.
- [23] A. K. Rai, H. Om, S. Chand, and S. Agarwal, "Reversible data hiding in encrypted image using two-pass pixel value ordering," *Journal of Information Security and Applications*, vol. 76, p. 103545, 2023.
- [24] G. Cao, Y. Zhang, R. Ni, and X. Wang, "High-capacity reversible data hiding in encrypted images based on pixel value ordering," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 26, no. 3, pp. 523-536, 2016.
- [25] C.-Y. Weng, T.-W. Lin, H.-Y. Weng, and C.-T. Huang, "Enhanced reversible data hiding in encrypted images via median preserving pixel value ordering and block-wise difference preservation," *Signal, Image and Video Processing*, 2025.
- [26] G. Kumar, S. Singh, R. Rani, R. Kumar, and A. Malik, "High-quality reversible data hiding scheme using sorting and enhanced pairwise PEE," *IET Image Processing*, vol. 16, no. 4, pp. 1096-1110, 2022.
- [27] K. Kumar, S. Chand, and S. Singh, "A reversible high-capacity data hiding scheme using combinatorial strategy," *International Journal of Multimedia Intelligence and Security*, vol. 3, no. 2, pp. 146-161, 2018.
- [28] 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.

- [29] W. Hong, T.-S. Chen, and H.-Y. Wu, "An improved reversible data hiding in encrypted images using side match," *IEEE Signal Processing Letters*, vol. 19, no. 4, pp. 199-202, 2012.
- [30] Z. Zhang, "Reversible data hiding in encrypted images using patch-level sparse representation," *IEEE Transactions on Multimedia*, 2015.
- [31] R. Kumar and S. Chand, "Enhanced reversible data hiding in encrypted image using modified block centroid," *Multimedia Tools and Applications*, vol. 74, no. 17, pp. 6985-7007, 2015.
- [32] Y. Wang, G. Xiong, and W. He, "High-capacity reversible data hiding in encrypted images based on pixel-value-ordering and histogram shifting," *Journal of Visual Communication and Image Representation*, vol. 86, p. 103541, 2022.
- [33] P. Shiu, W. Tai, J. Chang, and C. Lin, "An interpolative AMBTC-based high-payload reversible hiding in encrypted images," *Signal Processing: Image Communication*, vol. 74, pp. 64-77, 2019.
- [34] X.-K. Cao, C.-C. Chang, and J.-H. Horng, "Image covert communication with block regulation," *IEEE Signal Processing Letters*, vol. 30, pp. 1217-1221, 2023.

Index-Modulated High-Capacity Reversible Data Hiding in Encrypted Images

Ajit Kumar Yadav

Department of Computer Science and Engineering
Delhi Technological University, Delhi, India
ajitkryadav2511@gmail.com

Rajeev Kumar

Department of Computer Science and Engineering
Delhi Technological University, Delhi, India
rajeevkumar@dtu.ac.in

Abstract—Lossless recovery of both encrypted image content and embedded auxiliary information is a key requirement in secure image storage and transmission systems. However, reversible data hiding performed after encryption typically achieves limited payload size because encryption disrupts exploitable image structures. This paper presents an index-modulated reversible data hiding method that enables high-capacity embedding directly in the encrypted domain. The scheme applies a dual-stage encryption process and adaptive block regulation to impose a deterministic ordering on encrypted image blocks. Secret data are conveyed through permutation indices derived from these ordered blocks, avoiding any modification of encrypted pixel values. Complete reconstruction of both the hidden message and the original image is guaranteed by inverting the index mapping and decryption steps. Experimental results show that the proposed method achieves an embedding capacity of 0.512 bits per pixel on the Lena image, and exceeds 0.52 bits per pixel on the BOSSBase and BOWS-2 datasets. Perfect reconstruction is consistently obtained, yielding infinite PSNR and SSIM equal to 1.0 across all tested images. These results demonstrate that index-based modulation provides a practical path toward high-capacity, distortion-free reversible data hiding in encrypted images.

Index Terms—RDHEI, Block Regulation, VRAE, Embedding Capacity, Permutation-Index

I. Introduction

With the rapid growth of digital communication and cloud-based storage, a large amount of sensitive visual information, such as medical images, surveillance data, and personal photographs, is transmitted and stored online. Ensuring the confidentiality and integrity of such data has therefore become a critical challenge [1]. **Data Hiding (DH)** offers a practical solution by embedding secret information within digital media, allowing covert communication without revealing hidden content [2]. However, traditional data hiding techniques cause permanent distortion to the cover image, making them unsuitable for scenarios where exact recovery of the original media is required [3]. To overcome this limitation, **Reversible Data Hiding (RDH)** techniques have been developed, which allow both hidden data and original media to be perfectly restored after extraction [4]. This property makes RDH particularly useful in secure transmission and cloud storage systems, where both confidentiality and perfect reconstruction of the original content are essential. Despite their effectiveness, most RDH methods operate in plain-text domain, leaving the cover content vulnerable to unauthorized access during data embedding or transmission [5].

To ensure both privacy and reversibility, researchers introduced Reversible Data Hiding in Encrypted Images (RDHEI),

which first encrypts the cover media/image and then embeds secret data into the encrypted domain. This framework preserves image confidentiality while ensuring lossless recovery. Existing RDHEI methods can be broadly categorized into two types: *Reserving Room Before Encryption (RRBE)* and *Vacating Room After Encryption (VRAE)*.

In RRBE schemes, the content owner pre-processes the original plaintext image to reserve space for data embedding before encryption. This approach exploits the strong spatial correlations present in the plaintext domain, allowing mature reversible data hiding techniques such as prediction error expansion (PEE), pixel value ordering (PVO), or histogram shifting to be applied effectively [6]. Since it operates on unencrypted data, RRBE achieves high EC and efficient utilization of spatial redundancy. However, its main limitation lies in security and practicality because reserving space requires direct access to the original image, which increases the risk of data exposure and adds computational overhead. Furthermore, once the image is encrypted, the EC becomes fixed, leaving little flexibility for adaptive data hiding in the encrypted domain.

In contrast, the VRAE paradigm first encrypts the image and subsequently embeds secret data directly within the ciphertext. This approach completely shields the original content from the data hider, ensuring full privacy preservation. By employing specially designed encryption mechanisms that retain partial inter-pixel correlation, VRAE methods enable decent EC and fully separable embedding without requiring plaintext access [7]. Nevertheless, many existing VRAE techniques suffer from limited embedding flexibility and overflow/underflow issues, as encryption often destroys pixel predictability, reducing the number of stable embeddable regions.

Motivated by these limitations, this paper presents a high-capacity RDHEI method under the VRAE framework based on a Permutation-Index Embedding strategy. Rather than altering ciphertext values, the proposed approach represents secret data through the positional arrangement of pixels inside each encrypted block. Each permutation pattern corresponds to a distinct data index, enabling lossless and fully reversible embedding without disturbing pixel statistics.

The key contributions of this paper are summarized as follows:

- A two-stage encryption mechanism that maintains block-level consistency within the ciphertext, enabling reliable and secure permutation-based data embedding.

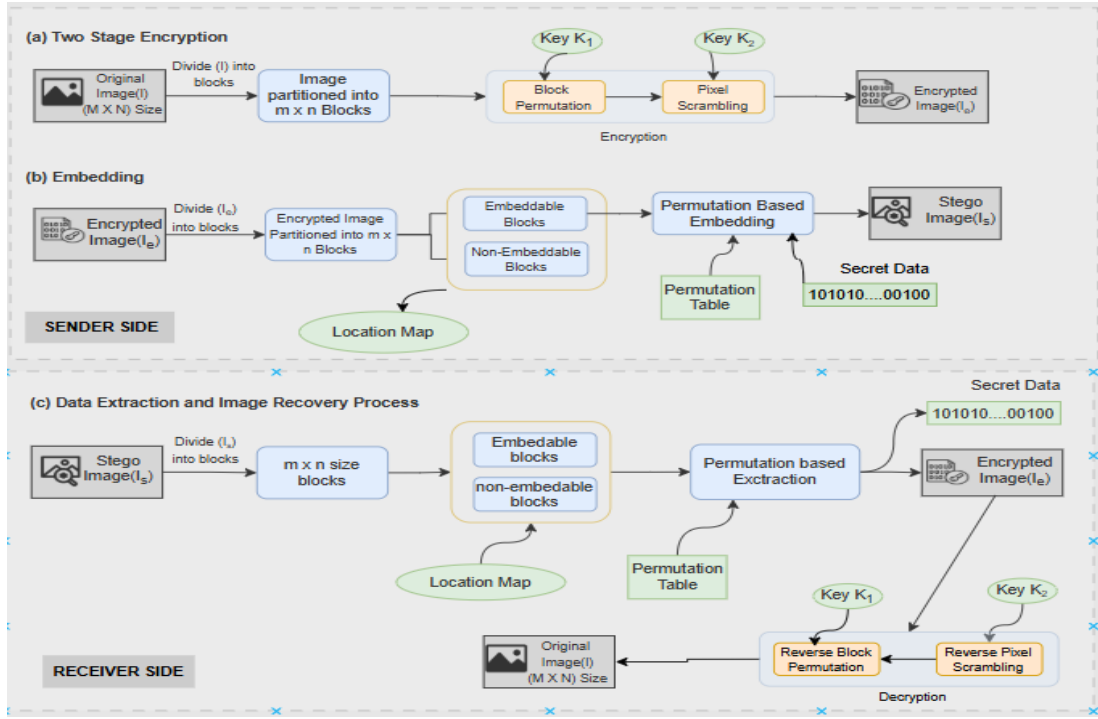


Fig. 1: Block diagram of the proposed RDHEI method.

- A permutation index embedding strategy that represents secret bits through specific permutation states, significantly enhancing EC.
- Comprehensive experimental validation demonstrating that the proposed method achieves higher EC than existing VRAE-based RDHEI schemes while preserving infinite PSNR, perfect reversibility, and strong visual security.

The remainder of this paper is organized as follows. Section II reviews related work. Section III presents the detailed methodology of the proposed scheme. Section IV discusses the experimental results and analysis, and finally Section V concludes the paper.

II. Related Work

RDHEI under the VRAE paradigm was pioneered by Zhang [8], who embedded data by flipping the least significant bits (LSBs) of encrypted pixels in block units. Although this method demonstrated the feasibility of separable RDHEI, it suffered from low EC and extraction errors due to the unpredictable nature of ciphertext/encrypted image. To improve robustness, Hong et al. [9] proposed a fluctuation-based block selection mechanism that embeds data only in smoother encrypted regions, thereby reducing extraction errors while retaining confidentiality.

Zhang later introduced an LSB-compression-based VRAE approach [10] that improved EC without requiring a location map. However, these pixel-modification-based VRAE techniques continued to face overflow/underflow issues and limited embedding flexibility. To mitigate this, Liao et al. [11] classified encrypted blocks using a local complexity measure, embedding data only in predictable blocks to minimize errors.

Khanam et al. [12] refined this strategy by enhancing complexity estimation, further improving reversibility. Bhardwaj et al. [13] incorporated symmetric encryption and pixel difference prediction to preserve image fidelity while increasing security.

Subsequent research focused on increasing EC without compromising full encryption. Wang et al. [14] and Ankur et al. [7] retained block-level correlation during encryption so that conventional RDH techniques such as Pixel Value Ordering (PVO) and histogram shifting could be applied in the ciphertext domain. Although these methods significantly increase EC, they still rely on pixel-value modification in the encrypted domain, which introduces auxiliary information overhead and requires overflow/underflow handling to maintain reversibility. Shiu et al. [15] utilized Interpolative Absolute Moment Block Truncation Coding (IAMBTC) and Huffman compression to free space in encrypted blocks, achieving higher payload at the cost of increased computational complexity. Recently, Sandhu et al. [16] proposed an interpolation-based reversible steganographic technique that improves embedding efficiency by exploiting spatial interpolation relationships between neighboring pixels.

To further enhance embedding efficiency, adaptive VRAE methods emerged. Yang et al. [17] introduced difference-preserving encryption to maintain pixel-value relationships post-encryption. More recently, the Adaptive Two-Stage (ATS) RDHEI scheme [18] combined PEE-based embedding with a two-stage encryption framework, preserving prediction error structures to enhance capacity. However, these schemes still modify encrypted pixel values and rely on overflow/underflow checks and location maps.

To overcome these limitations, recent research has shifted from pixel-value modification to structure-preserving VRAE,

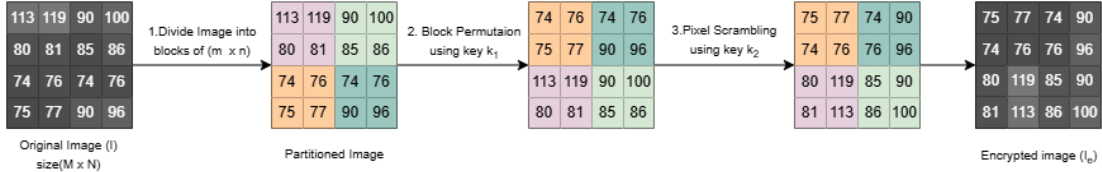


Fig. 2: Illustration of the Two-Stage Encryption Process

where secret data is embedded by rearranging encrypted pixels rather than altering their values. Inspired by Xu et al. [19], this paper develops a **Permutation-Index-Based VRAE** approach, in which secret bits are mapped to permutation states of encrypted blocks. This unique structural adaptation enables the high EC associated with VRBE while directly addressing the computational burden and plaintext manipulation risks of conventional methods, thereby improving both security and efficiency.

III. PROPOSED METHODOLOGY

The proposed methodology presents a VRAE-based reversible data hiding framework in which both data embedding and image recovery are performed entirely in the encrypted domain. As shown in Fig. 1, the system consists of two sides: the sender side and the receiver side. At the sender side, the original image undergoes a two-stage encryption process that ensures pixel confidentiality while preserving block-wise consistency required for reversible embedding. Secret data is then embedded using a permutation-index strategy, where information is encoded through controlled rearrangement of encrypted pixels rather than modifying their values, thereby avoiding distortion and overflow/underflow issues. At the receiver side, authorized users extract the embedded data by identifying permutation indices of encrypted blocks, while the content owner can further decrypt the image to achieve perfect reconstruction. This framework ensures high embedding capacity, zero distortion, strong confidentiality, and full reversibility. The complete workflow of sender and receiver side is discussed in phase wise manner as follows:

A. Two-Stage Encryption

Prior to embedding, the original image is encrypted to conceal its content. The encryption approach follows the method proposed by Rai et al. [20], incorporating image partitioning, block permutation, and pixel scrambling. Let I denote the original grayscale image of size $M \times N$. The image is divided into non-overlapping blocks of size $m \times n$, where block size affects both encryption strength and EC.

Stage 1: Block Permutation- Blocks are randomly rearranged using a *Block Permutation Key* K_1 , disrupting global spatial correlations.

Stage 2: Pixel Scrambling Pixels within each block are scrambled using a *Pixel Scrambling Key* K_2 , introducing intra-block randomness. The resulting encrypted image I_e appears statistically noise-like and visually unintelligible.

The procedure is summarized as follows:

This two-stage encryption also depicted in Fig. 2 effectively conceals both local and global spatial features, ensuring a secure foundation for subsequent reversible data embedding.

Algorithm 1 Two-Stage Image Encryption-

Input: Original image I of size $M \times N$, keys K_1, K_2
Output: Encrypted image I_e of size $M \times N$
Step1: Divide I into non-overlapping blocks of size $m \times n$
Step2: Permute the blocks randomly using key K_1
Step3: Scramble pixels within each block using key K_2

B. Embedding

To overcome the low EC of conventional VRAE schemes, the proposed framework operates in the partially encrypted domain. The method begins with a two-stage encryption process (as discussed in previous sub-section) that ensures pixel-level privacy while preserving the multiset of pixel values within each block. This preservation enables the subsequent Block Regulation process to reorganize each $m \times n$ encrypted block into a monotonic arithmetic sequence, creating predictable structures suitable for data embedding. Secret bits are mapped to specific permutation states of the regulated blocks through the Permutation Index Embedding strategy, achieving high embedding capacity with zero distortion and perfect reversibility. Non-overlapping blocks (typically 2×2 or 4×4) are processed sequentially to maintain uniformity and reversibility across the image. Finally, the modified and unmodified blocks are reassembled to generate the stego-encrypted image. The overall embedding workflow, as illustrated in Fig. shows the transformation from raw encrypted blocks V to regulated blocks A , and finally to permuted stego-blocks A' carrying the embedded data. The detailed step-by-step process is discussed as follows:

Step 1: Extract Pixel Sequence from Each Encrypted Block- Each $m \times n$ block in I_e is traversed in row-major order to form a one-dimensional pixel sequence $V = \{t(1), t(2), \dots, t(m \times n)\}$, where $t(i)$ denotes the encrypted grayscale value (0–255) at position i .

Example: For a 2×2 block,

$$\begin{bmatrix} p_{11} & p_{12} \\ p_{21} & p_{22} \end{bmatrix},$$

the sequence $V = \{p_{11}, p_{12}, p_{21}, p_{22}\}$. Since encryption alters pixel positions but not magnitudes, V retains the same multiset of values as the original image, ensuring perfect reversibility.

Step 2: Block Regulation via Arithmetic Progression- To make the encrypted block suitable for reversible embedding, the pixel sequence V is transformed into a regulated sequence:

$$A = \{a(1), a(2), \dots, a(m \times n)\}, \quad (1)$$

using an arithmetic-progression vector:

$$R = \{0, d, 2d, \dots, (m \times n - 1)d\}, \quad (2)$$

where d is a small integer step value (typically 1–4). Each pixel is updated as

$$a(i) = t(i) + (i - 1) \cdot d, \quad i = 1, 2, \dots, m \times n. \quad (3)$$

This regulation introduces a controlled linear progression, transforming the randomized pixel distribution into a predictable and sortable pattern, while ensuring that all modified intensities remain within the valid range $[0, 255]$.

After regulation, each block undergoes two critical validation checks to determine its eligibility for data hiding.

1. Monotonicity Check: The adjusted pixel sequence must follow a non-decreasing (or non-increasing) order, verified by the condition

$$a(i) \leq a(i + 1), \quad \forall i \in [1, n - 1] \quad (4)$$

2. Valid Range Check: All pixel intensities within the block must remain within the valid grayscale range, typically $[0, 255]$. Blocks containing values that result in an overflow (greater than 255) or underflow (less than 0) are immediately flagged as invalid.

Only blocks that satisfy both the monotonicity condition and the valid intensity range are designated as embeddable. Blocks failing either condition are designated as non-embeddable. A location map L is then used to record the embedding eligibility of every block, providing a compact means of identifying usable embedding regions.

Example: Given an original block $V = \{75, 77, 74, 76\}$ and an arithmetic progression $R = \{0, 3, 6, 9\}$ with common difference $d = 3$, the regulated block becomes $A = V + R = \{75, 80, 80, 85\}$. The resulting sequence is non-decreasing ($75 \leq 80 \leq 80 \leq 85$), qualifying the block as *embeddable*.

The proposed regulation serves two main purposes: it enforces a consistent structural pattern across encrypted blocks and primarily enhances embedding capacity by maximizing the number of distinct permutations available for data hiding, which effectively reserves the embedding room.

Step 3: Permutation-Index-Based Embedding. For each valid regulated block A , the number of distinct permutations P is computed as

$$P = \frac{(m \times n)!}{\prod_{i=1}^k m_i!}, \quad (5)$$

where m_i is the multiplicity of the i_{th} unique value, and k is the count of distinct pixel intensities. The block's EC is then given by:

$$B = \lfloor \log_2 P \rfloor \quad (6)$$

The next B bits from the secret data stream are converted into a decimal index $idx \in [0, P - 1]$, representing a specific permutation of A . Using a precomputed permutation table, such as one arranged in lexicographic order, the idx^{th} permutation forms the embedded block A' . In this process, only the order of pixels is altered, while their intensity values remain unchanged. As a result, the embedding introduces no distortion or no statistical change to the encrypted image. Furthermore, the permutation table need not be stored as auxiliary data, as it can be implicitly defined by the pixel composition of each regulated block. The auxiliary information is embedded in the image as done in [19].

Example: For $A = \{75, 80, 80, 85\}$, unique values are $75(1), 80(2), 85(1)$, giving

$$P = \frac{4!}{1! \times 2! \times 1!} = 12 \Rightarrow B = \lfloor \log_2 12 \rfloor = 3.$$

as shown in Fig. 3. If the next 3 secret data bits are “101”, they correspond to index 5 in the permutation table, selecting the 5th permutation (e.g., $\{80, 80, 75, 85\}$) as the embedded block.

During extraction, sorting A' reconstructs the regulated base A , and locating its permutation index recovers the embedded bits precisely, ensuring perfect reversibility. All embedded blocks A' and untouched blocks (non-embeddable) are then reassembled to produce the final stego-encrypted image I_s . The auxiliary information, such as the location map, regulation parameter d , and block size, is either stored separately or embedded in a reserved region.

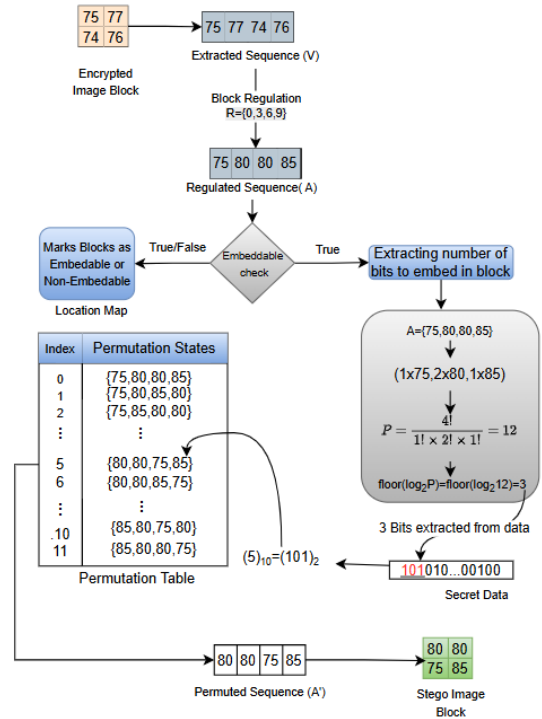


Fig. 3: Data embedding workflow showing transformation from encrypted block V to regulated block A , and finally to permuted block A' .

C. Data Extraction and Image Recovery

Upon receiving the stego-encrypted image, the receiver first extracts the auxiliary information, including the compressed location map and the regulation parameter d , as done in [19]. Using the location map, the image is divided into embeddable and non-embeddable blocks. For each embeddable encrypted block, both the secret data and the original encrypted block content are recovered as follows:

- 1) Retrieve the received permuted block:** Let the received block be

$$A' = \{a'(1), a'(2), \dots, a'(m \times n)\}. \quad (7)$$

- 2) **Recover the regulated sequence A :** The elements of A' are sorted to restore the regulated sequence used during embedding:

$$A = \begin{cases} \text{sort}(A') \text{ in ascending order,} & \text{if } d > 0 \\ \text{sort}(A') \text{ in descending order,} & \text{if } d < 0. \end{cases}$$

- 3) **Extract embedded secret data:** Since A may contain repeated pixel values, all its unique permutation states are enumerated based on the multiplicities of its elements (as defined in (7)). The index idx of the received sequence A' in this permutation table is determined. This index is then converted to its binary representation to recover the embedded secret bits for the block:

$$\text{Secret bits} = \text{bin}(idx).$$

- 4) **Recover the original encrypted block:** Once the secret bits are extracted, the regulated sequence A is used to reconstruct the original encrypted pixel sequence by inverting the block regulation: $t(i) = a(i) - (i-1) \cdot d$, where $i = 1, 2, \dots, m \cdot n$. This yields the original encrypted block $V = \{t(1), t(2), \dots, t(m \cdot n)\}$, which is then placed back into its original position in row-major order.

This process is repeated for all embeddable blocks, while non-embeddable blocks remain unchanged. After processing all blocks, the complete encrypted image is fully reconstructed.

Example: Consider a stego-encrypted block with received permuted sequence $A' = \{80, 80, 75, 85\}$ and regulation difference $d = 3$.

Secret data extraction: Sorting A' gives the regulated sequence $A = \{75, 80, 80, 85\}$. All distinct permutations of A are enumerated, and A' is found at index $idx = 5$. Thus, the embedded bits are: $(5)_{10} = (101)_2$.

Encrypted block restoration: Using $t(i) = a(i) - (i-1) \cdot d$, the original encrypted block is recovered as: $\{75, 77, 74, 76\}$, which is placed back in row-major order. This confirms that both secret data and the encrypted image content are perfectly restored.

Once the hidden data has been successfully extracted and the original pixel values have been restored, the final step involves decrypting the image to recover the original content. The decryption process follows the inverse of the encryption steps applied earlier. First, each block's pixels are reverse scrambled using the pixel scrambling key K_2 to restore their original order. Then, the blocks are reverse permuted using the same block permutation key K_1 to return them to their original positions in the image. Together, these steps fully decrypt the image, recovering its original form.

IV. EXPERIMENTAL RESULTS

This section evaluates the performance of the proposed RDHEI scheme through a series of experiments conducted on four standard grayscale benchmark images: Baboon, Barbara, Boat, and Lena, each having a resolution of 512×512 pixels (as shown in Fig. 5). The primary performance metric considered is the ER, represented in terms of Embedding Rate (ER) and measured in bits per pixel (bpp).

Table I presents the variation in ER for each test image with respect to increasing values of the common difference d . The results show that ER increases steadily with higher values of

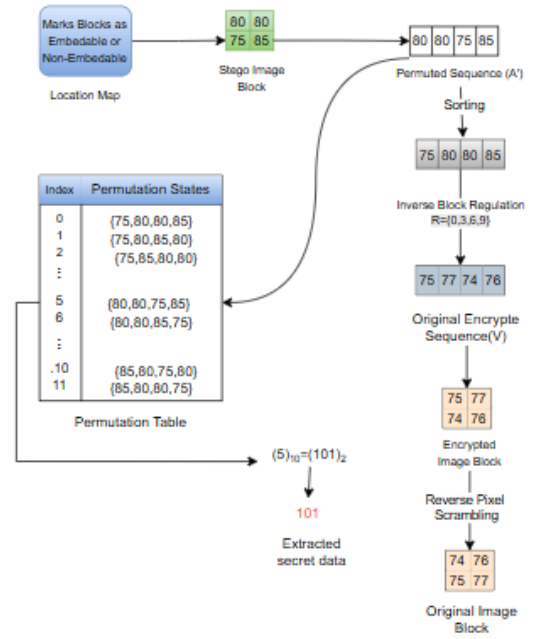


Fig. 4: Example illustrating Data extraction

d across all test images. For instance, for the Lena image, ER improves from 0.161 bpp ($d = 1$) to 0.512 bpp ($d = 4$). This trend indicates that as d increases, the probability of obtaining more monotonic sequences also increases, thereby providing a larger number of embeddable blocks and consequently higher embedding efficiency.

TABLE I: ER for different test images

Test images	$d = 1$	$d = 2$	$d = 3$	$d = 4$
Baboon	0.070	0.096	0.127	0.162
Barbara	0.153	0.254	0.360	0.450
Boat	0.120	0.188	0.265	0.346
Lena	0.161	0.278	0.402	0.512

To further validate the effectiveness of the proposed method, Table II compares the ER performance against several state-of-the-art RDHEI techniques [8], [9], [17], [21], and [20]. The proposed method outperforms all existing methods across all test images. The proposed method consistently outperforms all existing approaches across all benchmark images. For instance, on the Baboon image, the proposed scheme attains an ER of 0.162 bpp, surpassing the best-performing existing method by Rai et al. [20], which achieved only 0.055 bpp. Likewise, for the Lena image, the proposed method achieves 0.512 bpp, nearly three times higher than the value reported by [20], and considerably superior to the other referenced techniques.

The scalability and generalization capability of the proposed scheme were further evaluated using two large-scale datasets, BOSSBase and BOWS-2. As summarized in Table III, the proposed method maintains consistently high performance, achieving an average ER of 0.5575 bpp on BOSSBase and 0.5262 bpp on BOWS-2. The corresponding PSNR values remain at infinity and SSIM equals 1, confirming that the marked encrypted images are visually and statistically identical to their original counterparts. In comparison, the average ERs

TABLE II: Performance comparison for different test images

Images	Proposed	[20]	[17]	[21]	[8]	[9]
Baboon	0.162	0.055	0.045	0.005	0.005	0.001
Barbara	0.450	0.116	0.065	0.057	0.012	0.001
Boat	0.346	0.114	0.094	0.099	0.012	0.005
Lena	0.512	0.173	0.113	0.091	0.015	0.006

reported by [20] are 0.3524 bpp (BOSSBase) and 0.3387 bpp (BOWS-2), which are notably lower than those obtained by the proposed method. These results demonstrate the robustness, scalability, and superior embedding efficiency of the proposed RDHEI scheme across diverse image datasets.

TABLE III: Performance Comparison on two big datasets

	BOSSBase	BOWS-2
Average ER (bpp)	0.5575	0.5247
PSNR (dB)	∞	∞
SSIM	1	1

These results conclusively demonstrate that the proposed RDHEI scheme not only surpasses existing methods in terms of EC but also ensures perfect reversibility and imperceptibility, thereby making it well-suited for large-scale and secure data hiding applications.

V. CONCLUSION

In this paper, a high-capacity reversible data hiding method for encrypted images is presented to overcome the limitations of conventional VRAE schemes, which often suffer from low EC due to the loss of pixel correlation during encryption. The proposed approach employs a two-stage encryption process with block permutation and pixel scrambling for enhanced confidentiality, followed by block regulation to restore structured predictability in encrypted blocks. Secret data are represented as permutation indices rather than altering pixel values, achieving distortion-free embedding and perfect recovery of both the hidden data and the original image. Experimental results show that the method achieves higher EC than existing VRAE techniques while maintaining infinite PSNR and unity SSIM across standard datasets such as BOSSBase and BOWS-2. The proposed framework demonstrates strong scalability and reliability, making it suitable for secure medical imaging, cloud storage, and forensic applications.

Future research can extend the proposed permutation-index based RDHEI framework to color images and high-resolution multimedia data, where channel-wise or cross-channel permutation strategies may further improve embedding efficiency. Another promising direction is the development of adaptive block regulation mechanisms that dynamically adjust the regulation parameter based on local image characteristics to maximize the number of embeddable blocks. In addition, investigating security robustness against advanced steganalysis techniques and integrating lightweight cryptographic primitives for secure cloud-based image storage and transmission would further enhance the practical applicability of the proposed scheme.



Fig. 5: Test images: (a) Baboon (b) Barbara (c) Boat (d) Lena

References

- [1] S. Gandhi and R. Kumar, "Survey of reversible data hiding: Statistics, current trends, and future outlook," *Computer Standards & Interfaces*, p. 104003, 2025.
- [2] R. Kumar, S. Chand, and S. Singh, "A reversible high capacity data hiding scheme using combinatorial strategy," *International Journal of Multimedia Intelligence and Security*, vol. 3, no. 2, pp. 146–161, 2018.
- [3] A. Malik, R. Kumar, and S. Singh, "A new image steganography technique based on pixel intensity and similarity in secret message," in *2018 International Conference on Advances in Computing, Communication Control and Networking (ICACCCN)*. IEEE, 2018, pp. 828–831.
- [4] G. Kaur, S. Singh, R. Rani, R. Kumar, and A. Malik, "High-quality reversible data hiding scheme using sorting and enhanced pairwise pee," *IET Image Processing*, vol. 16, no. 4, pp. 1096–1110, 2022.
- [5] S. Kumar, A. Gupta, and G. S. Walia, "Reversible data hiding: A contemporary survey of state-of-the-art, opportunities and challenges," *Applied Intelligence*, vol. 52, no. 7, pp. 7373–7406, 2022.
- [6] 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.
- [7] Ankur, R. Kumar, and A. K. Sharma, "Bit-plane based reversible data hiding in encrypted images using multi-level blocking with quad-tree," *IEEE Transactions on Multimedia*, vol. 26, pp. 4722–4735, 2024.
- [8] X. Zhang, "Reversible data hiding in encrypted image," *IEEE Signal Processing Letters*, vol. 18, no. 4, pp. 255–258, 2011.
- [9] W. Hong, "An improved reversible data hiding in encrypted images using side match," *IEEE Signal Processing Letters*, vol. 19, no. 4, pp. 199–202, 2012.
- [10] X. Zhang, "Separable reversible data hiding in encrypted image," *IEEE Transactions on Information Forensics and Security*, vol. 7, no. 2, pp. 826–832, 2012.
- [11] X. Liao, X. Shu, and et al., "Reversible data hiding in encrypted images based on local complexity," *Journal of Visual Communication and Image Representation*, vol. 25, no. 5, pp. 1172–1180, 2014.
- [12] R. Khanam and S. Chand, "Enhanced reversible data hiding method in encrypted images using improved block complexity," *Multimedia Tools and Applications*, 2015.
- [13] M. Bhardwaj and P. Verma, "Reversible data hiding in encrypted images using secret sharing," *Multimedia Tools and Applications*, vol. 74, no. 17, p. 6985–7007, 2015.
- [14] Y. Wang, G. Xiong, and W. He, "High-capacity reversible data hiding in encrypted images based on pixel-value-ordering and histogram shifting," *Unpublished*, 2020.
- [15] P. Shiu, W. Tai, J. Jan, C. Chang, and C. Lin, "An interpolative ambtc-based high-payload rdh scheme for encrypted images," *Signal Processing: Image Communication*, vol. 74, pp. 64–77, 2019.
- [16] P. S. Sandhu, A. Malik, S. Singh, and V. Tyagi, "An innovative interpolation based reversible steganographic technique," *Procedia Computer Science*, vol. 259, pp. 1855–1862, 2025.
- [17] C.-H. Yang, C.-Y. Weng, and J.-Y. Chen, "High-fidelity reversible data hiding in encrypted image based on difference-preserving encryption," *Soft Computing*, vol. 26, pp. 1727–1742, 2022.
- [18] P. exact authors from your PDF here, "Adaptive two-stage reversible data hiding in encrypted images using prediction error expansion," (*You can add exact journal or write "Unpublished"*), 2023.
- [19] S. Xu, C.-C. Chang, and J.-H. Horng, "Image covert communication with block regulation," *IEEE Signal Processing Letters*, vol. 30, pp. 1217–1221, 2023.
- [20] A. K. Rai, H. Om, S. Chand, and S. Agarwal, "Reversible data hiding in encrypted image using two-pass pixel value ordering," *Journal of Information Security and Applications*, vol. 76, p. 103545, 2023.
- [21] R. Anushidevi and R. Amirtharajan, "Separable reversible data hiding in an encrypted image using the adjacency pixel difference histogram," *Journal of Information Security and Applications*, vol. 72, p. 103407, 2023.

Adaptive Weighted Predictor for Checkerboard based Reversible Data Hiding in Encrypted Images

Ajit Kumar Yadav

Department of Computer Science and Engineering
Delhi Technological University, Delhi, India
ajitkryadav2511@gmail.com

Rajeev Kumar

Department of Computer Science and Engineering
Delhi Technological University, Delhi, India
rajeevkumar@dtu.ac.in

Abstract—With the widespread use of cloud computing, protecting the privacy of off-site multimedia data has become a key challenge. Reversible Data Hiding in Encrypted Images (RDH-EI) addresses this by enabling data embedding in encrypted images while ensuring exact recovery. In RRBE-based schemes, embedding capacity depends on prediction accuracy, motivating the use of advanced predictors. Recent approaches leverage deep learning for improved prediction; however, they rely on checkerboard partitioning with mutual prediction between dot and cross pixel sets, which is not applicable in RDH-EI. In practice, prediction is only feasible from cross to dot pixels, leaving the reverse direction unused and resulting in reliance on Adaptive Mean Prediction (AMP) for second-stage prediction. Although AMP is efficient and direction-aware, it assumes equal reliability of neighboring pixels and neglects dot-set prediction errors, limiting performance in textured and edge regions. To address this limitation, this paper proposes a new Adaptive Weighted Predictor (AWP) within a two-stage framework. dot prediction is first performed using a rhombus mean predictor to obtain a reconstruction error map. This error information is then exploited by the proposed AWP to assign adaptive weights to cross-set neighbors, with a damping factor introduced to suppress noise amplification in smooth regions. Experimental results show that the proposed method consistently outperforms the RMP plus AMP baseline, achieving average reductions of 1.48% in MAE, 3.48% in MSE, and 3.50% in prediction error variance.

Keywords: Reversible Data Hiding, Encrypted Images, Adaptive Prediction, Weighted Error Correction, Image Privacy.

I. INTRODUCTION

Cloud computing enables convenient storage and sharing of multimedia data, but it also raises serious security concerns such as unauthorized access, data leakage, and content manipulation. Hence, ensuring confidentiality, controlled access, and data integrity has become a primary need in modern day cloud-based image management systems. For this purpose, Data Hiding (DH) is being recognized as an effective method against security threats vulnerability in terms of embedding secret information in multimedia content. However, a major drawback associated with traditional techniques of data hiding is that the original content is altered in an irreversible manner, which is not suitable in situations where exact restoration of data is required.

Reversible data hiding (RDH) enables embedding auxiliary data while allowing perfect recovery of the original image. Early RDH techniques, such as difference expansion [1], [2],

and histogram shifting [3] demonstrated that extra data can be appended without losing the possibility of recovering the image [4]. Thus, this extraordinary feature has turned RDH into a research topic for industries with the utmost demand on visual content integrity. However, traditional RDH methods leave the cover image exposed, which poses a serious privacy risk. Reversible Data Hiding in Encrypted Images (RDH-EI) was introduced to meet this requirement by allowing data to be embedded into encrypted images without compromising reversibility. RDH-EI scenario, three parties are involved: the content owner who encrypts the image, the data hider who adds auxiliary data, and the recipient who retrieves the data and/or recovers the original image. Depending on when the embedding space is generated, existing RDH-EI methods are generally classified into vacating room after encryption (VRAE) and reserving room before encryption (RRBE) frameworks. VRAE-based methods attempt to create embedding space directly in the encrypted domain however, strong encryption significantly reduces spatial redundancy, resulting in limited embedding capacity [5], [6]. To overcome this limitation, Ma et al. [7] introduced the RRBE framework, which exploits spatial correlation in the plaintext image to reserve embedding space prior to encryption. Because the reserved space is preserved during encryption, RRBE-based methods are widely regarded as the most effective solution for high-capacity RDH-EI in cloud storage scenarios.

Within the RRBE framework, embedding capacity is largely determined by the distribution of prediction errors generated during preprocessing. More accurate prediction produces prediction-error histograms that are concentrated around zero, enabling efficient space reservation using multi-MSB substitution [7], [8]. Consequently, improving prediction accuracy has become an important research focus in high-capacity RDH-EI systems.

Existing prediction techniques, however, face a persistent trade-off between computational efficiency and prediction performance. Classical linear predictors, including the Median Edge Detector (MED) and Gradient-Adjusted Predictor (GAP), are computationally efficient but rely on fixed coefficients or uniform averaging, their lack of directional adaptivity often leads to large prediction errors in complex image regions, resulting in inefficient room reservation in RDH-EI applications. [8], [9]. To overcome the limitations of fixed-

context predictors, AMP was proposed to dynamically select the most suitable directional context based on local image characteristics. By adapting the prediction direction, AMP effectively reduces directional mismatch and has demonstrated superior performance over conventional predictors in RDH-EI systems. Consequently, AMP has been widely adopted as a strong baseline predictor in recent RDH-EI literature. Despite its effectiveness, AMP treats all prediction contexts uniformly and does not consider the local reliability of prediction information. In practical RDH-EI systems, prediction accuracy can vary significantly across different image regions, and ignoring such reliability differences may lead to error propagation. Consequently, the resulting prediction-error distributions may become less compact, reducing the efficiency of room reservation. This limitation becomes more pronounced in hierarchical or multi-stage prediction frameworks, which are discussed in the next section. More recently, deep learning-based predictors have been introduced to model complex spatial dependencies and achieve superior prediction accuracy [10], [11]. Despite their effectiveness, these methods incur substantial computational overhead and require GPU acceleration, making them less suitable for large-scale or resource-constrained cloud environments.

To address these limitations, this paper proposes a lightweight error-guided prediction strategy tailored for RRBE-based RDH-EI in cloud storage systems. The proposed approach employs a two-stage prediction framework in which reconstruction errors generated in the first stage guide prediction in the second stage. By adaptively weighting neighboring pixels according to their local reliability and introducing a damping mechanism to suppress error amplification in smooth regions, the proposed method improves prediction accuracy without increasing computational complexity or requiring additional side information.

The remainder of this paper is organized as follows. Section II reviews related work on prediction techniques for RDH-EI. Section III presents the proposed adaptive weighted prediction methodology. Section IV describes the data embedding process, while Section V explains the data extraction and image recovery procedure. Section VI presents the experimental results and analysis. Finally, Section VII concludes the paper.

II. RELATED WORK

The evolution of prediction algorithms in RDH-EI has been driven by the continuous pursuit of minimizing the prediction error to maximize the reserved room for data embedding. As categorized in recent surveys [12], [13], existing prediction methodologies can be broadly classified into static linear predictors, block-wise multi-predictors, adaptive linear predictors, and deep learning-based approaches.

A. Static and Linear Predictors

Early RRBE schemes relied on standard predictors adapted from lossless compression standards due to their simplicity and reversibility. The Median Edge Detector (MED), originally used in JPEG-LS, was widely adopted in early RDH-EI works.

For instance, Puteaux et al. [8] utilized MED to predict pixels and embed data into the Most Significant Bits (MSB). While MED effectively handles sharp edges in simple images, its fixed coefficients are insufficient for capturing complex local variations. To address this, Ren et al. [14] proposed a scheme combining MED with the Gradient-Adjusted Predictor (GAP) to generate label maps for non-reference pixels. Similarly, Ren et al. [15] explored bit-plane redundancy in prediction errors but relied on basic gradient estimations. Prediction-based RDHEI methods have also been explored to improve embedding efficiency. For instance, Yi and Zhou [16] proposed a prediction-error encoding strategy for reversible data hiding in encrypted images. These static filters are computationally efficient ($O(1)$) but lack the flexibility to adapt to the rapidly changing textures found in natural images, resulting in suboptimal embedding capacities. Recent works such as Sandhu et al. [17] have also explored interpolation-based RDH techniques for improving embedding capacity while maintaining good visual quality.

B. Block-Wise and Multi-Predictor Schemes

To overcome the limitations of single-model predictors, researchers introduced Ensemble or multi-predictor strategies. The core idea is to test multiple predictors on a block of pixels and select the one that yields the minimum error. Dragoi and Coltuc [18] proposed a Multiple Predictor approach for VRAE, where pixels are estimated using four distinct predictors (Average, Weighted Average, Median, and Midpoint), and the optimal result is selected. Extending this concept to the RRBE framework, Zhang et al. [19] introduced a Block-Wise Multi-Predictor, which tests up to 16 different prediction models for each image block. While these ensemble methods significantly improve prediction accuracy, they incur a "Payload Penalty": the index of the optimal predictor for each block must be stored and transmitted as side information. In high-capacity scenarios, this overhead consumes a significant portion of the reserved room, negating some of the gains from improved prediction.

C. Adaptive and Edge-Directed Prediction

A significant leap in efficiency was achieved with adaptive predictors that adjust their coefficients based on local context without requiring side information. Qiu et al. [9] utilized Edge-Directed Prediction (EDP) combined with adaptive entropy coding to dynamically allocate errors based on edge strength. Ping et al. [10] proposed the Adaptive Mean Predictor (AMP) within a two-stage checkerboard framework. AMP has been widely adopted in recent RRBE-based RDH-EI schemes due to its low computational complexity, deterministic behavior, and strict reversibility constraints. Although AMP exhibits inferior standalone prediction accuracy compared to CNN-based predictors, it is intentionally designed to cooperate with stronger predictors in multi-stage RDH-EI frameworks to further increase embedding capacity rather than to minimize prediction error alone. As demonstrated in [10], AMP serves as an effective secondary predictor that complements

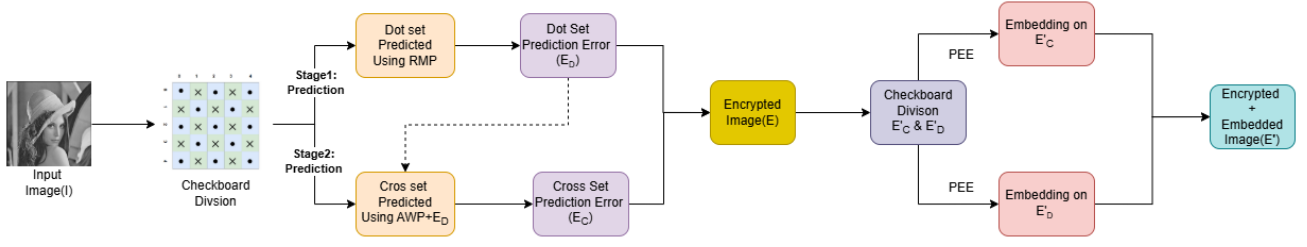


Fig. 1: The overall framework of the proposed method.

complex predictors by enabling flexible and reversible embedding. Therefore, AMP-based methods constitute a fair and representative baseline for evaluating prediction enhancement strategies in RDH-EI systems.

D. Deep Learning Approaches

Most recently, Convolutional Neural Networks (CNNs) have been applied to RDH-EI to capture non-linear spatial correlations. Ping et al. [10] introduced an Asymmetric CNN Predictor (ACNNP) to predict the first stage (dot set) with high accuracy. [11] proposed SCAM-Net, utilizing spatial and channel attention mechanisms to refine features across multiple scales. Saini and Singh [20] further integrated CNN predictors with Pixel Error Value Ordering (PEVO). While these methods achieve state-of-the-art MSE, they require millions of floating-point operations (FLOPs) per image and specialized hardware (GPUs). This creates a bottleneck for encrypted domain processing, which is often performed on resource-constrained cloud servers or edge devices.

There remains a distinct gap for a lightweight, error-aware linear predictor. Current linear methods (like AMP) are fast but "blind" to local outliers, while Multi-Predictors lose capacity to side information, and CNNs are computationally prohibitive. The proposed method fills this gap. Unlike AMP, which blindly averages neighbors, or Multi-Predictors, which require side information, our Adaptive Weighted Predictor utilizes the already available reconstruction errors from the first stage to intelligently weigh neighbors. This allows for high-accuracy prediction with the low computational cost of linear filters, offering an optimal trade-off for practical RDH-EI systems.

III. PROPOSED METHODOLOGY

In this section, we present the proposed Adaptive Weighted Predictor (AWP) designed for the RRBE framework. The overall workflow of the proposed scheme is illustrated in Fig. 1. The core innovation lies in the two-stage prediction mechanism, where the reconstruction errors generated in the first stage are utilized to guide the reliability weighting in the second stage.

A. Image Partitioning

To implement the two-stage prediction strategy, the original grayscale image I of size $H \times W$ is partitioned into two disjoint sets: the Dot set (S_D) and the Cross set (S_C),

following a checkerboard pattern, demonstrated in Fig 2 using following Eqs.

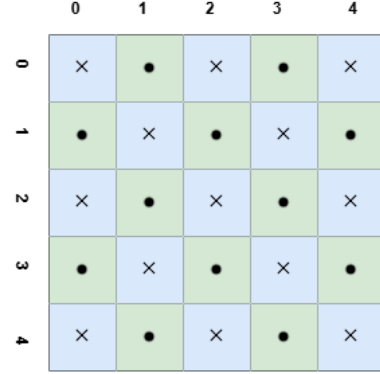


Fig. 2: Checkerboard division of a 5×5 image into dot' (•, green) and cross' (×, blue) sets, where $(i+j) \bmod 2 = 1$ for dot' and $(i+j) \bmod 2 = 0$ for cross', alternating correctly across the grid.

$$S_D = \{(i, j) \mid (i + j) \bmod 2 \neq 0\}, \quad (1)$$

$$S_C = \{(i, j) \mid (i + j) \bmod 2 = 0\}, \quad (2)$$

where $1 \leq i \leq H$ and $1 \leq j \leq W$. The pixels in S_D are predicted first to serve as reference neighbors for predicting pixels in S_C .

B. Stage 1: Dot Prediction and Error Map Generation

Since pixels in S_D do not have immediate causal neighbors available during the restoration process, we employ the Rhombus Mean Predictor (RMP) with a stride of 2. For a pixel $I(i, j) \in S_D$, the prediction value $\hat{I}_D(i, j)$ is calculated as the average of its four cardinal neighbors:

$$\hat{I}_D(i, j) = \left\lfloor \frac{I(i-2, j) + I(i+2, j) + I(i, j-2) + I(i, j+2)}{4} \right\rfloor \quad (3)$$

Once the Dot set is predicted, we calculate the Prediction Error Map(E) for the entire set S_D . This step is crucial, as E contains the "reliability history" of the local region.

$$E_D(i, j) = I(i, j) - \hat{I}_D(i, j), \quad \forall (i, j) \in S_D \quad (4)$$

For pixels belonging to S_C , the value in E is initially set to 0. This map E is stored and used as an input for the second stage.

C. Stage-2 Prediction: Cross set

After completing the first-stage prediction of the dot set, pixels in the Cross set S_C are predicted in the second stage using a weighted, error-guided predictor. This stage exploits spatial correlation among cross-set pixels while incorporating reliability information from dot-set prediction errors, thereby improving prediction accuracy without violating causality or reversibility.

Fig. 3 illustrates the prediction configuration adopted in this stage.

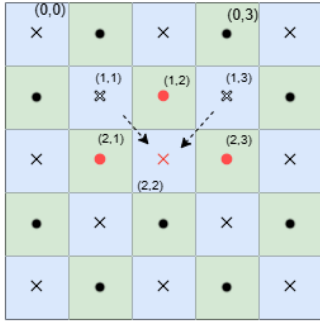


Fig. 3: Weighted average prediction for a ‘Cross’ pixel at $(2,2)$ in the top direction, using neighbors $P_1 = I_C(1,1)$, $P_2 = I_C(1,3)$ and errors $E_1 = E_D(2,1)$, $E_2 = E_D(2,3)$, $E_s = E_D(1,2)$.

1) *Prediction Context and Error Association:* Consider a target pixel $I_C(2,2) \in S_C$, as shown in Fig. 3. Two diagonal neighboring pixels belonging to the same Cross set are selected as spatial predictors:

$$P_1 = I_C(1,1), \quad P_2 = I_C(1,3).$$

The diagonal configuration ensures that the predictors are drawn from the same checkerboard set as the target pixel, thereby avoiding circular dependency and guaranteeing deterministic prediction during both embedding and recovery.

In addition to the spatial predictors, three prediction errors obtained from the error map E of the dot set are associated with the target pixel. These errors correspond to the prediction residuals generated during the first-stage rhombus prediction of the dot-set pixels. Specifically,

- $E_1 = E_D(2,1)$, associated with predictor P_1 ,
- $E_2 = E_D(2,3)$, associated with predictor P_2 ,
- $E_s = E_D(1,2)$, referred to as the shared error located between P_1 and P_2 .

Here, $E_D(i,j)$ denotes the prediction error at location (i,j) in the dot-set error map E , representing the residual obtained during the first-stage prediction.

2) *Weight Computation:* Each associated dot-set error is mapped to a reliability weight inversely proportional to its magnitude. The weights are defined as

$$w_1 = \frac{1}{1 + k|E_1|}, \quad w_2 = \frac{1}{1 + k|E_2|}, \quad w_s = \frac{1}{1 + k|E_s|}, \quad (5)$$

where k is a positive scaling constant controlling the sensitivity of the weights to prediction errors. In this work, k is fixed to 0.03. Larger prediction errors yield smaller weights, thereby reducing the influence of unreliable predictors.

3) *Weighted Prediction Model:* Using the above definitions, the predicted value of the target cross-set pixel at $(2,2)$ is computed as a normalized weighted combination of the diagonal predictors and their mean, with error compensation controlled by a damping factor λ . The prediction is expressed as

$$\hat{I}_C(2,2) = \frac{w_1 (P_1 + \lambda E_1) + w_2 (P_2 + \lambda E_2) + w_s \left(\frac{P_1 + P_2}{2} + \lambda E_s \right)}{w_1 + w_2 + w_s} \quad (6)$$

where λ regulates the contribution of dot-set errors to the prediction. In this paper, λ is set to 0.50 to balance error compensation and prediction stability.

4) *Prediction Error Computation:* After obtaining the predicted value $\hat{I}_C(2,2)$, the prediction error for the cross-set pixel is computed as

$$E_C(2,2) = I_C(2,2) - \hat{I}_C(2,2). \quad (7)$$

The prediction errors obtained from the two-stage prediction process are combined to form a unified error map $E(i,j)$, defined as

$$E(i,j) = \begin{cases} E_D(i,j), & (i,j) \in S_D, \\ E_C(i,j), & (i,j) \in S_C. \end{cases} \quad (8)$$

The resulting prediction error map $E(i,j)$ is utilized to identify reliable embedding locations based on prediction accuracy. Pixels associated with small prediction errors are considered suitable for data embedding, forming the basis for room reservation in the subsequent stage.

D. Encryption

After completing the two-stage prediction of the dot set and cross set, the original image is encrypted by the content owner to protect the image content.

A stream cipher is employed for encryption, where a pseudo-random key stream generated from the encryption key K_e is combined with the original image using a bitwise exclusive-or (XOR) operation. The encryption process is defined as

$$Ie(i,j) = I(i,j) \oplus K_e(i,j), \quad (9)$$

where $I(i,j)$ and $Ie(i,j)$ denote the pixel values of the original image and the encrypted image, respectively.

After encryption, the encrypted image I_e is delivered to the data hider, to perform data embedding according to the proposed two-stage prediction framework. The prediction error map $E(i,j)$ is retained as side information to guide embedding location selection in the subsequent data embedding stage.

IV. DATA EMBEDDING

In this section, the data hider performs secret data embedding in the encrypted image using prediction guidance obtained from the proposed two-stage framework. The embedding strategy follows the reserving room before encryption (RRBE) paradigm introduced in [7], and adopts a label-map guided bit substitution mechanism similar to [21].

The prediction error map $E(i, j)$ is utilized to identify reliable embedding locations. Pixels with small prediction errors are considered suitable for embedding, thereby defining the reserved room.

To ensure reversibility, the least significant bits (LSBs) of the selected pixels are extracted from the original image and stored as auxiliary information:

$$\mathcal{R} = \{\text{LSB}(I(i, j)) \mid (i, j) \in \Omega\}, \quad (10)$$

where Ω denotes the set of selected embeddable pixels.

After encryption, the payload bitstream $b = \{b_1, b_2, \dots, b_L\}$ (encrypted using data hiding key K_d) is embedded into the encrypted image I_e by LSB substitution:

$$I'_e(i, j) = \begin{cases} I_e(i, j) - \text{LSB}(I_e(i, j)) + b_k, & (i, j) \in \Omega, \\ I_e(i, j), & \text{otherwise.} \end{cases} \quad (11)$$

Embedding is performed sequentially to maintain consistency with the prediction structure: first on the dot set S_D , followed by the cross set S_C .

Finally, the auxiliary information $\mathcal{R}$ is compressed and embedded along with the payload to generate the marked encrypted image I'_e .

V. DATA EXTRACTION AND IMAGE RECOVERY

At the receiver side, the marked encrypted image I'_e is processed depending on the availability of the data-hiding key K_d and the encryption key K_e .

If the receiver possesses the data-hiding key K_d , the embedded payload can be extracted directly from I'_e . Since embedding modifies the least significant bit (LSB) of selected encrypted pixels, the embedded bit b_k at location (i, j) is obtained as

$$b_k = \text{LSB}(I'_e(i, j)). \quad (12)$$

After extracting the bit, the original encrypted pixel value is restored as

$$I_e(i, j) = I'_e(i, j) - b_k. \quad (13)$$

If the receiver possesses the encryption key K_e , the original image can be recovered by decrypting the restored encrypted image as

$$I_d(i, j) = I_e(i, j) \oplus K_e(i, j), \quad (14)$$

where $\oplus$ denotes the bitwise XOR operation. When both keys K_d and K_e are available, data extraction and image recovery can be performed jointly. Since embedding is carried

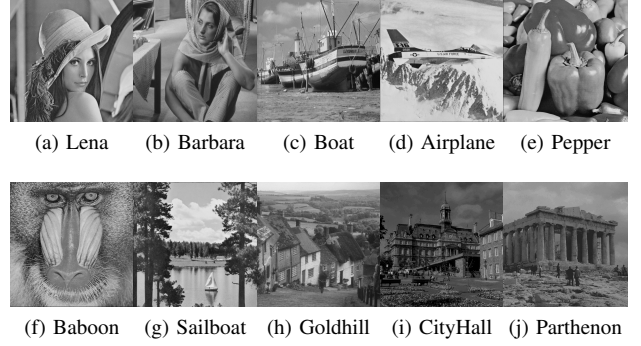


Fig. 4: Standard 512×512 grayscale test images used for evaluation.

TABLE I: Prediction performance comparison between the proposed method and the RMP+AMP baseline on 512×512 grayscale test images.

Image	RMP+AMP Baseline			Proposed Method		
	MAE	MSE	VAR	MAE	MSE	VAR
CityHall	9.814	254.036	253.966	9.679	244.034	243.834
Pepper	5.388	90.227	90.061	5.316	85.313	85.168
Parthenon	6.988	120.902	120.847	6.844	115.034	114.941
Baboon	17.503	631.599	631.421	17.241	613.566	613.371
Barbara	11.622	403.855	403.744	11.467	394.946	394.809
Lena	5.413	87.843	87.757	5.323	84.941	84.813
Sailboat	8.681	200.024	199.891	8.560	191.894	191.730
Boat	8.346	183.304	183.205	8.248	174.848	174.702
Goldhill	7.618	128.837	128.801	7.532	125.002	124.936
Airplane	5.900	130.920	130.758	5.772	124.203	124.052
Average	8.727	223.155	223.051	8.598	215.378	215.239

out in two stages, extraction is performed in the reverse order, processing cross-set pixels first followed by dot-set pixels. This guarantees correct payload extraction and perfect reconstruction of the original image.

VI. EXPERIMENTAL RESULTS

To evaluate the effectiveness of the proposed Adaptive Damped Weighted Predictor (ADWP), we conducted extensive experiments on standard grayscale test images. The proposed method was compared directly against the state-of-the-art linear baseline, the Adaptive Mean Predictor (AMP) [2], within the same RRBE framework. The RMP+AMP baseline is selected because it combines a classical local linear predictor with AMP, representing a strong non-learning-based benchmark under identical reversibility and encryption constraints.

A. Experimental Setup

The experiments were conducted on a dataset of standard 512×512 grayscale images, including *Airplane*, *Lena*, *Boat*, *Pepper*, *Baboon*, *Barbara*, *Sailboat*, *Goldhill*, *CityHall* and *Parthenon*. as shown in Fig. 4. These images exhibit diverse texture characteristics, ranging from smooth regions (*Airplane*) to complex high-frequency textures (*Baboon*).

B. Quantitative Analysis

Table I compares the prediction performance of the proposed method with the RMP+AMP baseline. The proposed

predictor consistently achieves lower MAE, MSE, and variance across all test images. On average, the proposed method reduces MSE and prediction error variance by approximately 4%, indicating a more compact prediction-error distribution that is beneficial for RRBE-based reversible data hiding. As observed in Table I, the proposed method achieves a consistent reduction in prediction error across all test images.

Table II presents the sensitivity of the proposed predictor to different values of the weighting parameter k and damping factor λ . It can be observed that moderate values of k combined with $\lambda = 0.50$ achieve the lowest prediction error across all metrics, indicating a balanced trade-off between error compensation and prediction stability.

TABLE II: Parameter sensitivity analysis of the proposed predictor for different values of k and damping factor λ

k	λ	MSE	MAE	Variance
0.03	0.50	215.378	8.598	215.239
0.03	0.75	216.027	8.624	215.902
0.05	0.75	216.687	8.684	216.565
0.07	0.75	217.362	8.653	217.134
0.05	0.90	218.223	8.682	218.106

VII. CONCLUSION

This paper introduced an adaptive weighted predictor for RRBE-based reversible data hiding in encrypted images. By utilizing reconstruction errors from the first prediction stage to guide neighbor weighting in the second stage, the proposed method effectively suppresses unreliable contexts in textured regions. A damping mechanism is further incorporated to prevent noise amplification in smooth areas, ensuring stable and reversible prediction. Experimental results on standard grayscale images show that the proposed predictor consistently outperforms the RMP+AMP baseline, achieving average reductions of 1.48% in MAE and approximately 3.50% in both MSE and prediction error variance. These improvements lead to a more compact prediction error distribution while maintaining constant-time computational complexity, making the method suitable for practical cloud-based RDH-EI systems. Future work will focus on extending this error-aware mechanism to color images by exploiting inter-channel correlations and exploring its application in 3D mesh watermarking.

REFERENCES

- [1] X. Li, J. Li, B. Li, and B. Yang, "High-fidelity reversible data hiding scheme based on pixel-value-ordering and prediction-error expansion," *Signal Processing*, vol. 93, no. 1, pp. 198–205, 2013.
- [2] 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.
- [3] 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. 13445–13457, 2018.
- [4] H. Honsinger, P. Jones, M. Rabbani, and J. Stoffel, "Lossless recovery of an original image containing embedded data," U.S. Patent 6 278 791, Aug. 2001.
- [5] X. Zhang, "Reversible data hiding in encrypted image," *IEEE Signal Processing Letters*, vol. 18, no. 4, pp. 255–258, 2011.

- [6] X. Zhang, "Separable reversible data hiding in encrypted image," *IEEE Transactions on Information Forensics and Security*, vol. 7, no. 2, pp. 826–832, 2012.
- [7] 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.
- [8] 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.
- [9] Y. Qiu, K. Chen, X. Lin, G. Li, and H. Zeng, "High-capacity reversible data hiding in encrypted images with edge-directed prediction and adaptive entropy coding," *IEEE Signal Processing Letters*, vol. 32, pp. 3012–3016, 2025.
- [10] 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.
- [11] S. Gandhi and R. Kumar, "SCAM-Net: Spatial-channel attention multi-scale network for reversible data hiding in encrypted images," in *Proc. IEEE International Conference*, 2024.
- [12] S. Gandhi and R. Kumar, "Survey of reversible data hiding: Statistics, current trends, and future outlook," *Computer Standards & Interfaces*, vol. 94, 2025.
- [13] R. Kumar and A. K. Sharma, "A review of different prediction methods for reversible data hiding," *Journal of Information Security and Applications*, vol. 78, 2023.
- [14] H. Ren, G. Bai, T. Chen, Z. Yue, and R. Ren, "Reversible data hiding in encrypted images with multi-prediction and adaptive Huffman encoding," *Scientific Reports*, vol. 13, p. 23104, 2023.
- [15] F. Ren, Z. Wu, Y. Xue, and Y. Hao, "Reversible data hiding in encrypted image based on bit-plane redundancy of prediction error," *Mathematics*, vol. 11, 2023.
- [16] S. Yi and Y. Zhou, "Reversible data hiding in encrypted images using prediction-error encoding," in *Proc. IEEE SMC*, 2018.
- [17] P. S. Sandhu, A. Malik, S. Singh, and V. Tyagi, "An innovative interpolation based reversible steganographic technique," *Procedia Computer Science*, vol. 259, pp. 1855–1862, 2025.
- [18] I. C. Dragoi and D. Coltuc, "Reversible data hiding in encrypted images based on reserving room after encryption and multiple predictors," in *Proc. IEEE ICASSP*, 2018.
- [19] H. Zhang, L. Li, and Q. Li, "Reversible data hiding in encrypted images based on block-wise multi-predictor," *IEEE Access*, vol. 9, 2021.
- [20] A. K. Saini and S. Singh, "Reversible data hiding using pixel error value ordering and new CNN predictor," *Signal, Image and Video Processing*, vol. 19, 2025.
- [21] Ankur, R. Kumar, A. K. Sharma, and P. Ranjan, "High-capacity reversible data hiding in encrypted images based on difference image transfiguration," *Signal, Image and Video Processing*, vol. 19, no. 5, p. 365, 2025.