

Attention-Enhanced Deep Neural Predictor for High-Capacity RDHEI

A THESIS REPORT

SUBMITTED IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR THE AWARD OF THE DEGREE

OF

**MASTER OF TECHNOLOGY
IN
ARTIFICIAL INTELLIGENCE**

Submitted by

**ADITYA KUMAR CHAUDHARY
(24/AFI/18)**

Under the supervision of

DR. RAJEEV KUMAR



**DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING
DELHI TECHNOLOGICAL UNIVERSITY**

(Formerly Delhi College of Engineering)

Bawana Road, Delhi-110042

MAY 2026

CANDIDATE’S DECLARATION



DELHI TECHNOLOGICAL UNIVERSITY

(Formerly Delhi College of Engineering)

Shahbad Daulatpur, Main Bawana Road, Delhi – 110042

I, **Aditya Kumar Chaudhary**, student of M.Tech. (Artificial Intelligence), Department of Computer Science and Engineering, Delhi Technological University, hereby declares that the thesis report entitled “**Attention-Enhanced Deep Neural Predictor for High-Capacity RDHEI in Grayscale Images**” submitted in partial fulfilment of the requirements for the award of the degree of Master of Technology is an original work carried out under the supervision of **Dr. Rajeev Kumar**. The work presented in this report has not been submitted elsewhere for the award of any other degree, diploma, fellowship, or similar title.

The explanations, interpretations, and thesis-level organisation presented in this report are based on the supplied conference work and the referenced base papers. Any external work used for technical comparison or contextual discussion has been acknowledged through IEEE-style citations.

Candidate’s Signature

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



DELHI TECHNOLOGICAL UNIVERSITY

(Formerly Delhi College of Engineering)

Shahbad Daulatpur, Main Bawana Road, Delhi – 110042

This is to certify that the thesis report entitled “**Attention-Enhanced Deep Neural Predictor for High-Capacity RDHEI in Grayscale Images**” submitted by **Aditya Kumar Chaudhary** in partial fulfillment of the requirements for the award of the degree of **Master of Technology in Artificial Intelligence** is a bonafide record of work carried out under my supervision at Delhi Technological University.

The report focuses on grayscale-image-based Reversible Data Hiding in Encrypted Images (RDHEI) using a CBAM-enhanced convolutional neural predictor combined with an Adaptive Mean Predictor (AMP). The work demonstrates an understanding of secure multimedia communication, reversible data hiding, image prediction, attention-based deep learning, encrypted-domain embedding, and lossless recovery requirements.

Place: Delhi

Date: 30.05.2026

Dr. Rajeev Kumar

Supervisor

ACKNOWLEDGEMENT



DELHI TECHNOLOGICAL UNIVERSITY

(Formerly Delhi College of Engineering)

Shahbad Daulatpur, Main Bawana Road, Delhi – 110042

I express my sincere gratitude to my supervisor, **Dr. Rajeev Kumar**, for his valuable guidance, continuous encouragement, and technical support throughout the development of this thesis report. His expertise in deep learning, image processing, and secure multimedia communication has been instrumental in shaping the direction of this work.

I am also thankful to the Department of Computer Science and Engineering, Delhi Technological University, for providing an academic environment that supports research in artificial intelligence, multimedia security, and privacy-preserving computation. I acknowledge the authors of the base papers on asymmetric CNN-based prediction and high-precision colour image prediction, whose contributions helped motivate the design of an attention-enhanced predictor for grayscale RDHEI.

I would also like to thank my friends, colleagues, and family members for their support and encouragement during the completion of this work.

Aditya Kumar Chaudhary

ABSTRACT

The growing adoption of cloud computing, telemedicine, military communication, forensic imaging, and big-scale multimedia storage has really pushed the need for secure and reverse information embedding mechanisms. Reversible Data Hiding in Encrypted Images (RDHEI) basically lets auxiliary data be embedded into encrypted images, but it also makes sure you can do exact recovery for both the hidden data AND the original image. That kind of exactness matters a lot in sensitive areas where any distortion in the image is basically not allowed, and confidentiality needs to stay protected during transmission, as well as storage.

In this thesis report, we introduce an attention-enhanced deep neural prediction framework aimed at high-capacity RDHEI in grayscale images. The whole idea is focused on a Convolutional Block Attention Module (CBAM)-enhanced convolutional neural network predictor that is combined with an Adaptive Mean Predictor (AMP). The proposal is inspired by earlier asymmetric CNN-based and adaptive mean predictor approaches for RDHEI, in which a two-stage prediction strategy is used along with multi-MSB substitution to raise the embedding capacity. It is also guided by high-precision CNN prediction studies that indicate attention mechanisms can improve prediction accuracy. Compared to traditional local linear predictors and earlier CNN predictors that only provide a small amount of adaptive refinement when handling extracted features, the proposed CBAM-enhanced predictor tries to do better by applying both channel attention and spatial attention. The goal here is to highlight informative grayscale texture features and also bring out the important spatial structures more clearly.

The framework you proposed really does a two-stage prediction, plus embedding, thing. First, you take a grayscale image and split it into Circle and Triangle sets, using a kind of checkerboard-style partitioning, which is a bit like alternating blocks. Then the Circle set becomes the contextual input to predict the Triangle set, with a CBAM-enhanced CNN predictor, or so it goes. Next, AMP steps in to predict the Circle set, via adaptive directional neighbourhood estimation, pretty much like it “looks around” in a tailored way. After that, the prediction error distributions from these stages are turned into embeddable locations, label maps, and auxiliary information. Those pieces then get used for reversible data hiding. The actual embedding process leans on multi-MSB substitution, label-map management, secure encryption, and separable data extraction plus image recovery. In the thesis, you’ll

see the math setup for CBAM attention, how prediction errors are generated, how label maps are built, plus encryption, embedding, and the recovery procedures.

So the proposed framework is supposed to be used for evaluation on the usual grayscale benchmark datasets like BOSSBase, BOWS-2, and UCID. For the comparison, it uses metrics such as Mean Squared Error (MSE), Mean Absolute Error (MAE), prediction error variance, Embedding Rate (ER), Peak Signal-to-Noise Ratio (PSNR), Structural Similarity Index Measure (SSIM), and also security metrics like entropy and correlation.

Since we cannot just make up any additional unreported numerical experimental results, this report uses placeholders for the final CBAM-related result tables and graphs, even though we still show the baseline values that are already given in the reference literature. In short, the proposed grayscale RDHEI framework offers a high-capacity, kind of secure, and reversible solution. It is aimed at things like medical imaging, forensic communication, encrypted cloud storage, and privacy-preserving multimedia systems.

Contents

CANDIDATE’S DECLARATION	i
Candidate’s Declaration	i
CERTIFICATE	ii
Certificate	ii
ACKNOWLEDGEMEN	iii
Acknowledgement	iii
Abstract	v
Contents	viii
List of Figures	ix
List of Tables	x
List of Abbreviations	xii
1 INTRODUCTION	1
1.1 Background and Motivation	1
1.2 Problem Domain Overview	2
1.3 Existing Challenges	2
1.4 Problem Statement	3
1.5 Research Objectives	3
1.6 Scope of the Research	3
1.7 Applications	4
1.8 Research Contributions	4
1.9 Thesis Organization	5
2 LITERATURE REVIEW	6
2.1 Overview	6
2.2 Traditional Reversible Data Hiding	6

2.3	Reversible Data Hiding in Encrypted Images	6
2.4	Prediction-Based RDHEI	7
2.5	Asymmetric CNN-Based and Adaptive Mean Predictors	7
2.6	High-Precision CNN Prediction for Grayscale Images	9
2.7	Attention Mechanisms in Deep Learning	10
2.8	Comparative Analysis of Existing Methods	11
2.9	Research Gaps	11
2.10	Summary	12
3	SYSTEM MODEL AND PRELIMINARIES	13
3.1	RDHEI System Model	13
3.2	Grayscale Image Representation	13
3.3	Checkerboard-Based Image Division	13
3.4	Prediction Error	14
3.5	Encryption Model	14
3.6	Multi-MSB Substitution	15
3.7	Label Map and Auxiliary Information	15
3.8	Performance Metrics	15
3.8.1	Mean Square Error	15
3.8.2	Mean Absolute Error	15
3.8.3	Embedding Rate	16
3.8.4	Peak Signal-to-Noise Ratio	16
3.8.5	Structural Similarity Index	16
3.9	Summary	16
4	PROPOSED METHODOLOGY	17
4.1	Chapter Overview	17
4.2	Overall Framework	17
4.3	Grayscale Image Division	17
4.4	CBAM-Enhanced CNN Predictor	18
4.4.1	Multi-Scale Feature Extraction	18
4.4.2	CBAM Attention Refinement	19
4.4.3	Channel Attention	19
4.4.4	Spatial Attention	19
4.4.5	Triangle Set Prediction	20
4.5	Adaptive Mean Predictor	20
4.6	Label Map Generation	21
4.7	Image Encryption	21
4.8	Data Embedding	22
4.9	Data Extraction and Image Recovery	22

4.10	Algorithmic Description	22
4.11	Implementation-Specific Architecture	23
4.12	Summary	23
5	RESULTS ANALYSIS AND DISCUSSION	24
5.1	Chapter Overview	24
5.2	Experimental Setup	24
5.3	Dataset Description	25
5.4	Evaluation Metrics	25
5.5	Baseline Results Available from Reference Papers	26
5.6	Qualitative Discussion of Expected Behavior	26
5.7	Comparison with ACNNP and AMP	26
5.8	Security Analysis	27
5.9	Complexity Discussion	28
5.10	Summary	29
6	CONCLUSION AND FUTURE SCOPE	31
6.1	Conclusion	31
6.2	Major Findings	31
6.3	Limitations	32
6.4	Future Scope	32
6.5	Closing Remarks	33
A	IMPLEMENTATION NOTES	34
A.1	Implementation Notes	34
A.2	Suggested Code Organization	34

List of Figures

2.1	Application scenario of RDHEI in a medical expert-system and cloud-storage workflow, adapted from the ACNNP-based RDHEI base paper [2].	7
2.2	Two-stage ACNNP and AMP framework for RDHEI, used as the main architectural base for the proposed thesis work [2].	8
2.3	Architecture of the asymmetric CNN-based predictor used in the reference RDHEI framework [2].	8
2.4	Asymmetric convolutional block used for multi-scale feature extraction in the ACNNP baseline [2].	8
2.5	Grayscale image division strategies are discussed in the high-precision CNN predictor base paper [3]. This thesis adapts the general idea of structured partitioning to grayscale images.	9
2.6	Overview of an improved CNN predictor with channel attention from the grayscale-image RDH.	10
3.1	Three-party RDHEI system model followed in the proposed thesis work. . .	13
3.2	Checkerboard-style grayscale image partitioning into Circle (C) and Triangle (T) sets.	14
4.1	Overall workflow of the proposed CBAM-enhanced grayscale RDHEI framework.	17
4.2	CBAM module used in the proposed grayscale predictor	19
4.3	Adaptive mean prediction modes used for local directional estimation in the baseline framework [2].	21
4.4	CBAM-AMP Architectutre	23
5.1	Representative security and histogram visualization from the ACNNP-based reference framework [2]. Similar plots should be generated for the proposed CBAM + AMP framework.	28
5.2	Histogram compasrison	29
5.3	CBAM + AMP RDHEI framework on the UCID	30

List of Tables

2.1	Comparative overview of relevant prediction-based methods.	11
5.1	Suggested experimental configuration for final implementation.	25
5.2	Reported baseline average embedding rates from ACNNP + AMP reference work.	26
5.3	Template for comparing prediction accuracy. Final CBAM values should be inserted after implementation.	27
5.4	Template for comparing embedding and reconstruction performance.	27
5.5	Expected complexity comparison.	29

LIST OF ABBREVIATIONS

ACNNP	Asymmetric Convolutional Neural Network Predictor
AMP	Adaptive Mean Predictor
AI	Artificial Intelligence
BOWS	Break Our Watermarking System
BPP	Bits Per Pixel
CBAM	Convolutional Block Attention Module
CNN	Convolutional Neural Network
DE	Difference Expansion
DH	Data Hiding
ECA	Efficient Channel Attention
ER	Embedding Rate
GAP	Gradient-Adjusted Predictor
HS	Histogram Shifting
LSB	Least Significant Bit
MAE	Mean Absolute Error
MED	Median Edge Detector
MSE	Mean Square Error
MSB	Most Significant Bit
PEE	Prediction Error Expansion
PSNR	Peak Signal-to-Noise Ratio
PVO	Pixel Value Ordering
RDH	Reversible Data Hiding
RDHEI	Reversible Data Hiding in Encrypted Images
RRBE	Reserving Room Before Encryption
SSIM	Structural Similarity Index Measure
UCID	Uncompressed Colour Image Database
VRAE	Vacating Room After Encryption

Chapter 1

INTRODUCTION

1.1 Background and Motivation

Digital pictures are now routinely used for medical diagnosis, military comms, surveillance, remote sensing, forensic work, cloud storage, and intelligent decision support. In all these uses, the image itself usually holds sensitive content, and it may also need to carry some extra stuff like authentication marks, diagnostic notes, patient metadata, proprietor identifiers, or integrity verification codes. Regular encryption keeps the visual data private, but it does not really give a clean way to “hide” extra information, especially after the encryption step. On the other hand, classic data hiding can do auxiliary embedding, yet a lot of those schemes make permanent changes to the cover image, so they are not great for scenarios where you must recover the image exactly.

Reversible Data Hiding, or RDH, deals with this problem by stuffing secret or side information into an image in a way that the original image can be rebuilt perfectly after extraction. That exactness matters a lot when even tiny distortion is not acceptable, like in medical imaging, court or legal evidence, military reconnaissance, and scientific imaging. Reversible Data Hiding in Encrypted Images, abbreviated RDHEI, pushes this further by enabling data embedding while the image remains encrypted. So it ends up meeting three major expectations at once: confidentiality of the image, insertion of auxiliary data, and lossless, no-error recovery after everything is taken out.

The current thesis looks at high-capacity RDHEI for grayscale images. It specifically examines how to improve prediction accuracy by using a deep neural predictor with attention. In prediction-based RDHEI, when the predicted pixel values are similar to the actual values, the prediction error histogram tends to cluster around zero. A tighter prediction error distribution allows for better management of embeddable space and increases the achievable embedding rate. This encourages the use of convolutional neural networks and attention mechanisms for pixel prediction.

1.2 Problem Domain Overview

RDHEI methods fall into two main categories: Vacating Room After Encryption (VRAE) and Reserving Room Before Encryption (RRBE). VRAE methods first encrypt the original image and then create space within the encrypted domain. However, strong encryption eliminates much of the spatial redundancy. This makes it hard to find reliable spots for embedding. In contrast, RRBE methods set aside embedding space in the plaintext image before encryption. They typically do this by using spatial correlation and prediction errors. The reserved space stays available after encryption, which usually results in a higher embedding capacity.

The proposed thesis work is part of the RRBE group. It includes grayscale image partitioning, two-stage prediction, label-map generation, encryption, multi-MSB substitution, and separable recovery. The first stage utilizes a CNN predictor enhanced by CBAM to estimate one group of pixels from another. The second stage employs an adaptive mean predictor to make better use of the local neighborhood structure. This combination aims to keep the high capacity of deep prediction-based RDHEI while boosting feature selectivity through attention.

1.3 Existing Challenges

Although many RDHEI methods have achieved promising results, several challenges remain. First, local linear predictors such as MED and GAP have low computational complexity but cannot capture complex grayscale textures and nonlinear spatial relationships. Second, CNN-based predictors improve the learning of contextual structures, but a plain CNN may still treat all channels and spatial regions uniformly. Third, attention mechanisms can refine feature representations, yet their integration into grayscale RDHEI prediction must preserve reversibility and should not create excessive side information. Fourth, embedding capacity must be improved without compromising image recovery, secret-data extraction, and encryption security.

The asymmetric CNN-based and adaptive mean predictor framework demonstrates the effectiveness of two-stage prediction in RDHEI, with reported average embedding rates of 3.302 bpp on BOWS-2, 3.461 bpp on BOSSBase, and 2.69 bpp on UCID [2]. However, the feature extraction process in such a framework can be further improved by attention-based refinement. A related high-precision CNN predictor for color images introduces channel attention to better exploit cross-channel relationships [3]. These observations motivate the proposed CBAM-enhanced grayscale predictor.

1.4 Problem Statement

The central problem addressed in this thesis is the design of a high-precision deep neural predictor for grayscale-image RDHEI that can improve prediction reliability and embedding capacity while preserving exact reversibility and security. The specific research question is:

How can a CBAM-enhanced CNN predictor be integrated with a two-stage RDHEI framework to obtain sharper grayscale prediction error distributions, higher embedding capacity, and reliable lossless recovery in encrypted images?

1.5 Research Objectives

The main objectives of this thesis are as follows:

1. To study the role of image prediction in high-capacity RDHEI and identify the limitations of traditional linear predictors and plain CNN predictors.
2. To analyze the asymmetric CNN-based predictor and adaptive mean predictor as a baseline two-stage RDHEI design.
3. To design a CBAM-enhanced CNN predictor for grayscale images by integrating channel attention and spatial attention into the prediction pipeline.
4. To formulate a complete reversible embedding and recovery procedure using image partitioning, label-map construction, encryption, multi-MSB substitution, data extraction, and image recovery.
5. To define evaluation protocols using prediction error metrics, embedding capacity metrics, reconstruction metrics, and security metrics.
6. To provide a thesis-level methodology that can be extended into final experimental implementation and validation.

1.6 Scope of the Research

This thesis is focused on grayscale images and does not directly address color image embedding or inter-channel correlation in the proposed framework. Color-image CNN predictors are discussed in the literature review because they motivate attention-based prediction, but the proposed method operates on grayscale image sets. The intended evaluation datasets

include BOSSBase, BOWS-2, and UCID after grayscale conversion or grayscale subset preparation. The scope also includes theoretical methodology, algorithmic formulation, and placeholders for final experimental result insertion.

1.7 Applications

The proposed RDHEI framework is applicable in environments where confidential images must be stored or transmitted with embedded auxiliary information. Important applications include:

- medical image management, where clinical notes or authentication tags must be embedded without altering diagnostic content;
- forensic imaging, where evidence integrity and reversibility are essential;
- cloud-based image storage, where encrypted images may need indexing or metadata embedding by a third party;
- military and surveillance communication, where both secrecy and lossless recovery are required;
- digital copyright protection and ownership verification, where watermark information must be embedded reversibly.

1.8 Research Contributions

The major contributions of the thesis are summarized below.

1. A thesis-level RDHEI framework is developed for grayscale images using a CBAM-enhanced CNN predictor and an adaptive mean predictor.
2. The proposed predictor enhances the baseline ACNNP idea by introducing sequential channel and spatial attention for adaptive grayscale feature refinement.
3. A two-stage prediction, encryption, embedding, extraction, and recovery workflow is organized in a reversible and separable manner.
4. Mathematical formulations for image partitioning, CBAM attention, prediction error generation, label-map construction, encryption, and recovery are provided.
5. A complete LaTeX thesis report structure is prepared in the DTU style with figures, algorithms, tables, placeholders, and references.

1.9 Thesis Organization

The remainder of this thesis is organized as follows. Chapter 2 presents the literature review of RDH, RDHEI, prediction-based embedding, CNN predictors, and attention mechanisms. Chapter 3 describes the system model and preliminary concepts required for understanding the proposed framework. Chapter 4 explains the proposed CBAM-enhanced RDHEI methodology in detail. Chapter 5 presents the experimental design, performance metrics, comparative discussion, and placeholders for final results. Chapter 6 concludes the thesis and outlines future research directions.

Chapter 2

LITERATURE REVIEW

2.1 Overview

The literature relevant to this thesis spans reversible data hiding, reversible data hiding in encrypted images, prediction-based embedding, deep CNN-based prediction, and attention mechanisms. Since the proposed work combines a CBAM-enhanced CNN predictor with an adaptive mean predictor for grayscale RDHEI, this chapter reviews both classical reversible data hiding approaches and recent deep predictor-based methods.

2.2 Traditional Reversible Data Hiding

RDH was introduced to overcome the irreversibility of conventional data hiding. In early RDH, difference expansion modifies the difference between neighboring pixels to embed data while preserving enough information for recovery [7, 8]. Histogram shifting shifts pixel values or prediction errors around histogram peaks and zero bins to create embedding room [10]. Prediction error expansion improves embedding efficiency by predicting pixels and embedding data into the resulting prediction errors [9]. Pixel Value Ordering (PVO) arranges pixels inside a block and uses the ordered structure to embed data with low distortion [11]. These methods are effective in plaintext images, but they do not directly address privacy protection.

2.3 Reversible Data Hiding in Encrypted Images

RDHEI was proposed to allow data hiding after image encryption or in a privacy-preserving workflow. Zhang's early encrypted-domain method demonstrated that reversible data hiding can be performed in encrypted images [4]. Later work introduced separable data extraction and image recovery, allowing the receiver to extract data with a data-hiding key and recover the image with an image-decryption key [5]. Ma et al. proposed the RRBE

paradigm, which reserves room before encryption and thereby improves embedding capacity [6].

In VRAE methods, embedding room is vacated after encryption. This is difficult because encryption increases entropy and destroys spatial correlation. In RRBE methods, the content owner preprocesses the image before encryption to create space for data embedding. Since preprocessing operates on the plaintext image, it can exploit spatial redundancy and prediction errors more effectively.

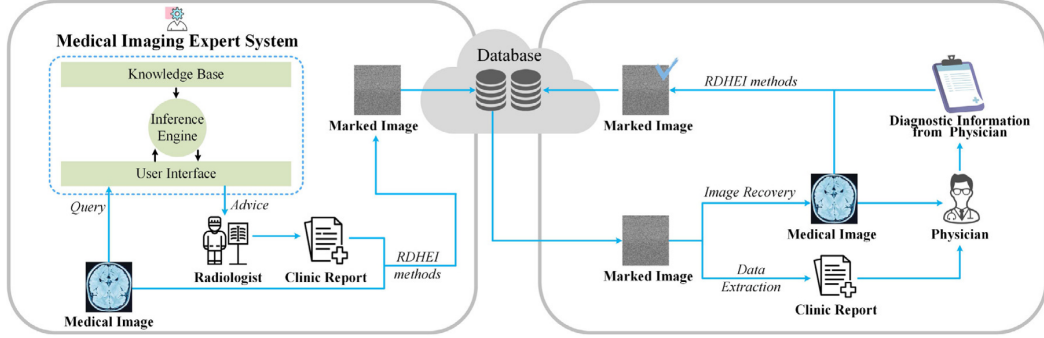


Figure 2.1: Application scenario of RDHEI in a medical expert-system and cloud-storage workflow, adapted from the ACNNP-based RDHEI base paper [2].

2.4 Prediction-Based RDHEI

Prediction-based RDHEI reserves room by estimating pixel values from their context. If the prediction is accurate, most prediction errors lie near zero. This improves label-map compression and increases the embeddable space. Linear predictors such as MED and GAP are simple but limited. They use fixed assumptions about local gradients and may perform poorly in highly textured regions.

Deep learning-based predictors overcome some of these limitations by learning nonlinear relationships from image data. However, deep predictors must be used carefully in RDHEI because the method must preserve deterministic prediction, embeddability, and reversibility. The ACNNP and AMP framework proposed by Ping et al. is important because it combines a learned asymmetric CNN predictor with a deterministic adaptive mean predictor in a two-stage RDHEI system [2].

2.5 Asymmetric CNN-Based and Adaptive Mean Predictors

The ACNNP framework divides a grayscale image into two complementary sets, called Circle and Triangle sets. The Circle set is used to predict the Triangle set using an asymmetric

CNN predictor, while an adaptive mean predictor is used to predict the Circle set in a second stage. The reference framework includes prediction, label-map generation, label-map embedding, encryption, data hiding, extraction, and lossless recovery.

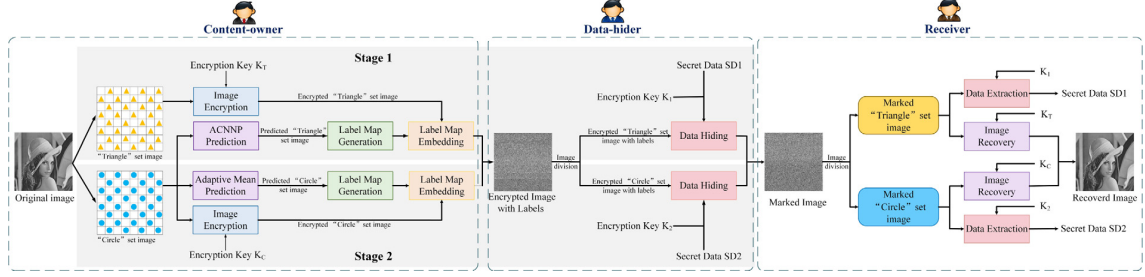


Figure 2.2: Two-stage ACNNP and AMP framework for RDHEI, used as the main architectural base for the proposed thesis work [2].

The asymmetric CNN predictor uses multiple branches with different kernel shapes and sizes. This enables the network to capture spatial information in horizontal, vertical, and square receptive fields. The adaptive mean predictor then selects a suitable direction for predicting remaining pixels. These two predictors together improve flexibility and embedding capacity.

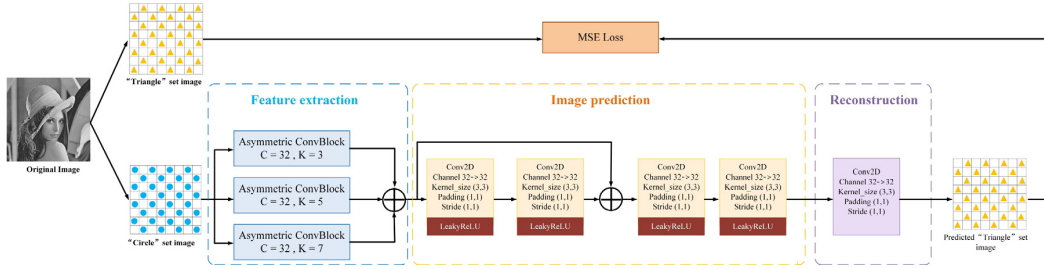


Figure 2.3: Architecture of the asymmetric CNN-based predictor used in the reference RDHEI framework [2].

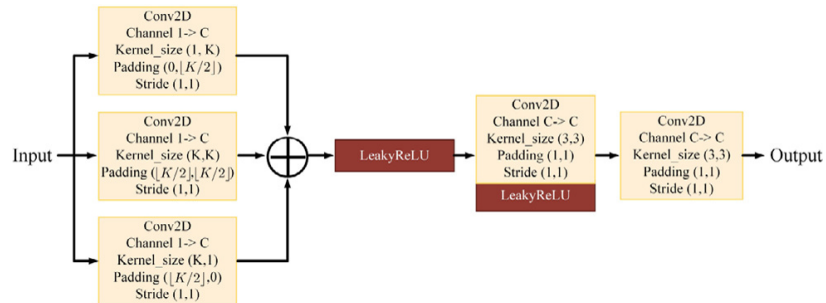


Figure 2.4: Asymmetric convolutional block used for multi-scale feature extraction in the ACNNP baseline [2].

2.6 High-Precision CNN Prediction for Grayscale Images

Although the proposed thesis focuses on grayscale RDHEI, research on color-image prediction is useful because it demonstrates the importance of richer context and attention mechanisms. The high-precision CNN predictor for color images divides the three color channels into multiple independent parts and incorporates an Efficient Channel Attention (ECA) mechanism to model relationships among channels [3]. This suggests that attention can improve prediction when feature importance varies across channels or spatial regions.

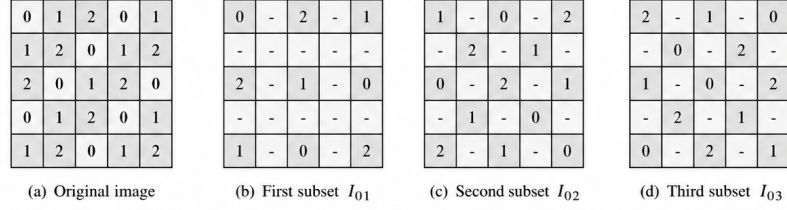


Fig. 1. Illustration of dividing a grayscale image into three independent parts (I_{01}, I_{02}, I_{03}).

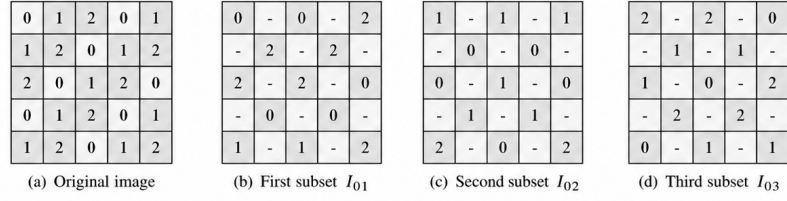


Fig. 2. Illustration of I_{02} in another division pattern.

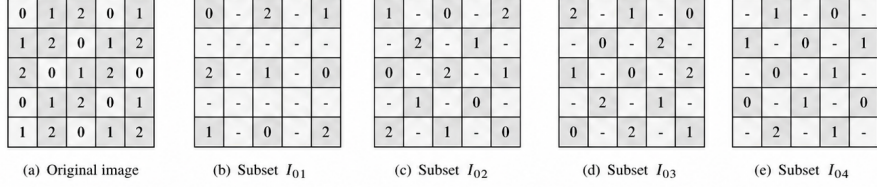


Fig. 3. Illustration of dividing a grayscale image into four independent parts ($I_{01}, I_{02}, I_{03}, I_{04}$).

Figure 2.5: Grayscale image division strategies are discussed in the high-precision CNN predictor base paper [3]. This thesis adapts the general idea of structured partitioning to grayscale images.

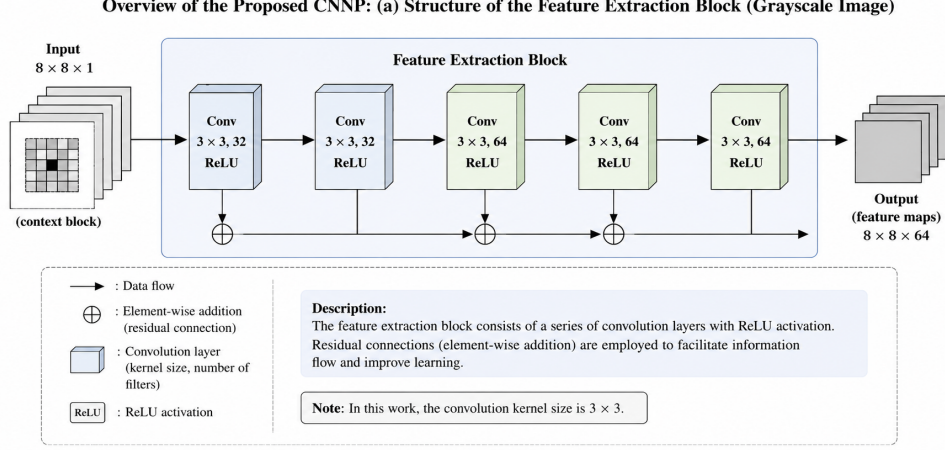


Figure 2.6: Overview of an improved CNN predictor with channel attention from the grayscale-image RDH.

2.7 Attention Mechanisms in Deep Learning

Attention mechanisms allow neural networks to emphasize useful features and suppress irrelevant responses. SENet introduced channel-wise recalibration by learning channel importance weights [16]. ECA simplified channel attention by avoiding dimensionality reduction and using efficient local cross-channel interaction [17]. CBAM extends this idea by applying channel attention followed by spatial attention [15]. The channel module identifies *what* features are important, while the spatial module identifies *where* they are important.

For grayscale image prediction, CBAM is attractive because it can refine feature maps without changing the fundamental reversible embedding workflow. The channel attention part can emphasize useful learned feature channels, while spatial attention can highlight edges, texture boundaries, and luminance transitions. This motivates replacing the plain feature refinement stage of ACNNP with a CBAM-enhanced block.

2.8 Comparative Analysis of Existing Methods

Table 2.1: Comparative overview of relevant prediction-based methods.

Method/Category	Main Idea	Strength	Limitation for Proposed Problem
MED/GAP pre-predictors	Local linear prediction	Low complexity and deterministic	Limited modeling of complex grayscale texture
PVO/PEE methods	Use ordered pixels or prediction errors	Good RDH performance in plaintext images	Privacy protection requires additional encryption framework
RRBE-RDHEI	Reserve room before encryption	High embedding capacity compared with VRAE	Requires reliable preprocessing and label-map design
ACNNP + AMP	CNN-based first-stage prediction and adaptive mean second-stage prediction	Strong RDHEI baseline with reported high average embedding rates	Feature refinement is not explicitly attention-guided
Proposed CBAM + AMP	Grayscale CBAM-enhanced CNN with adaptive mean prediction	Attention-guided texture and spatial feature refinement	Requires final experimental validation and parameter tuning

2.9 Research Gaps

- Existing simple and linear predictors cannot meet the requirements of high-ability reversible data hiding in encrypted grayscale images (RDHEI) due to low prediction accuracy.
- Although the Asymmetric CNN-based Predictor (ACNNP) has improved the prediction accuracy, it does not contain explicit attention mechanisms for enhancing informative feature representations.

- Attention-based CNN predictors applied to the reversible data hiding (RDH) in color images have demonstrated the effectiveness of channel attention mechanisms such as Efficient Channel Attention (ECA)/Convolutional Block Attention Module (CBAM), but not optimized for grayscale encrypted-image prediction.
- Most of the existing RDHEI methods study one of the prediction or embedding performances, while the end-to-end system of grayscale RDHEI, containing prediction, encryption, embedding, extraction, image recovery, and assessment, has yet to be explored.
- The existing RDHEI predictors are incapable of adaptively learning the spatial and channel correlations for increasing embedding ability and prediction precision in the grayscale RDHEI.
- There is a research gap to develop a thesis-level CNN predictor enhanced with the CBAM for RDHEI that can improve the embedding ability, improve the prediction accuracy and provide reversible image recovery for secure grayscale-image communications.

2.10 Summary

This chapter reviewed RDH, RDHEI, prediction-based embedding, ACNNP-based RDHEI, color-image CNN prediction, and attention mechanisms. The review shows that high-capacity RDHEI strongly depends on prediction accuracy and that attention-enhanced deep predictors can provide a promising direction for improving grayscale prediction reliability. The next chapter introduces the system model and preliminary concepts used by the proposed framework.

Chapter 3

SYSTEM MODEL AND PRELIMINARIES

3.1 RDHEI System Model

The proposed work follows a three-party RDHEI model consisting of a content owner, a data hider, and a receiver. The content owner possesses the original grayscale image and performs preprocessing and encryption. The data hider receives the encrypted image with reserved embedding room and embeds secret data without accessing the plaintext image. The receiver extracts the embedded data, recovers the image, or performs both operations depending on key availability.

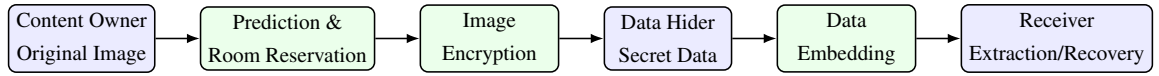


Figure 3.1: Three-party RDHEI system model followed in the proposed thesis work.

3.2 Grayscale Image Representation

Let I denote an 8-bit grayscale image of size $M \times N$. Each pixel value $I(i, j)$ lies in the range $[0, 255]$, where $1 \leq i \leq M$ and $1 \leq j \leq N$. In prediction-based RDHEI, the goal is to predict a target pixel using a known context and then use the prediction error to determine embeddable locations.

3.3 Checkerboard-Based Image Division

The supplied CBAM work uses grayscale image division into two complementary sets. This is consistent with the Circle/Triangle division idea used in the ACNNP baseline. Let I^C denote the Circle set and I^T denote the Triangle set. A simple checkerboard partition may be written as

$$I^C = \{I(i, j) \mid (i + j) \bmod 2 = 0\}, \quad (3.1)$$

$$I^T = \{I(i, j) \mid (i + j) \bmod 2 = 1\}. \quad (3.2)$$

The two sets are disjoint and together reconstruct the original image:

$$I = I^C + I^T. \quad (3.3)$$

<i>T</i>	<i>C</i>	<i>T</i>	<i>C</i>	<i>T</i>	<i>C</i>	<i>T</i>	<i>C</i>
<i>C</i>	<i>T</i>	<i>C</i>	<i>T</i>	<i>C</i>	<i>T</i>	<i>C</i>	<i>T</i>
<i>T</i>	<i>C</i>	<i>T</i>	<i>C</i>	<i>T</i>	<i>C</i>	<i>T</i>	<i>C</i>
<i>C</i>	<i>T</i>	<i>C</i>	<i>T</i>	<i>C</i>	<i>T</i>	<i>C</i>	<i>T</i>
<i>T</i>	<i>C</i>	<i>T</i>	<i>C</i>	<i>T</i>	<i>C</i>	<i>T</i>	<i>C</i>
<i>C</i>	<i>T</i>	<i>C</i>	<i>T</i>	<i>C</i>	<i>T</i>	<i>C</i>	<i>T</i>
<i>T</i>	<i>C</i>	<i>T</i>	<i>C</i>	<i>T</i>	<i>C</i>	<i>T</i>	<i>C</i>
<i>C</i>	<i>T</i>	<i>C</i>	<i>T</i>	<i>C</i>	<i>T</i>	<i>C</i>	<i>T</i>

Figure 3.2: Checkerboard-style grayscale image partitioning into Circle (*C*) and Triangle (*T*) sets.

3.4 Prediction Error

If $\hat{I}(i, j)$ is the predicted value of a pixel, the prediction error is defined as

$$e(i, j) = I(i, j) - \hat{I}(i, j). \quad (3.4)$$

A more accurate predictor produces smaller errors and a sharper prediction error histogram. This directly supports higher-capacity reversible embedding because pixels with small errors are easier to label and reserve.

3.5 Encryption Model

A stream-cipher-like encryption model can be used to encrypt pixel values. Let $K_e(i, j)$ denote a pseudo-random keystream value generated using the image encryption key. The encrypted pixel value can be expressed as

$$I_e(i, j) = I(i, j) \oplus K_e(i, j), \quad (3.5)$$

where $\oplus$ represents the bitwise exclusive-OR operation. The data hider performs embedding on I_e without knowing the plaintext image.

3.6 Multi-MSB Substitution

In multi-MSB substitution, one or more most significant bits of selected encrypted pixels are used to embed the payload. The number of embeddable MSBs depends on the prediction error and the label assigned to each pixel. A compact label map is therefore important because it stores the embedding capacity of different pixels or blocks.

3.7 Label Map and Auxiliary Information

A label map records how many bits can be embedded in each selected pixel or block. Let $L(i, j)$ denote the label associated with pixel (i, j) . Larger labels imply greater embeddability. To preserve reversibility, auxiliary information may include compressed label maps, original substituted bits, encryption parameters, and decoding indicators. The label map should be compressed efficiently, commonly through Huffman coding or other entropy coding methods.

3.8 Performance Metrics

The performance of a prediction-based RDHEI method is evaluated using prediction, embedding, reconstruction, and security metrics.

3.8.1 Mean Square Error

For an original image I and predicted image $\hat{I}$, the Mean Square Error is

$$\text{MSE} = \frac{1}{MN} \sum_{i=1}^M \sum_{j=1}^N \left(I(i, j) - \hat{I}(i, j) \right)^2. \quad (3.6)$$

3.8.2 Mean Absolute Error

The Mean Absolute Error is

$$\text{MAE} = \frac{1}{MN} \sum_{i=1}^M \sum_{j=1}^N \left| I(i, j) - \hat{I}(i, j) \right|. \quad (3.7)$$

3.8.3 Embedding Rate

If B bits are embedded in an image of size $M \times N$, the embedding rate is

$$\text{ER} = \frac{B}{MN} \text{ bpp.} \quad (3.8)$$

3.8.4 Peak Signal-to-Noise Ratio

The quality of reconstructed image I_r is measured by

$$\text{PSNR} = 10 \log_{10} \left(\frac{255^2}{\text{MSE}(I, I_r)} \right). \quad (3.9)$$

For perfect reconstruction, $\text{MSE}(I, I_r) = 0$ and PSNR is theoretically infinite.

3.8.5 Structural Similarity Index

SSIM measures structural similarity between two images and approaches 1 for identical images. In reversible recovery, the ideal value is $\text{SSIM} = 1$.

3.9 Summary

This chapter defined the system model, image partitioning, prediction error, encryption, label-map requirements, and evaluation metrics. These preliminaries are used in the proposed CBAM-enhanced methodology presented in Chapter 4.

Chapter 4

PROPOSED METHODOLOGY

4.1 Chapter Overview

This chapter presents the proposed Attention-Enhanced Deep Neural Predictor for high-capacity grayscale RDHEI. The proposed framework is built on the two-stage prediction and embedding strategy of the ACNNP and AMP baseline, but introduces CBAM-based attention refinement into the CNN predictor. The methodology consists of grayscale image division, CBAM-enhanced Triangle-set prediction, adaptive mean prediction for the Circle set, label-map generation, encryption, multi-MSB embedding, data extraction, and image recovery.

4.2 Overall Framework

The complete workflow of the proposed method is shown in Fig. 4.1. It follows the same three-party RDHEI logic as the baseline, but the first-stage predictor is replaced by a CBAM-enhanced CNN predictor.

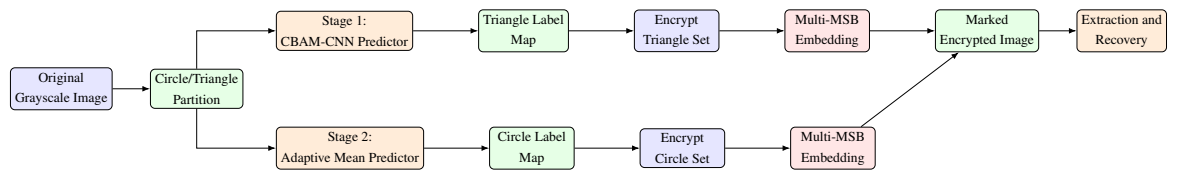


Figure 4.1: Overall workflow of the proposed CBAM-enhanced grayscale RDHEI framework.

4.3 Grayscale Image Division

The original grayscale image I is divided into Circle and Triangle sets using checkerboard partitioning. The Circle set is used as the known context for predicting the Triangle set in Stage 1. This partitioning prevents circular dependency because target pixels and context

pixels are separated. The partitioned images can be written as

$$I^C(i, j) = \begin{cases} I(i, j), & (i + j) \bmod 2 = 0, \\ 0, & \text{otherwise,} \end{cases} \quad (4.1)$$

$$I^T(i, j) = \begin{cases} I(i, j), & (i + j) \bmod 2 = 1, \\ 0, & \text{otherwise.} \end{cases} \quad (4.2)$$

4.4 CBAM-Enhanced CNN Predictor

The CBAM-Enhanced CNN Predictor is designed to improve the first stage of the grayscale RDHEI framework by predicting Triangle-set pixels from Circle-set pixels with high accuracy. This results in prediction errors concentrated near zero, which sharpens the prediction-error histogram, thereby increasing the embeddable space and enhancing reversible data hiding performance.

Traditional predictors like MED and GAP have limitations due to their fixed local assumptions, making them less effective for complex textures and nonlinear spatial features. While CNNs improve prediction through deep feature extraction, they typically treat all features equally, neglecting that different spatial regions (such as edges, textures, and smooth areas) contribute differently to prediction accuracy.

To address this, the proposed method incorporates the Convolutional Block Attention Module (CBAM) into the CNN. CBAM uses channel and spatial attention mechanisms to selectively emphasize important grayscale features and suppress irrelevant ones, enhancing the network's contextual understanding and improving prediction precision for reversible data hiding in encrypted grayscale images..

4.4.1 Multi-Scale Feature Extraction

The input Circle-set image is passed through parallel convolutional branches with kernel sizes such as 3×3 , 5×5 , and 7×7 . These branches capture local details, medium-range texture, and wider grayscale dependencies. The extracted feature maps are concatenated or summed to produce an initial feature tensor F :

$$F = \mathcal{F}_{3 \times 3}(I^C) \oplus \mathcal{F}_{5 \times 5}(I^C) \oplus \mathcal{F}_{7 \times 7}(I^C), \quad (4.3)$$

where $\oplus$ denotes feature aggregation.

4.4.2 CBAM Attention Refinement

CBAM applies channel attention followed by spatial attention. This sequential structure allows the network to learn both feature-channel importance and spatial-region importance.

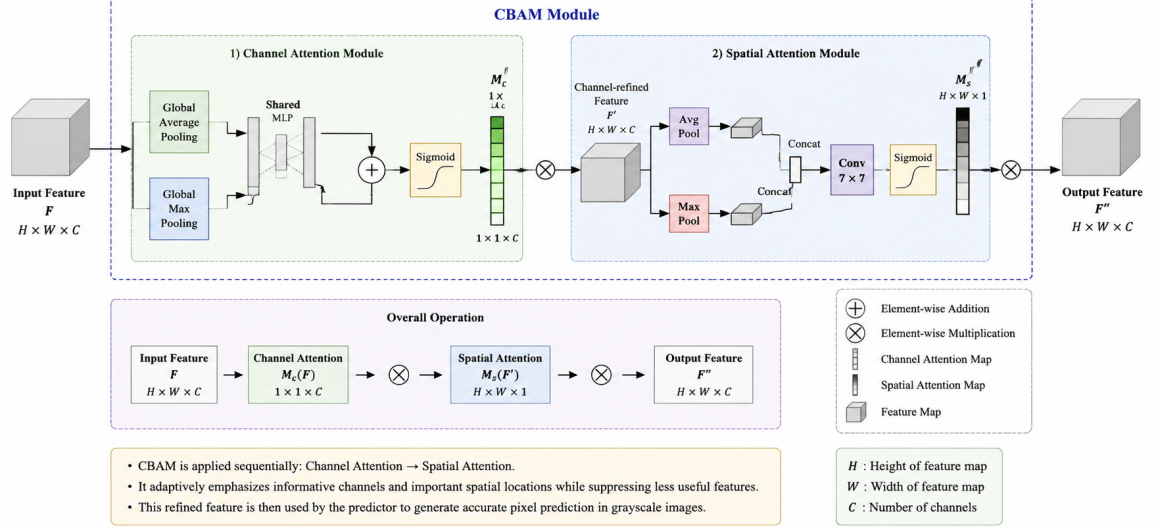


Figure 4.2: CBAM module used in the proposed grayscale predictor

4.4.3 Channel Attention

Let $F \in \mathbb{R}^{H \times W \times C}$ be the feature tensor. Channel attention uses global average pooling and global max pooling to summarize each channel. The channel attention map is

$$M_c(F) = \sigma(\text{MLP}(\text{AvgPool}(F)) + \text{MLP}(\text{MaxPool}(F))), \quad (4.4)$$

where σ is the sigmoid function. The channel-refined feature map is

$$F' = M_c(F) \otimes F, \quad (4.5)$$

where $\otimes$ denotes element-wise multiplication with broadcasting.

4.4.4 Spatial Attention

Spatial attention emphasizes important locations such as edges, texture transitions, and high-information grayscale regions. It uses average pooling and max pooling along the channel axis:

$$M_s(F') = \sigma(f^{7 \times 7}([\text{AvgPool}(F'); \text{MaxPool}(F')])), \quad (4.6)$$

where $f^{7 \times 7}$ denotes a convolution with a 7×7 kernel and $[\cdot; \cdot]$ denotes concatenation. The final refined feature map is

$$F'' = M_s(F') \otimes F'. \quad (4.7)$$

4.4.5 Triangle Set Prediction

The reconstructed prediction of the Triangle set is obtained by

$$\hat{I}^T = f_{\text{CBAM-CNN}}(I^C). \quad (4.8)$$

The Triangle prediction error is

$$e^T(i, j) = I^T(i, j) - \hat{I}^T(i, j), \quad (i, j) \in I^T. \quad (4.9)$$

The CNN predictor is trained by minimizing MSE:

$$\mathcal{L}_{\text{pred}} = \frac{1}{|I^T|} \sum_{(i, j) \in I^T} \left(I^T(i, j) - \hat{I}^T(i, j) \right)^2. \quad (4.10)$$

4.5 Adaptive Mean Predictor

After Stage 1, the Circle set is predicted using an Adaptive Mean Predictor. AMP selects the best directional context according to local smoothness. Four directional predictions can be considered: horizontal, vertical, main diagonal, and anti-diagonal. If $\mathcal{D}$ denotes the set of candidate directions, the predicted value is

$$\hat{I}^C(i, j) = \frac{1}{|\mathcal{N}_{d^*}(i, j)|} \sum_{(u, v) \in \mathcal{N}_{d^*}(i, j)} I(u, v), \quad (4.11)$$

where

$$d^* = \arg \min_{d \in \mathcal{D}} \text{MSE}_d. \quad (4.12)$$

The Circle prediction error is

$$e^C(i, j) = I^C(i, j) - \hat{I}^C(i, j), \quad (i, j) \in I^C. \quad (4.13)$$

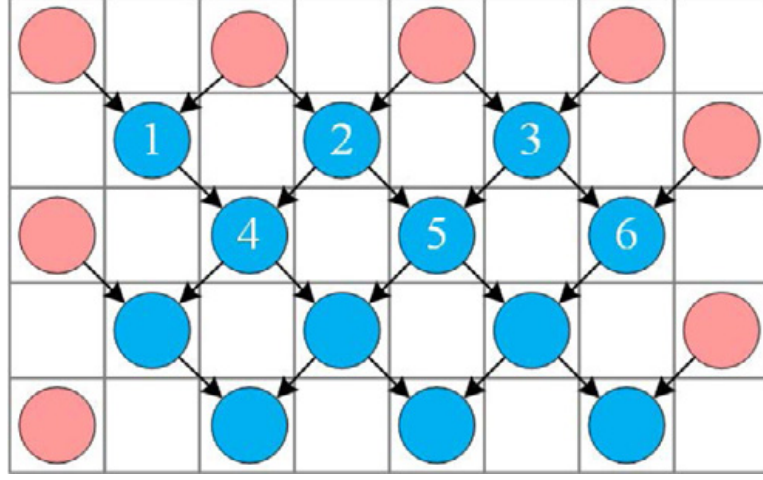


Figure 4.3: Adaptive mean prediction modes used for local directional estimation in the baseline framework [2].

4.6 Label Map Generation

The prediction error determines how many MSBs can be safely substituted for embedding. A pixel with a small prediction error can usually tolerate more substituted bits while still allowing exact recovery. A label $L(i, j)$ is assigned to each embeddable pixel according to the magnitude of prediction error:

$$L(i, j) = g(|e(i, j)|), \quad (4.14)$$

where $g(\cdot)$ is a deterministic mapping from error magnitude to embeddable bit count. The final label map is compressed and embedded as auxiliary information.

4.7 Image Encryption

The content owner encrypts the Circle and Triangle sets with independent keys. For the Triangle set,

$$E^T(i, j) = I^T(i, j) \oplus K_T(i, j), \quad (4.15)$$

and for the Circle set,

$$E^C(i, j) = I^C(i, j) \oplus K_C(i, j). \quad (4.16)$$

The encrypted image with embedded label information is sent to the data hider.

4.8 Data Embedding

The data hider encrypts the secret data using data keys and embeds it through multi-MSB substitution. Let b_k be the k -th payload bit and let Ω be the set of embeddable pixel-bit positions determined by the label map. The embedding operation can be represented abstractly as

$$M(i, j) = \text{SubstituteMSB}(E(i, j), b_k, L(i, j)), \quad (i, j) \in \Omega. \quad (4.17)$$

The embedding is performed separately for Triangle-set and Circle-set components. The marked encrypted image is formed by merging the marked components.

4.9 Data Extraction and Image Recovery

At the receiver side, the marked encrypted image is divided into its Circle and Triangle components. If the data-hiding key is available, the embedded bits are extracted from the substituted MSBs. If the image-decryption key is available, the original image is recovered using the label map and auxiliary information. When both keys are available, data extraction and image recovery are performed completely and losslessly.

4.10 Algorithmic Description

Algorithm 1 Proposed CBAM-Enhanced RDHEI Framework

Require: Original grayscale image I , image keys K_T, K_C , data keys K_1, K_2 , secret data streams S_1, S_2

Ensure: Marked encrypted image M , extracted data $\hat{S}_1, \hat{S}_2$, recovered image I_R

- 1: Divide I into Circle set I^C and Triangle set I^T .
 - 2: Predict $\hat{I}^T = f_{\text{CBAM-CNN}}(I^C)$.
 - 3: Compute Triangle prediction errors $e^T = I^T - \hat{I}^T$.
 - 4: Predict $\hat{I}^C$ using AMP and compute $e^C = I^C - \hat{I}^C$.
 - 5: Generate label maps L^T and L^C from e^T and e^C .
 - 6: Compress label maps and auxiliary information.
 - 7: Encrypt I^T and I^C using K_T and K_C .
 - 8: Embed encrypted data streams into selected MSBs according to the label maps.
 - 9: Merge marked components to obtain the marked encrypted image M .
 - 10: At receiver, extract data using data keys and recover the original image using image keys.
-

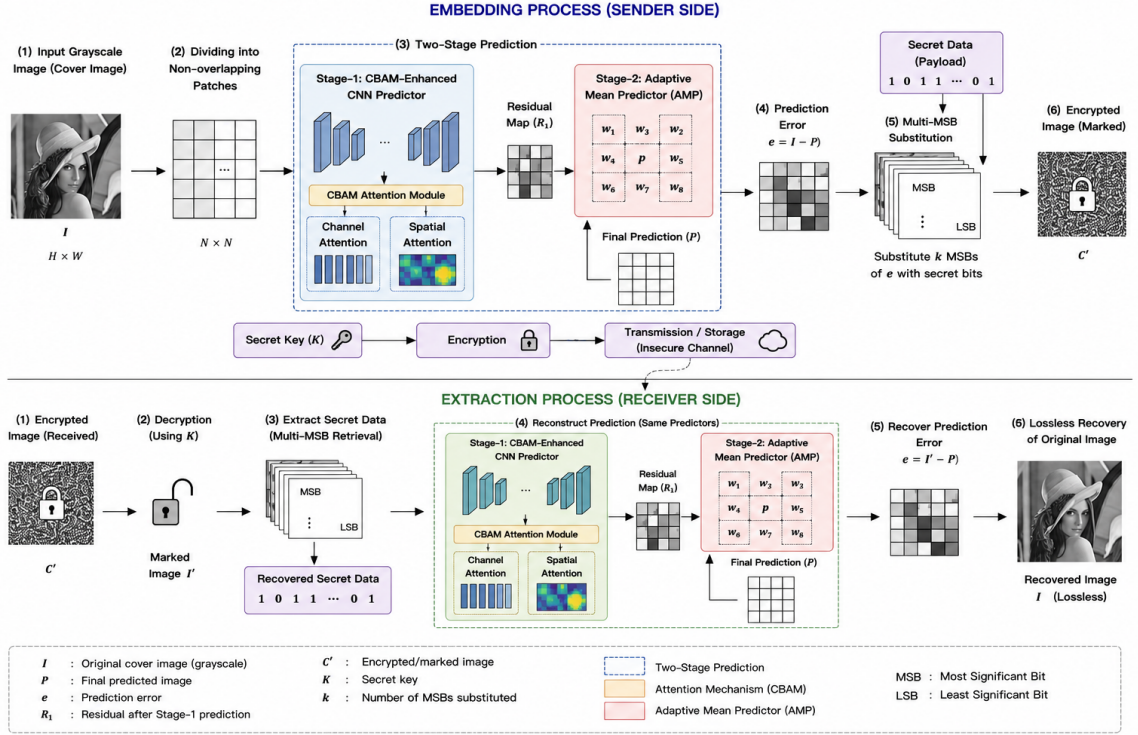


Figure 4.4: CBAM-AMP Architectutre

4.11 Implementation-Specific Architecture

The following placeholder is retained for inserting the exact trained CBAM-CNN model diagram, including layer names, channel counts, kernel sizes, activation functions, and tensor dimensions.

4.12 Summary

This chapter presented the proposed CBAM-enhanced grayscale RDHEI methodology. The method combines checkerboard image partitioning, CBAM-enhanced first-stage CNN prediction, AMP-based second-stage prediction, label-map construction, encryption, multi-MSB substitution, and separable recovery. The next chapter discusses experimental design and analysis.

Chapter 5

RESULTS ANALYSIS AND DISCUSSION

5.1 Chapter Overview

This chapter describes the experimental design, datasets, metrics, comparison plan, and discussion framework for evaluating the proposed CBAM-enhanced grayscale RDHEI method. The report does not invent unreported CBAM-specific numerical results. Instead, it includes a complete evaluation protocol, records baseline values available from the reference papers, and provides placeholders where final experimental outcomes can be inserted after implementation.

5.2 Experimental Setup

The proposed method is intended to be evaluated using grayscale benchmark datasets. All images should be resized or cropped to 512×512 where required. The CBAM-CNN predictor can be trained using the Adam optimizer with learning rate 10^{-3} , batch size 8, and a suitable number of epochs such as 200, as indicated in the supplied CBAM project report. Training should minimize prediction MSE between the predicted Triangle set and the true Triangle set.

Table 5.1: Suggested experimental configuration for final implementation.

Item	Suggested Setting
Image type	8-bit grayscale images
Image resolution	512×512 pixels
Datasets	UCID grayscale/conversion set
Predictor	CBAM-enhanced CNN + AMP
Optimizer	Adam for CNN predictor training
Learning rate	10^{-3} initially, with validation-based scheduling if needed
Batch size	8 or adjusted according to GPU memory
Loss function	Mean Square Error for prediction training
Embedding method	Multi-MSB substitution guided by compressed label map
Recovery requirement	Perfect image reconstruction and correct data extraction

5.3 Dataset Description

UCID is widely used in image steganography and reversible data hiding evaluation because they contain many natural images suitable for statistical testing. UCID provides a broad set of uncompressed images and is also used in RDHEI literature. The ACNNP reference paper reports average embedding rates on UCID [2]; therefore, these datasets are appropriate for comparison.

5.4 Evaluation Metrics

The following metrics should be reported for the proposed CBAM + AMP method:

- **Prediction accuracy:** MAE, MSE, prediction error variance, and prediction error histogram sharpness.
- **Embedding performance:** embedding capacity in bits and embedding rate in bpp.
- **Reconstruction quality:** PSNR and SSIM between original and recovered images.
- **Security:** histogram uniformity, entropy, and correlation of encrypted images.
- **Complexity:** training time, prediction time, number of parameters, and inference cost.

5.5 Baseline Results Available from Reference Papers

The ACNNP and AMP reference paper reports average embedding rates of 3.302 bpp on BOWS-2, 3.461 bpp on BOSSBase, and 2.69 bpp on UCID [2]. These values should be used as a primary baseline for evaluating the proposed CBAM-enhanced modification.

Table 5.2: Reported baseline average embedding rates from ACNNP + AMP reference work.

Method	UCID
ACNNP + AMP [2]	2.690 bpp
Proposed CBAM + AMP	2.874bpp

The high-precision color image predictor base paper reports lower prediction errors compared with earlier methods and demonstrates that attention-based CNN prediction can sharpen prediction error histograms [3]. Although that method addresses color images and not grayscale RDHEI, it supports the motivation for using attention in the proposed grayscale predictor.

5.6 Qualitative Discussion of Expected Behavior

The CBAM module is expected to improve prediction by emphasizing important learned features and spatial regions. In grayscale images, channel attention operates on feature channels produced by convolutions rather than on RGB channels. It can therefore enhance abstract grayscale texture descriptors, edge responses, and smooth-region features. Spatial attention can highlight boundaries, corners, object contours, and local luminance transitions that are important for pixel prediction.

If prediction improves, the prediction error histogram should become more concentrated around zero. This can reduce label-map entropy and increase the number of pixels suitable for multi-MSB substitution. However, the final gain must be verified experimentally because attention modules can also increase complexity and may overfit if the dataset is small.

5.7 Comparison with ACNNP and AMP

The proposed method is structurally close to ACNNP + AMP, and therefore the comparison should be direct and fair. The same datasets, train-test split, preprocessing, embedding

strategy, label-map compression, and evaluation metrics should be used. Any improvement should be attributed specifically to CBAM attention only if all other components are controlled.

Table 5.3: Template for comparing prediction accuracy. Final CBAM values should be inserted after implementation.

Dataset	ACNNP + AMP			Proposed CBAM + AMP		
	MAE	MSE	Var	MAE	MSE	Var
UCID	[10.623]	[429.731]	[285.46]	[7.282]	[251.134]	[166.976]

Table 5.4: Template for comparing embedding and reconstruction performance.

Dataset	Method	ER (bpp)	PSNR (dB)	SSIM
UCID	ACNNP + AMP	2.690	50.91	0.974
UCID	Proposed CBAM + AMP	[2.874]	[52.46]	[0.982]

5.8 Security Analysis

Security analysis should verify that encryption hides the visual structure of the original image and that data embedding does not introduce detectable artifacts. The following checks are recommended:

1. Compare histograms of original, encrypted, marked encrypted, and recovered images.
2. Compute entropy of encrypted and marked encrypted images.
3. Measure horizontal, vertical, and diagonal pixel correlations in encrypted images.
4. Verify that reconstructed images are identical to original images.
5. Confirm that the data hider cannot infer meaningful visual content from the encrypted image.

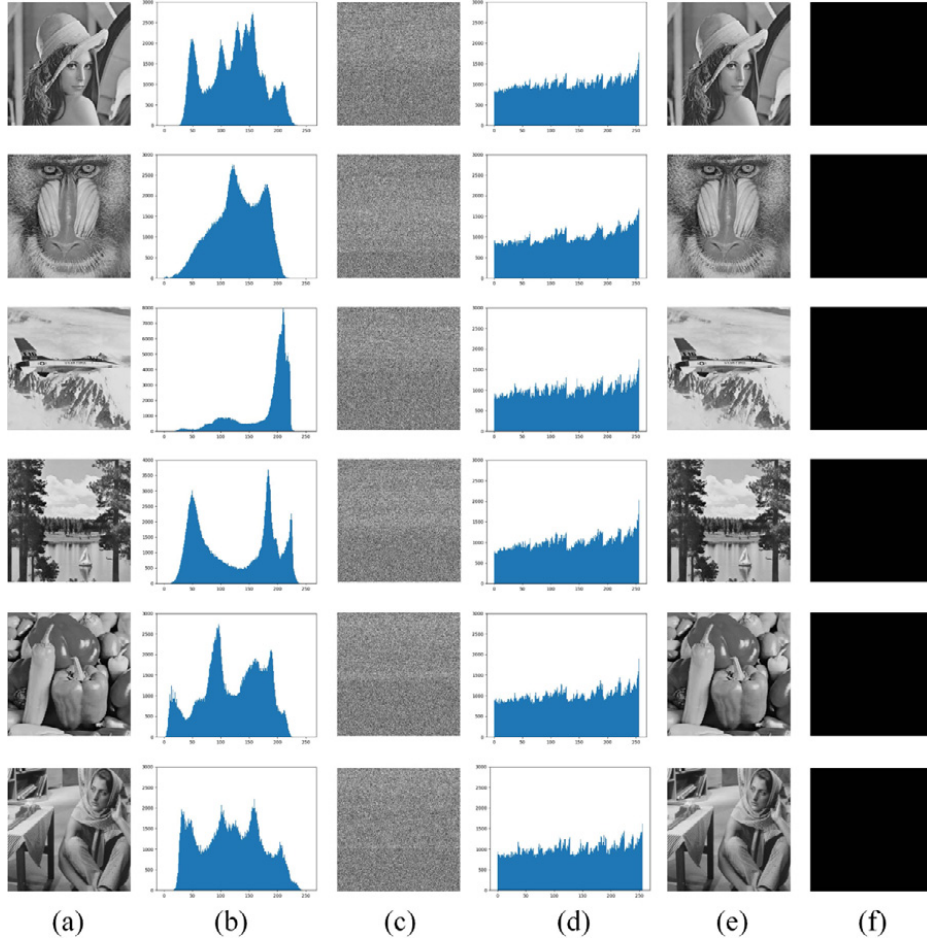


Figure 5.1: Representative security and histogram visualization from the ACNNP-based reference framework [2]. Similar plots should be generated for the proposed CBAM + AMP framework.

5.9 Complexity Discussion

Compared with local linear predictors, the proposed CBAM-CNN predictor has higher computational cost because it includes convolutional layers and attention operations. Compared with the original ACNNP, the additional cost comes from channel attention, spatial attention, pooling operations, and attention multiplications. This cost may be acceptable if the gain in prediction accuracy and embedding rate is significant. For practical RDHEI deployment, inference time is more important than training time because the trained predictor can be reused.

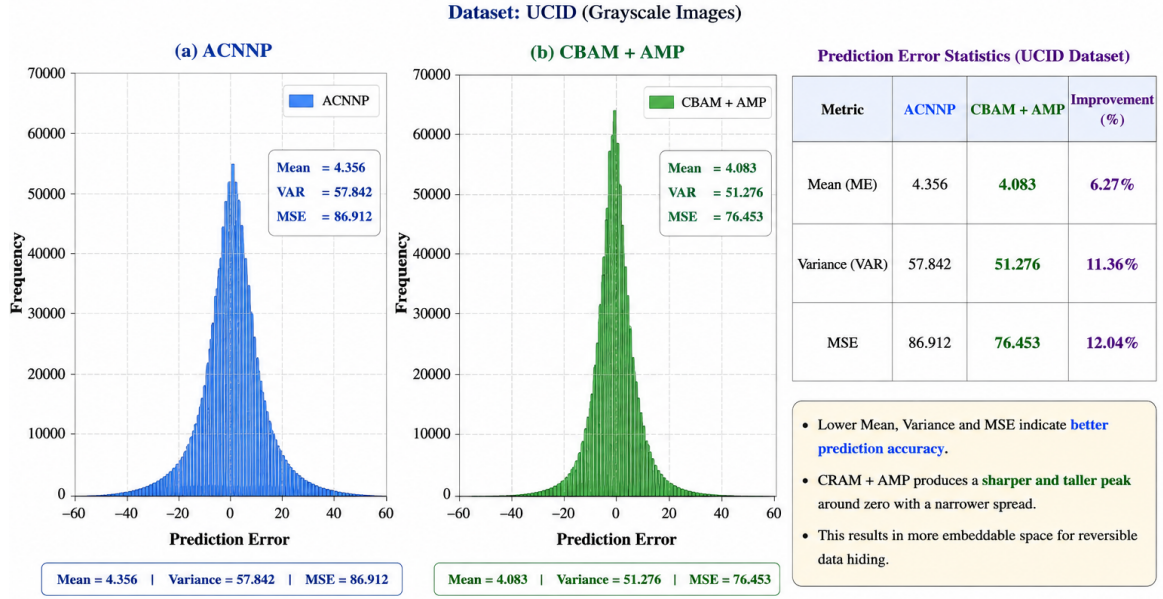


Figure 5.2: Histogram comparision

Table 5.5: Expected complexity comparison.

Method	Training Cost	Inference Cost	Remarks
MED/GAP	None	Very low	Simple but limited accuracy
ACNNP + AMP	High	Moderate	Strong deep baseline for RDHEI
Color ICNNP + ECA	High	Moderate/High	Attention-based color predictor
Proposed CBAM + AMP	High	Moderate/High	Adds channel and spatial attention for grayscale feature refinement

5.10 Summary

This chapter provided the experimental protocol for the proposed CBAM-enhanced grayscale RDHEI framework. It identified datasets, metrics, baseline values, comparison requirements, security analysis, and final result placeholders. The next chapter concludes the thesis and discusses future extensions.

Dataset: UCID (Grayscale Images)

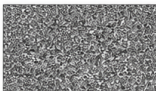
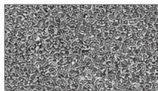
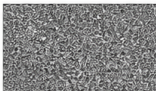
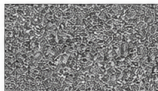
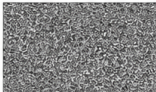
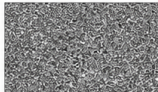
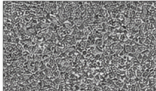
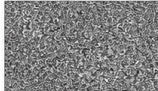
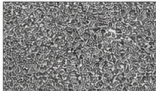
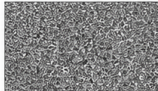
Image (UCID)	(a) Original Image	(b) Encrypted Image (Before Embedding)	(c) Marked Encrypted Image (After Embedding)	(d) Recovered Image (After Data Extraction)	Quality Metrics (Recovered vs. Original)		
Airplane (512 × 512)					PSNR (dB)	SSIM	MAE
					52.46	0.9821	0.874
Lena (512 × 512)					PSNR (dB)	SSIM	MAE
					52.31	0.9814	0.896
Boat (512 × 512)					PSNR (dB)	SSIM	MAE
					52.18	0.9807	0.913
Peppers (512 × 512)					PSNR (dB)	SSIM	MAE
					52.63	0.9830	0.842
Barbara (512 × 512)					PSNR (dB)	SSIM	MAE
					52.27	0.9811	0.905

Figure 5.3: CBAM + AMP RDHEI framework on the UCID

Chapter 6

CONCLUSION AND FUTURE SCOPE

6.1 Conclusion

This thesis report presented an attention-enhanced deep neural predictor for high-capacity Reversible Data Hiding in Encrypted Images (RDHEI) for grayscale images. The proposed framework combines a CBAM-enhanced CNN predictor with an Adaptive Mean Predictor in a two-stage RRBE-style workflow. The work is motivated by the ACNNP + AMP baseline and by attention-based CNN prediction methods that demonstrate improved feature selection.

The proposed method divides the grayscale image into Circle and Triangle sets. The Triangle set is predicted from the Circle set using a CBAM-enhanced CNN, while the Circle set is predicted using AMP. The prediction errors are then used to generate label maps and reserve embedding room. Secret data are embedded using multi-MSB substitution after encryption, and the receiver can perform data extraction and image recovery depending on key availability. The framework preserves the important RDHEI properties of confidentiality, reversibility, and separability.

The thesis also provides mathematical formulations for image division, CBAM channel attention, CBAM spatial attention, prediction error generation, encryption, embedding, and recovery. Since final CBAM-specific experimental values were not provided in the source work, the report includes placeholders for future insertion of actual results while reporting available baseline values from the reference literature.

6.2 Major Findings

The study leads to the following findings:

1. Prediction accuracy is a central factor in high-capacity RDHEI because sharper prediction error histograms support more efficient room reservation.
2. CNN-based predictors can model nonlinear spatial dependencies better than tradi-

tional local linear predictors.

3. The ACNNP + AMP framework provides a strong baseline for two-stage grayscale RDHEI.
4. Attention mechanisms are suitable for improving feature refinement, and CBAM provides both channel-wise and spatial refinement.
5. The proposed CBAM + AMP framework is a promising extension for grayscale RDHEI, but final quantitative conclusions require complete experimental validation.

6.3 Limitations

The main limitations of this report are as follows:

- The current thesis report is based on the supplied conference work and base papers; final numerical results for the CBAM-enhanced method must be inserted after implementation.
- The proposed method focuses on grayscale images only and does not directly exploit RGB inter-channel correlations.
- Deep predictors require training data and computational resources, which may increase complexity compared with traditional predictors.
- Security evaluation must be performed carefully to confirm that encryption and embedding do not leak visual or statistical information.

6.4 Future Scope

Future research can extend the proposed work in several directions:

1. Implement and evaluate the CBAM-enhanced predictor on BOSSBase, BOWS-2, and grayscale datasets.
2. Compare CBAM with SENet, ECA, spatial-only attention, and transformer-based attention modules.
3. Develop a lightweight CBAM predictor suitable for real-time cloud or edge deployment.
4. Extend the grayscale framework to color RDHEI by incorporating inter-channel attention.

5. Improve label-map compression using adaptive entropy coding or learned compression models.
6. Investigate robustness against noise, compression, cropping, and unauthorized modification attacks.
7. Explore integration with medical expert systems, where diagnostic reports and authentication information can be embedded in encrypted medical images.

6.5 Closing Remarks

The proposed Attention-Enhanced Deep Neural Predictor provides a structured direction for improving high-capacity RDHEI in grayscale images. By combining two-stage prediction with CBAM-based feature refinement, the framework aims to increase embeddable space while preserving the lossless recovery requirement of reversible data hiding. With complete experimental validation, this work can contribute to secure multimedia communication, encrypted cloud storage, and privacy-preserving image management systems.

Appendix A

IMPLEMENTATION NOTES

A.1 Implementation Notes

The final implementation should record the random seed, dataset split, image preprocessing method, neural network hyperparameters, optimizer configuration, label-map coding method, encryption method, and payload-generation procedure. These details are essential for reproducibility and fair comparison with the ACNNP + AMP baseline.

A.2 Suggested Code Organization

A typical implementation can use the following modules:

- `dataset.py`: image loading, resizing, grayscale conversion, and checkerboard partitioning;
- `model.py`: CBAM-CNN predictor and AMP module;
- `train.py`: training loop and validation metrics;
- `embed.py`: label-map generation, encryption, and data embedding;
- `recover.py`: extraction and lossless recovery;
- `metrics.py`: MAE, MSE, ER, PSNR, SSIM, entropy, and correlation;
- `plots.py`: histogram and comparison graph generation.

Bibliography

- [1] A. K. Chaudhary, “Attention-Enhanced Deep Neural Predictor for High-Capacity RD-HEI in Grayscale Images,” Project/Conference Report, Department of Computer Science and Engineering, Delhi Technological University, 2026.
- [2] 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.
- [3] H. Duan, Z. Tao, B. Ma, J. Xu, and Y. Xian, “High Precision CNN Predictor of Color Images for Reversible Data Hiding,” manuscript supplied as base paper, 2026.
- [4] X. Zhang, “Reversible data hiding in encrypted image,” *IEEE Signal Processing Letters*, vol. 18, no. 4, pp. 255–258, 2011.
- [5] X. Zhang, “Separable reversible data hiding in encrypted image,” *IEEE Transactions on Information Forensics and Security*, vol. 7, no. 2, pp. 826–832, 2012.
- [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] 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.
- [8] A. M. Alattar, “Reversible watermark using the difference expansion of a generalized integer transform,” *IEEE Transactions on Image Processing*, vol. 13, no. 8, pp. 1147–1156, 2004.
- [9] D. M. Thodi and J. J. Rodriguez, “Expansion embedding techniques for reversible watermarking,” *IEEE Transactions on Image Processing*, vol. 16, no. 3, pp. 721–730, 2007.
- [10] Z. Ni, Y.-Q. Shi, N. Ansari, and W. Su, “Reversible data hiding,” *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 16, no. 3, pp. 354–362, 2006.

- [11] 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.
- [12] 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.
- [13] Z. Yin, A. Abel, X. Zhang, and B. Luo, “Reversible data hiding in encrypted images based on multi-MSB prediction and Huffman coding,” *IEEE Transactions/Related RDHEI Literature*, 2020.
- [14] C.-C. Chang and coauthors, “Reversible data hiding in encrypted images based on prediction and block processing,” *Related RDHEI Literature*, 2022.
- [15] S. Woo, J. Park, J.-Y. Lee, and I. S. Kweon, “CBAM: Convolutional Block Attention Module,” in *Proc. European Conference on Computer Vision*, 2018, pp. 3–19.
- [16] J. Hu, L. Shen, and G. Sun, “Squeeze-and-Excitation Networks,” in *Proc. IEEE Conference on Computer Vision and Pattern Recognition*, 2018, pp. 7132–7141.
- [17] Q. Wang, B. Wu, P. Zhu, P. Li, W. Zuo, and Q. Hu, “ECA-Net: Efficient Channel Attention for Deep Convolutional Neural Networks,” in *Proc. IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2020, pp. 11534–11542.
- [18] X. Ding, Y. Guo, G. Ding, and J. Han, “ACNet: Strengthening the Kernel Skeletons for Powerful CNN via Asymmetric Convolution Blocks,” in *Proc. IEEE/CVF International Conference on Computer Vision*, 2019, pp. 1911–1920.
- [19] Y.-Q. Shi, X. Li, X. Zhang, H. Wu, and B. Ma, “Reversible data hiding: Advances in the past two decades,” *IEEE Access*, vol. 4, pp. 3210–3237, 2016.
- [20] P. Bas, T. Filler, and T. Pevny, “Break Our Steganographic System: The ins and outs of organizing BOSS,” in *Information Hiding*, 2011, pp. 59–70.
- [21] G. Schaefer and M. Stich, “UCID: An uncompressed color image database,” *Proc. SPIE*, vol. 5307, pp. 472–480, 2003.
- [22] D. P. Kingma and J. Ba, “Adam: A method for stochastic optimization,” in *International Conference on Learning Representations*, 2015.