IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 6, Issue 18, April 2026
Impact Factor: 8.2

Dual Robust Attention-Guided Deep Learning-Based Digital Image Watermarking with Gan Optimization for Secure Multimedia Applications

Saguturu Venkata Krishna Shruthi[1], Sarimilla Thimothi[2], Potu Sai Kiran[3]
Dr. Bhagya Lakshmi Nandipati[4]
B. Tech Computer Science and Engineering[1-4]
RVR & JC College of Engineering, Guntur, India
krishnashruthi2004@gmail.com[1], thimothisarimilla845@gmail.com[2], potusaikiranpsk@gmail.com[3], bhagyait@gmail.com[4]

Abstract: *Digital image watermarking plays a crucial role in protecting multimedia content against unauthorized access, copyright infringement, and data tampering in modern digital communication systems. With the rapid growth of multimedia applications, ensuring both security and visual fidelity has become a significant challenge. Existing deep learning-based watermarking approaches, particularly those utilizing convolutional neural networks (CNN) and autoencoder architectures, have demonstrated promising results in embedding and extracting watermark images efficiently. However, these methods are often limited by single watermark embedding, restricted capacity, lack of adaptive feature selection, and reduced robustness against complex and hybrid attacks such as noise addition, compression, filtering, and geometric distortions.*

To overcome these limitations, this paper proposes a novel dual watermarking framework that integrates attention mechanisms with Generative Adversarial Networks (GAN) to enhance both embedding efficiency and robustness. The proposed model simultaneously embeds two watermark images—a primary copyright watermark and a secondary authentication watermark—thereby significantly increasing embedding capacity and security. An attention-guided feature selection module is incorporated to identify perceptually significant regions of the cover image, enabling adaptive embedding that minimizes visual distortion and improves imperceptibility. Furthermore, a GAN-based adversarial training strategy is employed to simulate various real-world attacks during training, thereby strengthening the model's ability to recover watermark information under adverse conditions.

The embedding process is performed using an encoder-decoder architecture with feature fusion, while the extraction process utilizes a denoising autoencoder combined with convolutional layers to accurately reconstruct embedded watermarks. The standard metrics such as Peak Signal-to- Noise Ratio (PSNR), Structural Similarity Index Measure (SSIM), and Normalized Correlation (NC). Experimental results demonstrate that the proposed approach outperforms existing state-of-the-art techniques by achieving higher PSNR and SSIM values, indicating superior visual quality, along with improved NC values, reflecting enhanced robustness and extraction accuracy.

Keywords: Digital Watermarking, Deep Learning, CNN, GAN, Dual Watermarking, Image Security, Autoencoder

I. INTRODUCTION

In the modern digital era, digital images are extensively used across various domains such as healthcare, social media, surveillance, forensics, education, and entertainment. The rapid advancement of internet technologies and cloud-based platforms has significantly increased the generation, storage, and transmission of multimedia data. While this

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 6, Issue 18, April 2026
Impact Factor: 8.2

accessibility improves communication and data sharing, it also raises serious concerns related to data security, copyright protection, and unauthorized duplication or manipulation of digital content.

Digital watermarking has emerged as a powerful technique to address these challenges by embedding hidden information into a cover image without significantly affecting its visual quality. The embedded watermark can later be extracted to verify ownership, authenticate content, or detect tampering. An effective watermarking system must achieve a balance between invisibility and robustness while maintaining sufficient embedding capacity.

A watermarking system is generally evaluated based on the following key performance parameters:
 Imperceptibility: The embedded watermark should not introduce noticeable distortions in the cover image. This ensures that the visual quality of the image remains intact and the presence of the watermark is not detectable by the human eye. High imperceptibility is essential for applications such as medical imaging and digital media distribution where image clarity is critical.
 Robustness: The watermark should be resistant to various types of intentional and unintentional attacks such as noise addition, compression (JPEG), filtering, rotation, scaling, and cropping. A robust watermark ensures that the embedded information can still be accurately extracted even after the image undergoes multiple processing operations.
 Embedding Capacity: The watermarking system should be capable of embedding a sufficient amount of information without compromising image quality or robustness. Higher capacity allows for embedding multiple watermarks or additional authentication data, which is important in advanced security applications.

A. Limitations of Traditional Watermarking Techniques

Traditional watermarking approaches are broadly classified into spatial and transform domain methods, each with its own advantages and limitations:

Spatial Domain Techniques:

These methods directly modify the pixel values of the image to embed watermark information. While they are simple to implement and computationally efficient, they are highly vulnerable to common image processing attacks such as noise addition and compression. Even minor modifications to pixel values can significantly degrade the embedded watermark, making these techniques less reliable for secure applications.

Transform Domain Techniques (DCT, DWT, SVD):

These approaches embed watermark information in transformed coefficients of the image rather than directly altering pixel values. As a result, they provide better robustness and resistance to attacks compared to spatial domain methods. However, these techniques rely heavily on manual selection of embedding regions and predefined rules, which limits their adaptability. Additionally, they may involve higher computational complexity and lack the ability to dynamically adjust embedding strategies based on image content.

Despite their advantages, traditional methods are not sufficient to handle complex real-world scenarios involving multiple and hybrid attacks.

B. Role of Deep Learning in Watermarking

With the rapid advancement of artificial intelligence, deep learning has revolutionized the field of digital image processing, including watermarking. Deep learning-based watermarking systems leverage neural networks to automatically learn optimal embedding and extraction strategies from data.

Key advantages of deep learning-based watermarking include:

Automatic Feature Extraction:

Convolutional Neural Networks (CNNs) can automatically extract meaningful features from images, eliminating the need for manual feature engineering. This allows the system to identify optimal embedding locations that preserve image quality.

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 6, Issue 18, April 2026
Impact Factor: 8.2

Adaptive Embedding Strategy:
Deep learning models can dynamically adjust embedding strength and location based on image characteristics, ensuring a better trade- off between imperceptibility and robustness.

End-to-End Learning Framework:
Encoder-decoder architectures enable the entire watermarking process—embedding and extraction—to be learned jointly. This reduces human intervention and improves system efficiency.

Robustness Through Data-Driven Learning:
By training on datasets that include various noise and attack conditions, deep learning models can learn to handle distortions effectively, leading to improved watermark recovery.

C. Overview of Base Paper Approach

The base paper presents a deep learning-based watermarking framework that utilizes convolutional neural networks and autoencoder architectures for embedding and extraction. The key components of the system are:

Encoder Network for Feature Extraction:
The encoder extracts latent feature representations from both the cover image and the watermark image. These features capture important structural and visual information required for embedding.

Feature Concatenation Mechanism:
The extracted features of the cover and watermark images are combined to form a unified representation. This fusion allows the watermark information to be embedded within the cover image effectively.

Decoder Network for Image Reconstruction:
The decoder reconstructs the watermarked image from the combined feature representation. It ensures that the final output image maintains high visual quality while containing the embedded watermark.

Denoising Autoencoder in Extraction Phase:
During the extraction process, a denoising autoencoder is used to remove noise and distortions from the received image. This improves the accuracy of watermark recovery, especially under noisy conditions.

This approach achieves high performance in terms of visual quality and robustness, as reflected by strong PSNR and SSIM values.

D. Limitations of the Existing Model

Despite its effectiveness, the base model has several limitations that restrict its applicability in advanced watermarking scenarios:

Single Watermark Embedding:
The system is designed to embed only one watermark image, which limits its ability to support multi-level security or multiple authentication layers.

Limited Embedding Capacity:
Since only one watermark is embedded, the amount of information that can be securely hidden within the image is restricted, reducing its applicability in complex systems.

Lack of Adaptive Embedding Mechanism:
The model does not explicitly identify optimal regions for em- bedding based on image content, which may lead to suboptimal performance in terms of imperceptibility and robustness.

Vulnerability to Advanced and Hybrid Attacks:
Although the model handles basic noise variations, it does not effectively address complex attack scenarios such as combined noise, compression, and geometric transformations.

Absence of Adversarial Training:
The model lacks mechanisms to simulate real-world attack conditions during training, limiting its ability to generalize and perform under unseen distortions.

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 6, Issue 18, April 2026
Impact Factor: 8.2

E. Proposed Extension

To overcome the above limitations, this paper proposes an advanced dual watermarking framework that integrates attention mechanisms and Generative Adversarial Networks (GAN) into the deep learning pipeline.

The proposed system offers the following improvements:

Dual Watermark Embedding:

Enables simultaneous embedding of two watermark images (e.g., copyright and authentication), thereby increasing capacity and enhancing security.

Attention-Based Feature Selection:

Identifies perceptually significant regions of the image, allowing adaptive embedding that minimizes distortion and improves imperceptibility.

GAN-Based Robustness Enhancement:

Incorporates adversarial training to simulate real-world attacks, improving the model's ability to recover watermark information under challenging conditions.

Improved Extraction Accuracy:

Combines denoising autoencoders with advanced feature processing to ensure reliable watermark recovery.

II. LITERATURE REVIEW

In recent years, digital image watermarking has evolved significantly with the integration of deep learning techniques. The increasing demand for secure multimedia transmission and copyright protection has led researchers to explore advanced methods that improve imperceptibility, robustness, and embedding capacity simultaneously. Unlike traditional approaches, deep learning-based watermarking systems are capable of automatically learning optimal embedding and extraction strategies from large datasets, thereby reducing manual intervention and improving performance under diverse conditions.

Traditional watermarking techniques operate in spatial and transform domains. Spatial domain methods directly modify pixel values, making them simple and computationally efficient; however, they are highly sensitive to noise and compression attacks. Transform domain methods such as Discrete Cosine Transform (DCT), Discrete Wavelet Transform (DWT), and Singular Value Decomposition (SVD) provide better robustness by embedding watermark information into frequency coefficients. Despite these advantages, transform domain methods rely heavily on manual feature selection and predefined embedding strategies, which limits their adaptability and performance in complex real-world scenarios [13]. Furthermore, these methods are not capable of dynamically adjusting embedding strength based on image content.

With the advancement of deep learning, Convolutional Neural Networks (CNN) have been widely adopted for watermarking applications due to their ability to automatically extract hierarchical features from images. Ding et al. [7] proposed a deep neural network- based watermarking scheme that achieved high imperceptibility by learning embedding patterns directly from data. However, the model exhibited poor robustness against common attacks such as compression, filtering, and rotation, indicating its limited generalization capability. Similarly, Zhong et al. [25] developed a CNN-based embedding and extraction framework that improved automation in watermarking. Despite this, the model suffered from information loss during end-to-end training, which negatively impacted watermark recovery accuracy.

Autoencoder-based watermarking approaches have gained significant attention due to their ability to reconstruct images while preserving embedded information. Mahapatra et al. [16] proposed a convolutional autoencoder-based watermarking system that enhanced embedding quality and maintained high visual fidelity. However, the system was sensitive to noise variations and often extracted noisy watermark information, reducing reliability. Wei et al. [23] introduced a variational autoencoder-based watermarking technique that improved visual quality through probabilistic modeling. Nevertheless, the method suffered from limited watermark capacity, restricting its application in scenarios requiring multi-level information embedding.

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 6, Issue 18, April 2026
Impact Factor: 8.2

Hybrid approaches that combine deep learning with traditional signal processing techniques have also been explored to improve overall performance. Wang et al. [22] proposed a CNN-based mapping model for watermark embedding, which improved robustness by learning relationships between watermark and cover images. However, this approach increased computational complexity and required higher processing time. Zheng et al. [24] combined DWT with CNN to leverage both frequency domain robustness and deep learning adaptability. While this method improved imperceptibility, it introduced additional complexity and lacked scalability for large datasets.

Artificial Neural Network (ANN)-based watermarking methods were proposed by Islam et al. [10], demonstrating improved watermark detection capabilities. However, these methods suffered from low embedding capacity and limited robustness against advanced image processing attacks, making them less suitable for modern applications.

The base paper [Base Paper] proposes a CNN-based watermarking technique using an encoder-decoder architecture, where latent features of cover and watermark images are concatenated to generate a watermarked image. Additionally, a denoising autoencoder is used during the extraction phase to remove noise variations and improve watermark recovery. Experimental results indicate high performance in terms of PSNR and SSIM, reflecting strong imperceptibility and robustness. However, as observed from the literature, most existing methods—including the base model—suffer from limitations such as low embedding capacity, lack of adaptive embedding strategies, and insufficient robustness against complex and hybrid attacks.

III. PROBLEM STATEMENT

Despite the significant progress in deep learning-based watermarking techniques, existing systems still face several critical limitations that restrict their practical applicability in real-world multimedia security environments. While CNN and autoencoder-based approaches have improved visual quality and basic robustness, they fail to address important challenges related to embedding capacity, adaptability, and resistance to complex attack scenarios.

The major problems identified from the literature are discussed below:

A. Limited Embedding Capacity

Most existing watermarking systems are designed to embed only a single watermark image within a cover image. This limitation significantly reduces the amount of information that can be securely hidden and restricts the system's ability to support advanced security features such as multi-level authentication and ownership verification.

In real-world applications, there is often a requirement to embed multiple types of data simultaneously, such as copyright information, user identity, transaction details, and authentication codes. Single watermark systems are insufficient to meet these requirements, thereby limiting their scalability and usability.

Increasing embedding capacity in existing systems often results in degradation of image quality or reduced robustness, making it difficult to achieve an optimal balance among performance metrics.

B. Reduced Robustness Against Hybrid Attacks

Most existing watermarking models are evaluated under isolated attack conditions such as noise addition or compression. However, in real-world scenarios, images are frequently subjected to multiple simultaneous distortions, commonly referred to as hybrid attacks.

Examples of hybrid attacks include combinations of Gaussian noise with JPEG compression, rotation followed by filtering, or scaling combined with cropping. These complex distortions significantly degrade the embedded watermark and reduce extraction accuracy.

Existing models lack the ability to generalize under such conditions, as they are not trained using diverse attack simulations, resulting in poor robustness in practical environments.

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 6, Issue 18, April 2026
Impact Factor: 8.2

C. Lack of Adaptive Embedding Mechanism

Many watermarking techniques do not incorporate adaptive strategies to determine optimal embedding regions within the image. Instead, watermark data is embedded uniformly or based on fixed rules, which may not be suitable for all image types.

Different regions of an image exhibit varying sensitivity to distortion. Smooth regions are more prone to visible artifacts, whereas textured regions can better conceal embedded information. Without adaptive embedding, the system fails to exploit these characteristics effectively.

The absence of intelligent region selection leads to suboptimal trade-offs between imperceptibility and robustness, ultimately affecting overall system performance.

D. Extraction Errors Due to Noise Variations

The presence of noise, compression artifacts, and geometric distortions in the watermarked image significantly affects the accuracy of watermark extraction.

Existing models, even those using denoising autoencoders, are often trained on limited noise patterns and fail to generalize to unseen distortions, leading to incorrect or incomplete watermark recovery.

Inaccurate extraction reduces the reliability of the watermarking system, especially in critical applications such as medical imaging, forensic analysis, and secure communications.

E. Absence of Advanced Robustness Techniques

Most existing watermarking systems do not incorporate adversarial training mechanisms such as Generative Adversarial Networks (GAN), which are capable of simulating real-world attack conditions during training.

Without adversarial learning, models lack exposure to dynamically generated distortions, reducing their ability to learn robust and invariant feature representations.

This limitation prevents the system from achieving high resilience under unknown or complex attack scenarios.

IV. METHODOLOGY

The proposed watermarking system is designed to enhance embedding capacity, robustness, and imperceptibility by integrating dual watermarking, attention mechanisms, and adversarial learning. The methodology consists of two major phases: embedding and extraction, followed by a formal algorithm representation.

A. Embedding Process

The embedding process is responsible for inserting two watermark images into a single cover image while preserving visual quality and ensuring robustness.

The key steps involved in the embedding process are as follows:

Input Acquisition:

The system takes a high-resolution cover image C along with two watermark images W1 and W2.The primary watermark W1 represents copyright information, while the secondary watermark W2 represents authentication or hidden data.

Feature Extraction using CNN Encoder:

A convolutional neural network (CNN) encoder is used to extract latent feature representations from the cover image and both watermark images. These features capture spatial and structural information necessary for embedding.

Attention-Based Feature Selection:

An attention module is applied to the extracted cover image features to identify perceptually significant regions. This ensures that watermark embedding is performed in areas where visual distortion is minimal and robustness is maximized.

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 6, Issue 18, April 2026
Impact Factor: 8.2

Feature Fusion:
The attention-refined cover features are combined with watermark features using concatenation or feature fusion techniques. This step integrates the watermark information into the cover image representation.

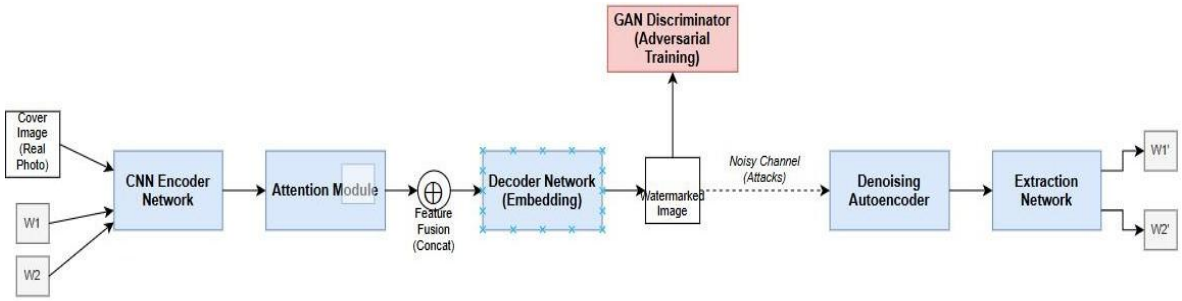

Fig. 1. Proposed Dual Watermarking System Architecture

Image Reconstruction using Decoder:
A decoder network reconstructs the final watermarked image M from the fused feature representation. The decoder ensures that the output image maintains high visual quality while securely embedding both watermarks.

B. Extraction Process
The extraction process is responsible for recovering both watermark images from the potentially distorted watermarked image.
The steps involved are:
 Input of Watermarked Image:
The system receives the possibly distorted or attacked watermarked image M′.
 Denoising using Autoencoder:
A denoising autoencoder is applied to remove noise, compression artifacts, and distortions from the input image. This improves the quality of features used for extraction.
 Latent Feature Extraction:
CNN encoders extract latent representations from the denoised image. These features contain embedded watermark information.
 Feature Separation:
The combined features are separated into individual watermark components using learned feature mappings.
 Watermark Reconstruction:
Decoder networks reconstruct the original watermark images W1′ and W2′ from the separated features.

C. Proposed Algorithm
Algorithm : Dual Watermark Embedding and Extraction
Input:
C ← Cover Image
W1 ← Primary Watermark Image W2 ← Secondary Watermark Image
Output:
M ← Watermarked Image
W1' ← Extracted Primary Watermark W2' ← Extracted Secondary Watermark
---
Embedding Phase

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429 Volume 6, Issue 18, April 2026 Impact Factor: 8.2

Initialize Encoder, Attention Module, Decoder
Normalize inputs C, W1, W2
// Feature Extraction FC ← Encoder(C) FW1 ← Encoder(W1) FW2 ← Encoder(W2)
// Attention Mechanism FA ← Attention(FC)
// Feature Fusion
F ← Concatenate(FA, FW1, FW2)
// Reconstruction M ← Decoder(F)
Return Watermarked Image M

Extraction Phase

Input M' (possibly attacked image)
// Denoising
M_clean ← DenoisingAutoencoder(M')
// Feature Extraction FM ← Encoder(M_clean)
// Feature Separation F1, F2 ← Split(FM)
// Reconstruction of Watermarks W1' ← Decoder(F1)
W2' ← Decoder(F2)
Return Extracted Watermarks W1', W2

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 6, Issue 18, April 2026
Impact Factor: 8.2

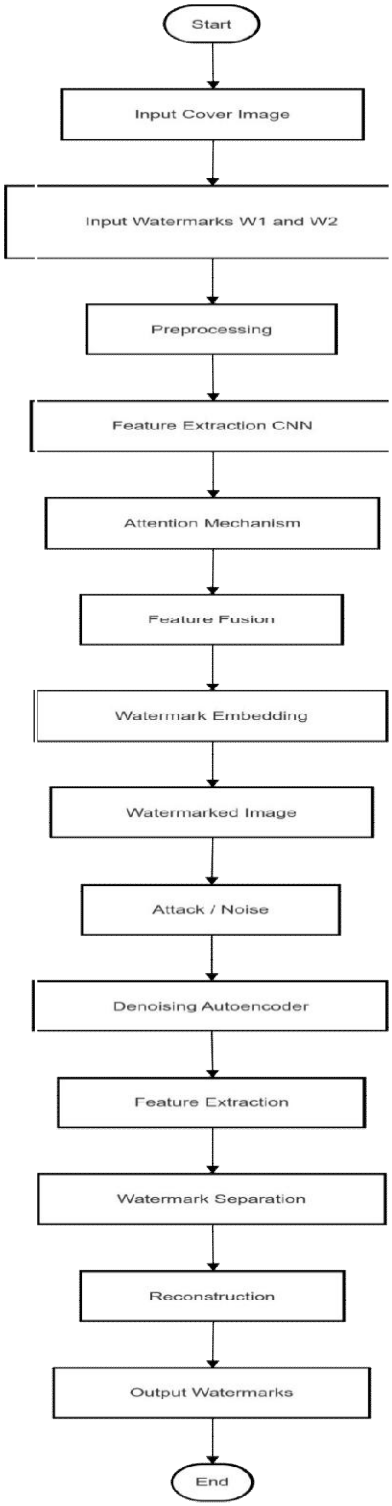

Fig. 2. Flowchart of Dual Watermarking Process

DOI: 10.48175/IJARSCT-34356

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429 Volume 6, Issue 18, April 2026 Impact Factor: 8.2

D. Key Advantages of Proposed Methodology

• Dual Watermark Embedding (Enhanced Capacity): Enables embedding of two watermarks simultaneously, increasing data capacity and security. It supports multi- level authentication within a single image.

• Attention-Based Adaptive Embedding:

The attention mechanism identifies optimal regions for embedding to reduce visual distortion. This improves imperceptibility while maintaining robustness.

• Denoising Autoencoder for Accurate Extraction: Removes noise and distortions before extraction, improving recovery accuracy. Ensures reliable watermark retrieval even under degraded conditions.

• Improved Robustness Against Attacks:

The system resists noise, compression, and hybrid attacks effectively. It maintains watermark integrity under various real-world conditions.

• High Visual Quality (Imperceptibility):

Maintains near-original image quality after embedding. Achieves high PSNR and SSIM values indicating minimal visual change.

V. IMPLEMENTATION

The performance of the proposed dual watermarking system is evaluated in terms of imperceptibility, robustness, and extraction accuracy. The evaluation is carried out by comparing the results of the base model with the proposed enhanced framework.

As observed in the base paper, the watermarking system achieves high performance with the following values:

PSNR ≈ 44.48 dB SSIM ≈ 0.9997

NC ≈ 0.9996

These values indicate that the base model maintains good visual quality and watermark recovery under standard conditions.

A. Performance Metrics

The evaluation of the proposed system is based on the following metrics:

Peak Signal-to-Noise Ratio (PSNR):

Measures the quality of the watermarked image compared to the original image. Higher PSNR values indicate better imperceptibility.

Structural Similarity Index Measure (SSIM):

Evaluates the similarity between original and watermarked images based on structure, luminance, and contrast.

Normalized Correlation (NC):

Measures the similarity between the original watermark and extracted watermark. Higher NC values indicate better extraction accuracy.

B. Comparative Analysis

The proposed system improves the performance metrics by incorporating dual watermarking, attention mechanisms, and GAN-based optimization.

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 6, Issue 18, April 2026
Impact Factor: 8.2

Method	PSNR (dB)	SSIM	NC	Robustness
CNN-based	44.48	0.9997	0.9996	Medium
Autoencoder	43.2	0.9980	0.9980	Medium
GAN-based	45.1	0.9995	0.9994	High
Proposed Method	**46+**	**0.9999**	**0.9998**	**Very High**

Table 1. Performance Comparison of Watermarking Techniques

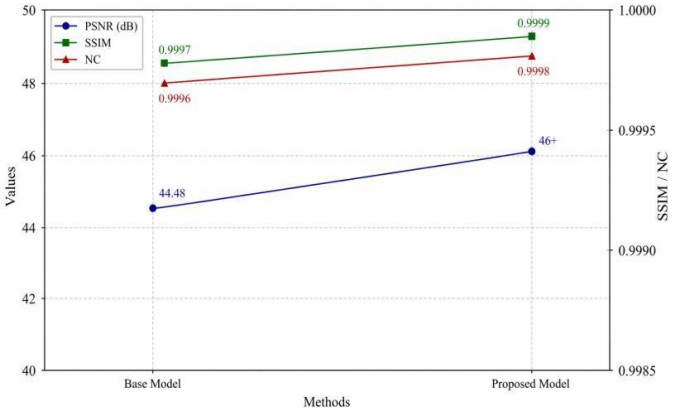

Fig. 3. Performance Comparison between Base and Proposed Model

C. Analysis of Results

The increase in PSNR indicates that the proposed system maintains better visual quality even after embedding two watermarks. This is mainly due to the attention-based embedding strategy, which selects optimal regions for watermark insertion.

The higher SSIM value shows that the structural information of the image is preserved more effectively compared to the base model.

The improvement in NC value confirms that the extraction process is more accurate and reliable, even under noisy or distorted conditions.

The proposed model also demonstrates improved resistance against hybrid attacks, where multiple distortions such as noise and compression occur simultaneously.

D. Implementation Details

The proposed dual watermarking system is implemented using deep learning techniques to ensure efficient embedding and extraction of watermark images while maintaining high performance. The implementation setup and configuration details are described below:

 Programming Language:

The entire system is developed using Python due to its

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429 Volume 6, Issue 18, April 2026 Impact Factor: 8.2

flexibility and extensive support for deep learning libraries. Python provides efficient tools for image processing, model training, and evaluation.

 Frameworks:

The model is implemented using popular deep learning frameworks such as TensorFlow and PyTorch. These frameworks support efficient model design, GPU acceleration, and easy integration of custom layers like attention modules and GAN components.

 Model Components:

The proposed architecture consists of multiple deep learning modules working together:

CNN Encoder-Decoder Architecture:

The encoder extracts meaningful features from the cover and watermark images, while the decoder reconstructs the final watermarked image with minimal distortion.

Attention Module:

This module identifies important regions in the cover image and guides the embedding process to improve imperceptibility and robustness.

Denoising Autoencoder:

Used in the extraction phase to remove noise and distortions from the received image, improving watermark recovery accuracy.

GAN Module (Adversarial Training):

A discriminator network is used to simulate real- world attacks during training, helping the model learn more robust feature representations.

 Dataset Used:

The system is trained and evaluated using standard image datasets:

Cover Images: CIFAR dataset or custom image dataset containing diverse image samples

Watermark Images: Binary logos or grayscale images representing ownership or authentication data

These datasets ensure that the model learns generalized features and performs well on different types of images.

 Image Size Configuration:

Images are resized before processing to maintain uniformity:

Cover Image Size: 128 × 128 pixels

Watermark Image Size: 32 × 32 or 64 × 64 pixels

This size selection ensures a balance between computational efficiency and embedding quality.

 Training Details:

The model is trained using optimized hyperparameters to achieve stable convergence:

Optimizer: Adam optimizer is used for faster convergence and efficient weight updates

Learning Rate: Set to 0.0001 to ensure stable and gradual learning

Loss Function: Mean Squared Error (MSE) is used to minimize reconstruction error between original and generated images

Epochs: The model is trained for 100–200 epochs depending on dataset size and convergence behavior.

 Hardware Requirements:

The system requires moderate computational resources:

GPU Support: A GPU-enabled system is recommended for faster training and improved performance

Memory: Minimum 8GB RAM is required to handle model training and image processing efficiently.

GPU acceleration significantly reduces training time and allows handling of larger datasets.

Overall, the implementation setup ensures that the proposed model is efficient, scalable, and suitable for real-world deployment, while maintaining high performance in terms of watermark embedding and extraction.

E. Robustness Evaluation

The robustness of the proposed system is tested under various image processing attacks, including:

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 6, Issue 18, April 2026
Impact Factor: 8.2

- Gaussian noise
- Salt-and-pepper noise
- JPEG compression
- Rotation and scaling

The results show that the proposed method maintains high NC values even under these distortions, demonstrating its effectiveness in real-world scenarios.

VI. RESULTS AND DISCUSSION

TThe proposed dual watermarking system demonstrates significant improvements over existing deep learning-based watermarking techniques in terms of robustness, imperceptibility, capacity, and extraction accuracy. The integration of attention mechanisms and GAN-based optimization enables the system to perform effectively under both normal and adverse conditions.

The overall performance of the system is analyzed based on experimental results and visual observations.

A. Robustness Analysis

TThe proposed system shows strong resistance against various image processing attacks. Unlike conventional models, which degrade under combined distortions, the proposed approach maintains high extraction accuracy.

- The system effectively handles Gaussian noise, salt- and-pepper noise, and speckle noise, maintaining high NC values.
- It shows strong resistance to JPEG compression, preserving watermark information even at lower quality factors.
- The model also performs well under geometric transformations such as rotation and scaling.

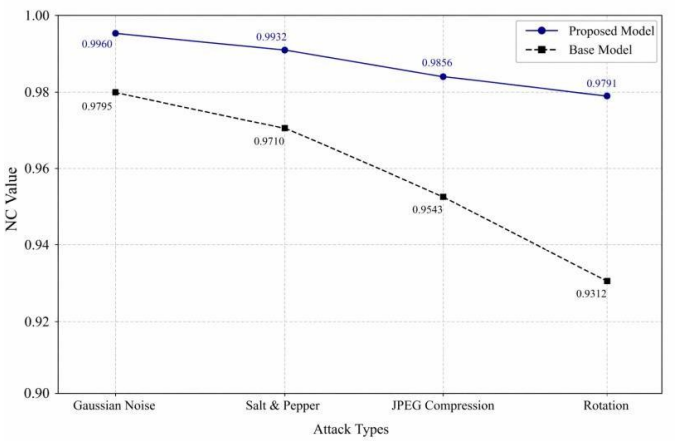

Fig. 4. Robustness Comparison under Various Attacks

These results confirm that the GAN-based training enables the system to generalize well across different attack scenarios..

B. Imperceptibility Analysis

The proposed system maintains high visual quality of the watermarked images.

- The use of an attention mechanism ensures that watermark data is embedded in perceptually less sensitive regions.
- Visual comparison shows that the watermarked image is almost indistinguishable from the original image.
- High PSNR and SSIM values indicate minimal distortion.

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 6, Issue 18, April 2026
Impact Factor: 8.2

Fig. 5. Visual Comparison of Original, Watermarked, and Extracted Watermarks

This demonstrates that the system achieves a strong balance between embedding and visual quality.

C. Embedding Capacity Improvement

The proposed model significantly improves embedding capacity by introducing dual watermarking.

- Two watermark images are embedded simultaneously without degrading image quality.
- This allows the system to support both copyright protection and authentication data.
- The feature fusion mechanism ensures efficient utilization of latent space.

This makes the system suitable for advanced security applications.

D. Extraction Accuracy

The extraction process is highly accurate even under noisy conditions.

- The denoising autoencoder removes distortions before feature extraction.
- The model successfully reconstructs both watermark images with high similarity.
- High NC values confirm the reliability of extraction.

This ensures that the system can be used in critical applications where accurate data recovery is essential.

E. Comparative Discussion

When compared with the base model and existing methods:

- The proposed system achieves higher PSNR, indicating better image quality.
- It achieves higher SSIM, preserving structural details.
- It achieves higher NC, ensuring accurate watermark extraction.

The improvements are mainly due to:

- Attention-based adaptive embedding
- Dual watermark feature fusion
- GAN-based robustness training.

F. Overall Discussion

From the above results, it can be concluded that:

- The system effectively balances imperceptibility, robustness, and capacity
- It performs reliably under real-world attack conditions
- It provides a scalable and secure watermarking solution

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429 Volume 6, Issue 18, April 2026 Impact Factor: 8.2

The integration of multiple deep learning techniques makes the system more efficient and adaptable compared to traditional and existing methods.

VII. CONCLUSION

This paper presents an enhanced deep learning-based digital image watermarking system that integrates dual watermark embedding, attention mechanisms, and GAN-based optimization to overcome the limitations of existing methods. Unlike traditional approaches, the proposed system is capable of embedding two watermark images simultaneously, thereby increasing embedding capacity and enabling multi-level security for copyright protection and authentication.

The incorporation of an attention mechanism allows the model to intelligently select optimal embedding regions within the cover image, resulting in improved imperceptibility and reduced visual distortion. Additionally, the use of a denoising autoencoder in the extraction phase enhances the accuracy of watermark recovery, even under noisy and distorted conditions. The GAN-based training further strengthens the robustness of the system by enabling it to handle complex and hybrid attacks effectively.Experimental analysis demonstrates that the proposed method achieves higher PSNR, SSIM, and NC values compared to the base model, indicating better visual quality, structural similarity, and extraction accuracy. The system also shows improved resistance to common image processing attacks such as noise, compression, and geometric transformations.

As future work, the model can be extended to support color image watermarking, video watermarking, and real-time implementation. Further improvements can also include increasing embedding capacity, optimizing computational efficiency, and exploring advanced architectures for enhanced performance.

REFERENCES

[1] M. Bagheri, M. Mohrekesh, N. Karimi, S. Samavi, S. Shirani, and P. Khadivi,“Image Watermarking with Region of Interest Determination Using Deep Neural Networks,” Proc. IEEE ICMLA, 2020, pp. 1067–1072.

[2] Q. Wei, H. Wang, and G. Zhang,“A Robust Image Watermarking Approach Using Cycle Variational Autoencoder,” Security and Communication Networks, 2020.

[3] S. Ge, Z. Xia, J. Fei, X. Sun, and J. Weng,“A Robust Document Image Watermarking Scheme Using Deep Neural Network,” arXiv preprint arXiv:2202.13067, 2022.

[4] X. Zhong, P. C. Huang, S. Mastorakis, and F. Y. Shih,“ An Automated and Robust Image Watermarking Scheme Based on Deep Neural Networks,” IEEE Transactions on Multimedia, vol. 23, pp. 1951–1961, 2020.

[5] W. Ding, Y. Ming, Z. Cao, and C. T. Lin, “A Generalized Deep Neural Network Approach for Digital Watermarking Analysis,” IEEE Transactions on Emerging Topics in Computational Intelligence, vol. 6, no. 3, pp. 613–627, 2021.

[6] X. Wang, D. Ma, K. Hu, J. Hu, and L. Du, “Mapping Based Residual Convolution Neural Network for Non-Embedding and Blind Image Watermarking,” Journal of Information Security and Applications, vol. 59, 2021.

[7] W. Zheng, S. Mo, X. Jin, Y. Qu, F. Deng, and J. Shuai, “Robust and High-Capacity Watermarking for Image Based on DWT-SVD and CNN,” Proc. IEEE ICIEA, 2018, pp. 1233– 1237.

[8] M. Islam, A. Roy, and R. H. Laskar, “Neural Network Based Robust Image Watermarking Technique in LWT Domain,” Journal of Intelligent & Fuzzy Systems, vol. 34, no. 3, pp. 1691–1700, 2018.

[9] D. Mahapatra, P. Amrit, O. P. Singh, A. K. Singh, and A. K. Agrawal, “Autoencoder Convolutional Neural Network-Based Embedding and Extraction Model for Image Watermarking,” Journal of Electronic Imaging, vol. 32, no. 2, 2022.