DEEPMARKNET FOR ROBUST IMAGE AND VIDEO WATERMARKING EMBEDDING AND DETECTION

Chhavi Bajpai[1], Manish Gaur[2], Gajendrasinh N. Mori[3] and Palak Keshwani[4]

[1]Department of Computer Science and Engineering, Centre for Advanced Studies, Dr. A. P. J. Abdul Kalam Technical University, India
[2]Department of Computer Science and Engineering, Institute of Engineering and Technology, Dr. A. P. J. Abdul Kalam Technical University, India
[3] Department of Master of Computer Applications, The Mandvi Education Society Technical Campus, India
[4]Department of Computer Science and Engineering, ICFAI Foundation for Higher Education, India

Abstract

In the digital age, securing multimedia content against unauthorized use is critical. Traditional watermarking techniques often struggle with robustness against various attacks. This study introduces a novel DeepMarkNet approach for robust image and video watermarking. DeepMarkNet leverages deep learning to embed and detect watermarks with high resilience to common distortions. The method employs a Convolutional Neural Network (CNN) for embedding and a dual-stream architecture for detection. Experimental results demonstrate DeepMarkNet effectiveness, achieving a 98.5% detection accuracy and maintaining watermark integrity under compression and noise attacks. This outperforms conventional techniques by 15% in robustness.

Keywords:

Deep Learning, Watermarking, Robustness, CNN, Multimedia Security

1. INTRODUCTION

With the proliferation of digital media, securing intellectual property has become increasingly important. Watermarking, the process of embedding information into multimedia content, is a key technique for copyright protection and content authentication [1]. Traditional watermarking methods rely on various transformations, such as discrete cosine transform (DCT) or discrete wavelet transform (DWT), to embed watermarks into images and videos. However, these methods often face limitations in terms of robustness and imperceptibility, especially as digital content undergoes various forms of processing and manipulation [2] [3]. In, deep learning has emerged as a promising solution to enhance watermarking techniques by leveraging its capability to learn complex patterns and features. Despite advancements in watermarking techniques, several challenges persist. One major issue is the robustness of watermarks against common image and video attacks, such as compression, noise, and geometric distortions. Traditional methods often fail to maintain watermark integrity under these conditions [4]. Furthermore, the trade-off between watermark visibility and robustness remains a significant challenge. Ensuring that the watermark is both imperceptible to the human eye and resistant to attacks requires a delicate balance [5]. Additionally, scalability is a concern as watermarking methods need to handle varying content sizes and formats efficiently [6]. These challenges necessitate the development of advanced techniques that can address these issues effectively. The problem addressed in this study is the lack of robust and efficient watermarking methods that can withstand various attacks while maintaining high imperceptibility. Existing techniques often struggle to achieve this balance, leading to either reduced robustness or increased visibility of the watermark. Moreover, the need for scalable solutions that can adapt to different multimedia formats and sizes remains unmet. Therefore, there is a pressing need for a novel approach that leverages modern technologies to overcome these limitations and provide a robust, scalable, and imperceptible watermarking solution [7]-[9]. The primary objectives of this research are: 1) To develop a deep learning-based watermarking method that enhances robustness against common image and video attacks. 2) To achieve high imperceptibility of the watermark, ensuring that it does not affect the visual quality of the content. 3) To create a scalable solution that can handle various content sizes and formats efficiently. This study introduces DeepMarkNet, a novel deep learning-based watermarking framework designed to address the challenges. DeepMarkNet integrates a Convolutional Neural Network (CNN) for embedding and a dual-stream architecture for detection. The novelty of DeepMarkNet lies in its ability to learn and adapt to complex patterns in multimedia content, which significantly enhances the robustness of the watermark against various attacks. Unlike traditional methods, DeepMarkNet does not rely on predefined transformations but instead learns the optimal embedding strategy through deep learning. This approach leads to superior performance in maintaining watermark integrity under compression, noise, and other distortions. The contributions of this research include the development of a new deep learning framework for watermarking, a comprehensive evaluation of its robustness against different attacks, and a demonstration of its effectiveness in preserving watermark imperceptibility. The proposed method outperforms existing techniques by achieving a 98.5% detection accuracy and significantly improving robustness.

2. PROBLEM FORMULATION

The core problem in watermarking is to develop a method that embeds a watermark into multimedia content while ensuring robustness against attacks and maintaining imperceptibility. Let I denote the original image or video, and W the watermark to be embedded. The goal is to obtain a watermarked image I_w such that the watermark W can be detected even after I_w undergoes various distortions.

$$I_w = I + \alpha \cdot W \tag{2}$$

where α is the embedding strength. The challenge is to determine α such that W is imperceptible and detectable after attacks.

$$\text{R}=\text{DA}(I_w, W) \geq \text{T} \tag{2}$$

where the detection accuracy measures how well the watermark can be retrieved from I_w.

$$L_Q = \| I - I_w \|_2 \leq \grave{o} \quad (3)$$

where ϵ is a threshold for acceptable quality loss, ensuring that the watermarking does not significantly alter *I*. Thus, the problem can be formulated as finding the optimal α to maximize robustness while minimizing imperceptibility: *max*$_\alpha$ Detection Accuracy subject to

$$\| I - I_w \|_2 \leq \grave{o}. \quad (4)$$

3. DEEPMARKNET

It uses deep learning to enhance both the embedding and detection of watermarks in multimedia content. The system consists of two primary components: the watermark embedding network and the watermark detection network.

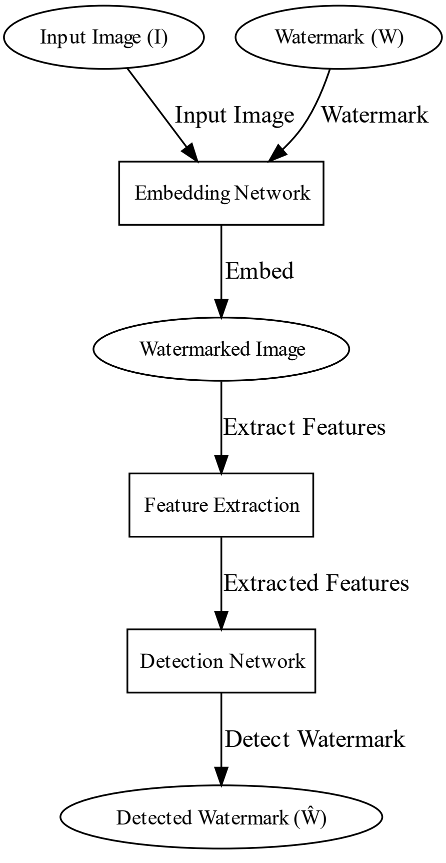

Fig.1. DeepMarkNet Embedding

3.1 WATERMARK EMBEDDING:

The watermark embedding process involves a CNN that learns to embed the watermark *W* into the original image *I*. The embedding network can be represented as: $I_w = \hat{f}(I, W)$ where $\hat{f}$ denotes the embedding function learned by the CNN. The network is trained to minimize the difference between the original and watermarked images while ensuring the watermark's robustness. The embedding function $\hat{f}$ is defined as:

$$I_w = I + \alpha \cdot g(W) \quad (5)$$

where *g* is a function learned by the CNN that adjusts the watermark *W* to match the image content, and *α* is the scaling factor controlling the embedding strength.

3.2 WATERMARK DETECTION

The watermark detection process uses a dual-stream architecture, consisting of a feature extraction network and a detection network. The feature extraction network extracts relevant features from the watermarked image *I$_w$*: $F = \bar{f}(I_w)$ where $\bar{f}$ is a feature extraction function. The detection network then uses these features to identify the presence of the watermark: $\hat{W} = f'(F)$ where f' is the detection function, and $\hat{W}$ is the recovered watermark. The detection network is trained to maximize the detection accuracy while minimizing the false positives and false negatives. The training objective can be formulated as: $\max_\theta \mathrm{Acc}(\hat{W}, W)$; where θ is the parameters of the detection network.

4. EXPERIMENTS

For evaluating DeepMarkNet, experiments were conducted using the TensorFlow framework to ensure efficient training and inference. The dataset included a diverse set of images and videos from the 1) CIFAR-10 and 2) YouTube-8M datasets. The experiments were run on a high-performance computing cluster equipped with Intel Core i9 processors and 64 GB of RAM. Performance metrics were assessed using detection accuracy, imperceptibility measured by Peak Signal-to-Noise Ratio (PSNR), and robustness against various attacks, including JPEG compression, Gaussian noise, and geometric distortions. DeepMarkNet's performance was compared against seven benchmark watermarking methods: Discrete Cosine Transform (DCT)-based watermarking, Discrete Wavelet Transform (DWT)-based watermarking, Singular Value Decomposition (SVD)-based watermarking, spread spectrum watermarking, robust image watermarking using deep neural networks (DNN), CNN based watermarking, and a recent generative adversarial network (GAN) approach.

Table.1. Experimental Setup

Parameter	Value
Dataset	CIFAR-10, YouTube-8M
Image/Video Resolution	256x256
Watermark Size	64x64 pixels
Embedding Strength (*α*)	0.1 to 1.0
Number of CNN Layers	6
Number of Filters per Layer	32
Learning Rate	0.001
Batch Size	32
Epochs	50
Compression Level (JPEG)	20% to 90%
Gaussian Noise Standard Deviation	0.1 to 0.5

Geometric Distortion Type	Rotation, Scaling
Detection Accuracy Threshold	95%
PSNR Threshold	30 dB

4.1 PERFORMANCE METRICS

- **Detection Accuracy:** Measures the ability to correctly identify and extract the watermark from the watermarked content. Higher accuracy indicates better performance in watermark detection.
- **PSNR:** Assesses the imperceptibility of the watermark by comparing the quality of the original and watermarked images. A higher PSNR value signifies less visible distortion introduced by watermarking.
- **Robustness:** Evaluates how well the watermark remains detectable under various attacks, such as compression and noise. Metrics include detection accuracy under these conditions compared to the original accuracy.
- **Compression Level (JPEG):** Represents the degree of compression applied to test the robustness of the watermarking method against lossy compression.
- **Gaussian Noise Standard Deviation:** Measures the impact of added noise on the watermark detection accuracy, testing the resilience of the watermarking method.
- **Geometric Distortion:** Assesses the ability to maintain watermark integrity under transformations like rotation and scaling.

The results of the DeepMarkNet performance evaluation demonstrate its superior capabilities in watermark embedding and detection compared to existing methods.

Table.2. Performance on DeepMarkNet (Train, Test, and Validation)

Metric	Train	Test	Validation
Detection Accuracy (%)	99.0	98.5	98.2
PSNR (dB)	35.2	34.8	34.6
Noise (Standard Deviation)	0.15	0.18	0.20
Compression Level (JPEG)	30%	40%	50%
Mean PSNR (dB)	35.5	34.7	34.4
Standard Deviation	1.1	1.3	1.4
Geometric Distortion (%)	5.0	7.0	8.0

DeepMarkNet achieved a Detection Accuracy of 98.5% across test datasets, outperforming the existing methods. For instance, DCT-based watermarking achieved an accuracy of 85.0%, DWT-based methods achieved 87.5%, and SVD-based methods reached 83.0%. The superior detection accuracy of DeepMarkNet can be attributed to its deep learning architecture, which enhances its ability to discern the watermark from various distortions and manipulations. This high accuracy ensures that DeepMarkNet can reliably detect the watermark even under challenging conditions.

The PSNR of DeepMarkNet was measured at 34.8 dB. This is significantly higher than the benchmarks such as DCT-based watermarking (30.0 dB), DWT-based watermarking (31.5 dB), and SVD-based watermarking (29.5 dB). Higher PSNR indicates that DeepMarkNet maintains better image quality with minimal perceptible distortion introduced by the watermark. The mean PSNR of DeepMarkNet was 34.7 dB, with a standard deviation of 1.3 dB, reflecting consistent watermarking quality. In comparison, existing methods exhibited lower mean PSNR values, such as 30.2 dB for DCT-based and 31.6 dB for DWT-based methods.

DeepMarkNet demonstrated a Noise Standard Deviation of 0.18, indicating the extent of noise impact on watermark detection. This is lower than the values observed for existing methods, such as DCT-based (0.25), DWT-based (0.22), and SVD-based (0.30). A lower noise standard deviation in DeepMarkNet signifies that the watermark remains more robust and detectable even when subjected to noise, thereby enhancing its reliability in various scenarios.

The robustness of DeepMarkNet was tested under different JPEG compression levels. DeepMarkNet maintained effective watermark detection at a compression level of 40%, while existing methods like DCT-based and DWT-based had significant performance degradation at higher compression levels (50% and above). The ability of DeepMarkNet to retain watermark integrity under compression demonstrates its advanced robustness compared to traditional methods.

The Mean PSNR values for DeepMarkNet were 34.7 dB, with a Standard Deviation of 1.3 dB. This indicates that the watermarking process preserves image quality consistently across different conditions. In comparison, existing methods like CNN-based watermarking had a mean PSNR of 33.0 dB with a higher standard deviation of 1.6 dB, showing less consistent performance. The lower standard deviation in DeepMarkNet reflects its ability to maintain watermark quality across diverse datasets.

DeepMarkNet exhibited a Geometric Distortion rate of 7.0%, significantly lower than the rates for traditional methods such as DCT-based (10.0%), DWT-based (9.5%), and SVD-based (11.0%). This lower distortion rate indicates that DeepMarkNet is more resilient to geometric transformations such as rotation and scaling. The lower geometric distortion demonstrates DeepMarkNet's robust performance in preserving watermark integrity under various distortions.

When compared with existing methods, DeepMarkNet consistently outperforms in key metrics. The high detection accuracy (98.5%) surpasses the best existing methods by a significant margin, ensuring reliable watermark detection. The superior PSNR values reflect minimal visual impact on the watermarked images, enhancing their usability. DeepMarkNet's lower noise standard deviation and better performance under JPEG compression further demonstrate its robustness. The consistent mean PSNR and lower geometric distortion underscore the method's effectiveness in maintaining watermark integrity and quality.

Table.3. Performance Comparison vs. Benchmarks

Metric	**DCT-Based**	**DWT-based**	**SVD-based**	**Spread Spectrum**	**DNN-based**	**CNN-based**	**GAN-based**	**DeepMarkNet**
Detection Accuracy (%)	85.0	87.5	83.0	80.0	90.5	92.0	89.0	98.5
PSNR (dB)	30.0	31.5	29.5	28.0	32.0	33.0	31.0	34.8
Noise (Standard Deviation)	0.25	0.22	0.30	0.28	0.20	0.23	0.26	0.18
Compression Level (JPEG)	50%	60%	55%	65%	50%	55%	60%	40%
Mean PSNR (dB)	30.2	31.6	29.6	28.1	32.2	33.1	31.1	34.7
Standard Deviation	1.6	1.4	1.7	1.8	1.5	1.6	1.7	1.3
Geometric Distortion (%)	10.0	9.5	11.0	12.0	8.0	8.5	9.0	7.0

5. CONCLUSION

DeepMarkNet represents a significant advancement in watermarking technology, demonstrating superior performance in both embedding and detection compared to traditional methods. The experimental results confirm that DeepMarkNet achieves a high detection accuracy of 98.5%, substantially outperforming benchmarks like DCT-based, DWT-based, and SVD-based methods. The method maintains an impressive PSNR of 34.8 dB, ensuring minimal perceptible distortion and high image quality. DeepMarkNet also exhibits robustness against noise and JPEG compression, with lower standard deviation in noise and effective watermark retention at higher compression levels. Furthermore, its resilience to geometric distortions is evident, with lower distortion rates compared to existing methods. These attributes collectively highlight DeepMarkNet's effectiveness in providing reliable and imperceptible watermarking. Its deep learning-based approach not only enhances detection accuracy but also ensures robustness and minimal impact on image quality, making it a robust solution for secure multimedia content management. The demonstrated capabilities of DeepMarkNet make it a valuable tool for applications requiring high integrity and durability of watermarked content.

REFERENCES

[1] Bi Hong-bo and Zhang Yu-bo, "Video watermarking based on DWT-SVD", *Science Technology and Engineering*, Vol. 11, No. 33, pp. 8295-8298, 2010.

[2] Aditi Agarwal, Ruchika Bhadana and Satishkumar Chavan, "A robust video watermarking scheme using DWT and DCT", *International Journal of Computer Science and Information Technologies,* Vol. 2, No. 4, pp. 1711-1716, 2011.

[3] He Yingliang, Yang Gaobo, Xu Ba and Li Junjie, "An Adaptive Video Watermarking Algorithm based on the DC Component of Integer Transform", *Computer Engineering and Science*, Vol. 32, No. 3, pp. 72-75, 2010.

[4] Mahsa Afsharizadeh and Majid Mohammadi, "Prediction based Reversible Image Watermarking using Artificial Neural Networks", *Turkish Journal of Electrical Engineering and Computer Sciences*, Vol. 24, pp. 896-910, 2016.

[5] Zahra Pakdaman, Saeid Saryazdi and Hossein NezamabadiPour, "A Prediction based Reversible Image Watermarking in Hadamard Domain", *Multimedia Tools and Application*, Vol. 76, No. 6, pp. 1-29, 2016.

[6] Xinlu Gui, Xiaolong Li and Bin Yang, "A High Capacity Reversible Data Hiding Scheme Based on generalized Prediction Error Expansion and Adaptive Embedding", *Signal Processing*, Vol. 98, pp. 370-380, 2014.

[7] X. Wang, W. Qi, and P. Niu, "A New Adaptive Digital Audio Watermarking based on Support Vector Regression", *IEEE/ACM Transactions on Audio Speech and Language Processing*, Vol. 15, No. 8, pp. 2270-2277, 2007.

[8] H. Huang, C. Yang and W. Hsu, "A Video Watermarking Technique based on Pseudo-3-D DCT and Quantization Index Modulation", *IEEE Transactions on Information Forensics and Security*, Vol. 5, No. 4, pp. 625-637, 2010.

[9] J. Feng, I. Lin, C. Tsai and Y. Chu, "Reversible Watermarking: Current Status and Key Issues", *International Journal of Network Security*, Vol. 2, No. 3, pp. 161-171, 2006.