Invisible Watermark for Copyright Protection

Cheng-Han Chiang, Yun-Hsuan Cheng, Zong-Yun Li
University of Illinois at Urbana Champaign
chc11@illinois.edu, yhcheng3@illinois.edu, zyl2@illinois.edu

1 Introduction

The problem this project aims to solve is the need for a robust and efficient method to authenticate digital images and protect them from unauthorized use. In the digital age, images are easily duplicated and shared, leading to copyright infringement and misuse. A reliable solution is needed to embed information in images to prove their authenticity and ownership.

The proposed solution is to implement a watermarking technique based on the Fast Fourier Transform (FFT) to ensure digital image authentication and copyright protection. This technique involves the insertion of hidden data (watermarks) into the frequency domain of an image using FFT. These watermarks are imperceptible to the human eye but can be detected and verified later to establish the image's authenticity and the copyright holder.

2 Related Work

In the initial phase, we extensively reviewed blogs [1] and [2] to gain a comprehensive understanding of the two-dimensional Fourier transform and its applications in digital watermarking. Furthermore, we referred to the GitHub repositories provided by [3] and [4], which offered valuable insights into the practical implementation aspects. As we delved deeper into the subject, we discovered an alternative approach for digital watermarking outlined in [5]. This methodology leverages a combination of Discrete Wavelet Transform (DWT), Discrete Cosine Transform (DCT), and Singular Value Decomposition (SVD) to embed watermarks by modifying singular values.

3 Methods

The WatermarkCore class provides functionalities for embedding and extracting watermarks within digital images. The approach utilizes Discrete Cosine Transform (DCT), Singular Value Decomposition (SVD), and random strategies for watermark embedding and extraction. The process involves several stages:

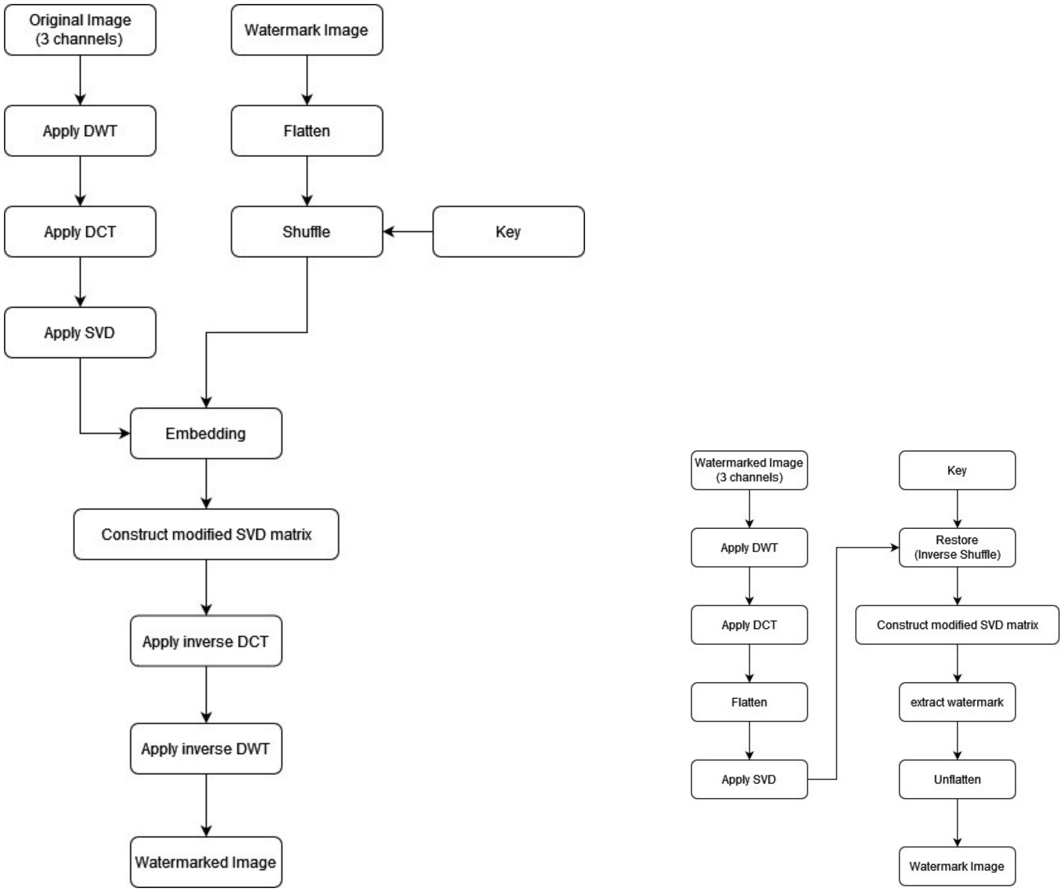
3.1 Image Preprocessing

- DWT Transformation: DWT (Discrete Wavelet Transform) is applied to the image, separating the CA (approximation) and HVD (horizontal-vertical-detail) components

3.2 Embedding Process

- Initialization and Information Allocation: Watermark bits are assigned to specific blocks based on a shuffling index. The watermark bits are embedded using a strategy involving DCT, SVD, and manipulation of singular values. The blocks' frequency domain data is modified based on the watermark bits to embed information.
- Inverse Transformations and Reassembly: The modified CA components and HVD components are combined using inverse DWT to reconstruct the modified image.

Preprint. Under review.



(a) Flow Chart for Embedded Process

(b) Flow Chart for Extract Process

Figure 1: Flow Charts for Embedded and Extract Processes

3.3 Image Post-Embedding Attack Fusion

- Strategy deploys post-embedding attacks for refining subversion in digital image manipulation.
- Fusion of crop, noise, rotation, and Gaussian attacks is orchestrated post-watermark embedding.
- Crop attack selectively trims significant regions, complicating watermark extraction.
- Noise, rotation, and Gaussian attacks introduce perturbations, spatial warping, and blur, masking the embedded watermark.

3.4 Extraction Process

- Reading Image and Initialization: The image is read, and its dimensions are processed to initialize the extraction process. Random strategies and block indices are created for extraction.
- Block-wise Watermark Extraction: Watermark bits are extracted from individual blocks using the DCT, SVD, and statistical analysis of singular values.
- Aggregation and Decoding: Extracted bits are aggregated and decoded to retrieve the watermark information.
- Include a threshold while generating the extracted watermark, particularly as our watermark operates in grayscale. The application of a threshold helps eliminate some noise from the extracted watermark.

4 Result

The results are presented as follows: initially, we attempted to embed various watermarks. Starting with the telephone icon (2), we subjected it to several different attacks (3). Subsequently, to showcase the algorithm's robustness, we employed a QR code as our watermark. Prior to embedding, we scanned the QR code, which directed us to the wiki page. Following the embedding and extraction processes, we were delighted to find that we could still successfully scan the extracted QR code from the image!

Table 1: Comparison of Original image and embedded image (telephone)

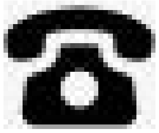
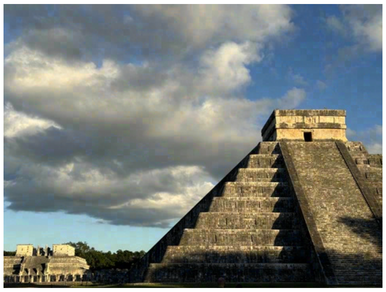
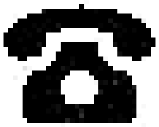
	Embedded image	Watermark
Original picture		
Embedded picture		

Table 2: Comparison of Original image and embedded image (QRcode)

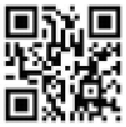
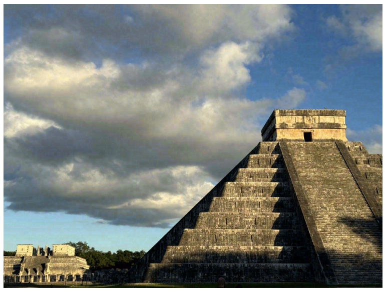
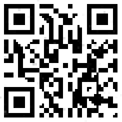
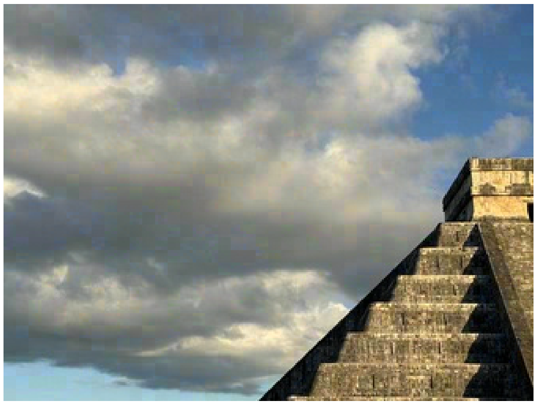
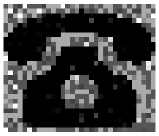
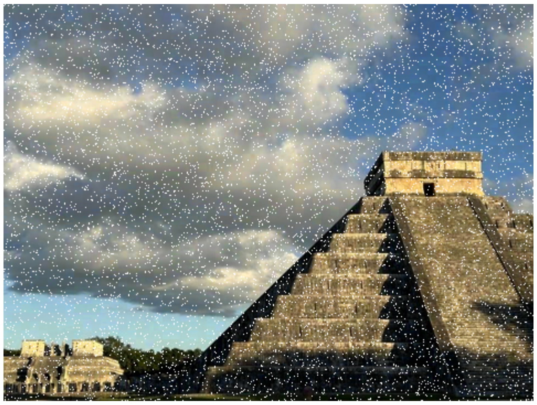
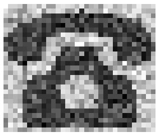
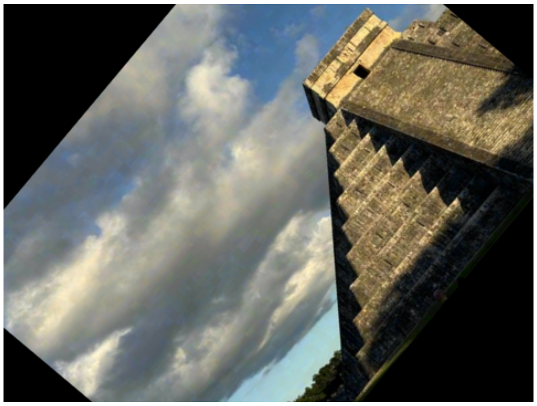
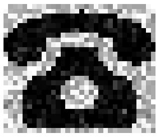
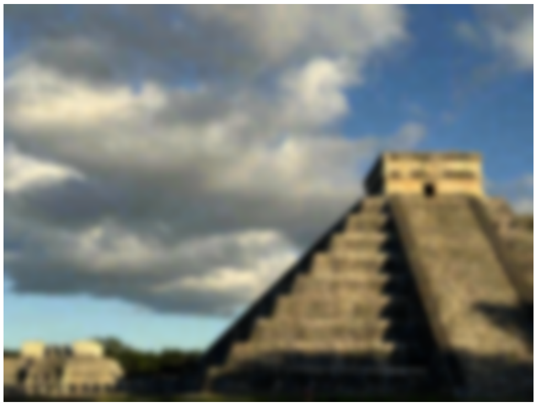
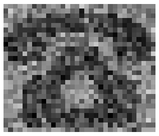
	Embedded image	Watermark
Original picture		
Embedded picture		

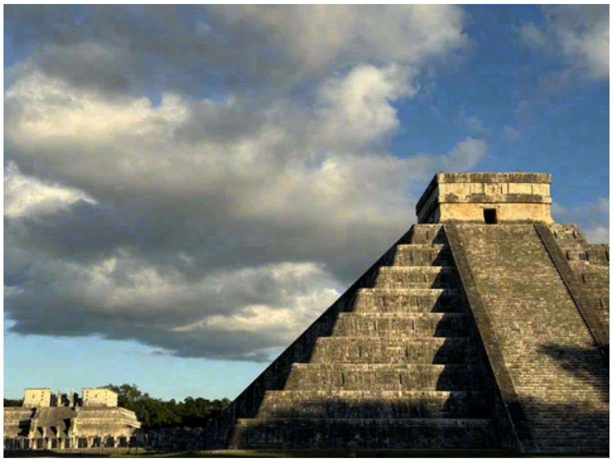
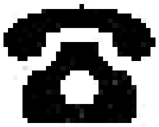
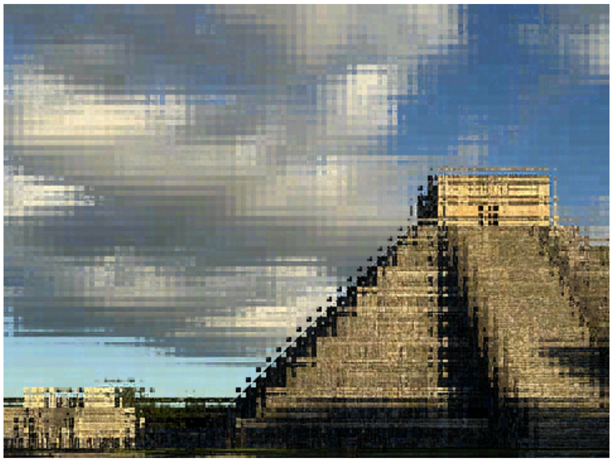
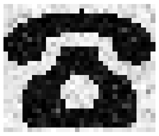
Table 3: Comparison between different attacks

	Embedded image	Watermark
Crop Attack		
Noise Attack		
Rotate attack		
Gaussian attack		

5 Discussion

Upon implementing FFT as our first attempted method, we successfully extracted the watermark from the embedded image, albeit with more noise compared to DCT. However, this approach resulted in a blurred image that didn't meet our objective of preserving the original image's appearance in the watermarked version. You can find the FFT results in 4.

Table 4: Comparison of DCT embedded image and FFT embedded image

	Embedded image	Watermark
DCT Embedded image		
FFT Embedded image		

As mentioned in related work, we discovered an alternative approach for digital watermarking outlined in [5]. This methodology leverages a combination of DWT, DCT, and SVD to embed watermarks by modifying singular values.

In our experiment, DCT might have shown advantages over the FFT due to several reasons:

- **Perceptual Similarity**: DCT has been observed to be more similar to the human visual system compared to FFT. Certain DCT coefficients represent higher frequency changes and edge information, which is closer to how humans perceive visual details.
- **Block-based Structure**: DCT is commonly applied in blocks and exhibits better properties for processing localized image regions. In certain applications, this block-based approach might offer advantages in watermark embedding and extraction.
- **Reduced Artefacts**: In some scenarios, DCT-based watermarking might result in fewer artefacts or visual distortions compared to FFT-based approaches, especially in preserving image quality during watermark embedding.

The WatermarkCore class, employed in the implementation, utilizes a combination of DWT, DCT, SVD, and random strategies for watermark embedding and extraction. The process involves image preprocessing through Discrete Wavelet Transform (DWT) and an embedding process encompassing initialization, information allocation, and inverse transformations for reassembly. Subsequently, a post-embedding attack fusion strategy, involving crop, noise, rotation, and Gaussian attacks, is introduced to fortify the resilience of the embedded watermarks against extraction attempts.

Among the attacks, the algorithm performs better with crop attacks and rotate attacks. However, noise attacks and Gaussian attacks tend to result in a significantly noisier extracted watermark compared to the former.

Crop attacks and rotate attacks primarily involve altering the spatial location or orientation of the image. In these cases, the algorithm might still be able to retrieve the watermark effectively as long as the structure of the image remains relatively intact. Hence, the algorithm performs relatively well under these conditions.

On the other hand, noise attacks and Gaussian attacks introduce additional elements or modify pixel values within the image, causing more significant alterations in the pixel values and structure of the image. As a result, these attacks might lead to a more distorted and noisy image, making it challenging for the algorithm to accurately extract the watermark. This increased level of distortion can consequently result in a noisier extracted watermark compared to the relatively less disruptive crop and rotate attacks.

In crop and rotation attacks, it's essential to restore the attacked image to its original size. Our algorithm heavily relies on the original image size and the pixel locations. Consequently, we encounter challenges when handling resize attacks as we struggle to accurately resize the image and position it correctly within the original picture. As a result, we face difficulties extracting the watermark in such cases.

We do not require any recovery process to extract the watermark in the case of Noise and Gaussian attacks because these attacks do not alter the image size.

The extraction process involves reading the image, initialization, block-wise watermark extraction using DCT, SVD, and statistical analysis of singular values, and finally, aggregation and decoding of extracted bits. The holistic approach presented in this report offers a comprehensive solution to digital image authentication and copyright protection.

6 Conclusions

In summary, this project tackles the crucial need for a strong and effective method to validate digital images and shield them from unauthorized use. In today's digital era, image duplication and widespread dissemination have led to rampant copyright infringement and misuse. The proposed solution outlined in this report employs a Fourier-based watermarking technique to embed imperceptible watermarks, ensuring authentication and protection of digital images.

Through the integration of advanced techniques and a strategic post-embedding attack fusion, the proposed methodology enhances the security and reliability of watermarking in the digital realm. As the landscape of digital watermark subversion continually evolves, this project contributes to the ongoing efforts to strengthen image authentication methodologies and protect the intellectual property rights associated with digital images. Future work may focus on further refining the technique, exploring additional security measures, and adapting to emerging challenges in the dynamic field of digital image processing and protection.

7 Statement of individual contribution

- **Zong-Yun Li** (zyl2): Implement one kind of watermark method, attack, change to DFT, project proposal writing
- **Yun-Hsuan Cheng** (yhcheng3): Implement one kind of watermark method, change to DFT, watermark selection, final report writing
- **Cheng-Han Chiang** (chc11): Implement one kind of watermark method, attack, watermark selection, progress report writing

We write the code independently and integrate it in meetings. We use Discord or meet at Grainger to talk to each other and share information. Regularly meet on Friday.

References

[1] Natuki. 2d-fft domain transformation blog; https://blog.csdn.net/natukiaaa/article/details/120259449, 2021.

[2] Jeremy Kun. The two-dimensional fourier transform and digital watermarking; https://jeremykun.com/2013/12/30/the-two-dimensional-fourier-transform-and-digital-watermarking/, 2013.

[3] linyacool. The two-dimensional fourier transform and digital watermarking; https://github.com/linyacool/blind-watermark, 2018.

[4] j2kun. fft-watermark github; https://github.com/j2kun/fft-watermark, 2014.

[5] Md. Maklachur Rahman. A dwt, dct and svd based watermarking technique to protect the image piracy. *International Journal of Managing Public Sector Information and Communication Technologies*, 4(2):21–32, June 2013.