Team 26

Title: Image Tampering Detection Based on Steganographic Watermarking

---

Names of team members

Name	Email
Peng Chen	pec020@ucsd.edu
Yifan Peng	yip009@ucsd.edu
Jianxiao Cai	jic154@ucsd.edu

Goal/Objective of the project

In today's digital landscape with the surge in digital media use, image tampering has become a pressing issue. This is a significant concern for fields like digital forensics, journalism, and security, where image authenticity is crucial. Whether it's preventing misinformation or safeguarding the integrity of visual data, the ability to reliably detect tampered images is essential for maintaining trust in digital content.

Our project aims to develop a solution that enhances the detection of tampered images by embedding invisible watermarks. These watermarks act as indicators, revealing any alterations made to the image. It provides a proactive method to verify image integrity, helping users identify manipulated visuals and reinforcing trust in what they see online.

Previous works

People have put a lot of effort into figuring out ways to detect when an image has been tampered with, and over time, several techniques have been developed to spot and locate any changes made. Generally, these techniques fall into two main categories: active and passive approaches.

Active Approaches

Active methods are all about embedding extra information into the image right when it's created. This way, when we need to verify the image's authenticity later, we have something to reference. Here are a few common active methods:

- Digital Watermarking: This involves adding a watermark into the image, which can be either visible or hidden. Later, this watermark can be used to check for tampering. The watermark can be fragile (detects any change) or robust

(resistant to minor modifications). For example, some researchers have developed watermarking systems that are not only hard to spot but also strong enough to survive resizing or compression without losing their effectiveness.
- Digital Signatures: In this approach, a unique “signature” is created based on the image’s content. This signature is stored separately, so any modification to the image would cause a mismatch with the original signature, flagging potential tampering.

Passive Approaches

Passive approaches don’t require any pre-added information. Instead, they analyze the image’s natural properties and look for anything that seems off. Here’s how they generally work:

- Pixel-Based Methods: These look for inconsistencies at the pixel level. For instance, Error Level Analysis (ELA) checks for variations in compression across different parts of the image. An altered section might show different compression levels, hinting at tampering.
- Format-Based Methods: Some methods analyze the format and compression of an image. For example, inconsistencies in JPEG quantization tables or signs of double compression can reveal that an image has been altered.
- Camera-Based Methods: These techniques focus on artifacts introduced by the camera itself, like sensor noise patterns or lens distortions. Any irregularities in these unique “fingerprints” can indicate tampering.
- Deep Learning-Based Methods: In recent years, machine learning has been applied to tamper detection. With enough training, neural networks can automatically learn to detect signs of tampering. This approach has shown promising results, but it does require large datasets and considerable computing power.

Challenges in Current Methods

While these techniques have made huge strides, they aren’t without their limitations. Some of the main challenges include:

- Balancing Robustness and Quality: In watermarking, for example, it’s tough to make the watermark robust enough to survive edits without making it visible.
- Complexity and Resources: Methods like deep learning are effective but require extensive labeled data and computational power, making them less practical in resource-constrained environments.

- Generalization: Many methods are tailored to specific types of tampering or image formats, limiting their effectiveness in real-world scenarios where tampering techniques vary widely.

Proposed work

Our goal is to build a tamper detection system that's both reliable and easy to use. The idea is simple: we'll embed a subtle but robust watermark into images, which acts like a unique digital fingerprint. This watermark will stay with the image no matter where it goes, helping us spot any changes made to it later. If anyone tampers with the image, the system will pick up on it by finding mismatches in the watermark.

The system has three main parts: **Watermark Embedding, Tamper Detection, and Analysis & Output.**

1. **Watermark Embedding**

Here's where we add that unique, hidden watermark to each image. To do this, we'll use steganography—a method that subtly embeds this fingerprint directly into the image pixels. It's completely invisible to the naked eye, but resilient enough to stay intact even if the image goes through standard processing, like resizing or compressing. This way, the watermark doesn't interfere with the image quality but is robust enough to withstand basic edits.

2. **Tamper Detection**

If someone tries to modify or tamper with the image, this part of the system kicks in. When we check the image, the system extracts the hidden watermark and compares it to the original. Any differences between the two watermarks will reveal tampered areas. It's a bit like comparing fingerprints—any mismatch tells us that something's changed.

3. **Analysis & Output**

Once we detect tampering, we'll dig deeper to pinpoint exactly where and how much the image was altered. This analysis can then be shared in a few ways: by visually marking the altered regions on the image itself or generating a detailed report. This can be especially useful for legal, security, or media purposes, where verifying the authenticity of an image is crucial.

Expected Results

- High Robustness: The watermark should stay intact through common edits like resizing, compressing, and even light touch-ups.
- Accurate Tamper Detection: We want the system to go beyond simply flagging an image as tampered—it should pinpoint exactly where any alterations occurred.
- Minimal Impact on Image Quality: The watermark should be invisible to the viewer, ensuring the image's original appearance is preserved.

Potential Challenges

- Balancing Robustness and Quality: It's a fine line to walk—creating a watermark that's resilient to edits without affecting image quality. If the watermark becomes visible, it loses its subtlety, which we want to avoid.
- Handling Complex Edits: Some advanced edits or highly skilled tampering techniques might be challenging for the system to detect accurately, especially when the alterations blend seamlessly with the original content.
- Processing Speed: Extracting and comparing watermarks in real-time, especially with large or multiple images, could be resource intensive. We'll need to find ways to make this process as efficient as possible without sacrificing accuracy.

Timeline

Step	Main Task	Timeline
Research & Design	Study existing methods, define project goals, and design system architecture and algorithms.	Week 1
Implementation	Develop the watermark embedding and tamper detection modules.	Week 2-3
Testing & Optimization	Test the system against different tampering techniques and optimize for robustness and efficiency	Week 4-5
Documentation & Presentation	Analyze results, Document findings, finalize project report, and prepare the presentation	Week 6