



A Peer Reviewed Journal

ISSN : 2456-1363

International Journal of Scientific Research & Growth

A multidisciplinary journal for empowering the research

Examining Image Tampering: Surveying Methods for Precise Localization of Alterations

¹Dr. P.L. Verma ²Dr. Shivendra Kumar Dwivedi

¹Professor, Department of Physics, Govt. Vivekananda PG College, Maihar

²Assistant Professor, Department of Computer Science, Sharda Mahavidyalaya, Sarlanagar

Article Info

Received 05 January 2023

Revised form 20 January 2023

Accepted for publication 25 February 2023

Abstract

The authenticity and integrity of digital images is a critical and challenging research problem. Powerful image editing tools like Adobe Photoshop and PaintShop Pro enable the creation of highly convincing forged or tampered images for various malicious purposes. Analyzing and reliably distinguishing tampered images from authentic originals is an extremely difficult task due to the complex nature of images and the sophistication of modern tampering techniques. This paper presents a comprehensive survey and overview of current state-of-the-art methods for precisely localizing and detecting tampering in digital images through comparative analysis and image authentication mechanisms.

A wide range of approaches are covered, including fragile watermarking schemes that embed imperceptible data for tamper detection, feature extraction and machine learning classifiers to identify statistical inconsistencies, double compression artifact analysis for JPEG images, geometric and photometric inconsistency detection, and more. The principles, algorithms, advantages and limitations of each major technique are discussed in detail. Particular focus is given to recent developments in using robust hashing, human perceptual models, invariant feature point matching, and block-level tampering localization through selective authentication data embedding.

The paper aims to provide an in-depth yet accessible technical review that can serve as a reference for researchers working on trustworthy image forensics and authentication. Challenges, open problems and promising directions for future research in this field are also highlighted. The ultimate goal is to advance toward more accurate, comprehensive and efficient algorithms capable of reliably determining image integrity and precisely localizing any malicious tampering, thereby enhancing trust in digital image sources.

I. Introduction

Powerful publicly available image processing software packages such as Adobe Photoshop or PaintShop Pro make digital forgeries a



A Peer Reviewed Journal

reality. Feathered cropping enables replacing or adding features without causing detectable edges. It is also possible to carefully cut out portions of several images and combine them together while leaving barely detectable traces. In the past, several techniques based on data hiding in images have been designed as a means for detecting tampering.

The redundancy of images can be utilized to insert some additional information for the purpose of detecting changes and for image authentication. If the inserted watermark is fragile so that any manipulation of pixels will disturb its integrity, one can easily detect the tampered areas by checking for presence of this fragile watermark. One of the first techniques used for detection of image tampering was based on inserting check-sums of gray levels determined from the seven most significant bits into the least significant bits (LSB) of pseudo-randomly selected pixels [1]. This technique provides very high probability of tamper detection, and it can be implemented in such a manner that creating forgeries from one or multiple authenticated images is highly unlikely without a secret key. Yeung and Wong [10, 11] generate a key-dependent binary valued function f , $f: \{0, 1, \dots, 255\} \rightarrow \{0, 1\}$, that maps integers from 0 to 255 to either 1 or 0. The gray scales are perturbed to satisfy the expression $L(i,j) = fg(g(i,j))$ for each pixel (i,j) . Error diffusion is further employed to preserve the original colors. The image authenticity is easily verified by checking the relationship $L(i,j) = fg(g(i,j))$ for each pixel (i,j) . There are some obvious advantages of this approach. First, the logo can visually represent a particular authentication device or software. By comparing the original logo with the recovered one, one can visually inspect the integrity of the image. Second, the authentication watermark is embedded not only in the LSBs of the image but somewhat deeper (± 5 gray scales).

This makes it more secure and harder to remove. Third, the method is fast, simple, and amenable to hardware implementation. This makes it very appealing for various applications. In this report, we show that if the same logo and the same image key are used for watermarking more than one image, it is typically possible to recover a large portion of the binary function, and subsequently the binary logo [10-13]. Van Schyndel et al. [2] modify the LSB of pixels by adding extended m-sequences to rows of pixels. For an $N \times N$ image, a sequence of length N is randomly shifted and added to the image rows. The phase of the sequence carries the watermark information. A simple cross-correlation is used to test for the presence of the watermark. Wolfgang and Delp [3] extended van Schyndel's work and improved the localization properties and robustness.

They use bipolar m-sequences of -1 's and 1 's arranged into 8×8 blocks and add them to corresponding image blocks. The watermark presence can be evaluated using classical correlation. The fact that the watermark has some robustness properties can be used to quantify the degree of tampering. Zhu et al. [4] propose two techniques based on spatial and frequency masking. Their watermark is guaranteed to be perceptually invisible, yet it can detect errors up to one half of the maximal allowable change in each pixel or frequency bin depending on whether frequency [5] or spatial [6] masking is used. The image is divided into blocks and in each block a secret random signature (a pseudo-random.



A Peer Reviewed Journal

II. Literature Review

The widespread proliferation of powerful digital image editing tools has made it increasingly easy to create highly convincing forged or manipulated images. Detecting and localizing such tampering is critical for ensuring image integrity and authenticity across many domains like news media, criminal investigations, intelligence, evidence documentation, and more. As a result, image forensics has become a vital area of research, with numerous techniques proposed over the past few decades.

This literature review provides a comprehensive survey of major methods for tamper detection and localization in digital images. One of the earliest approaches was based on fragile watermarking, where authentication data in the form of checksums or signatures derived from the image content is embedded into the image file itself (Fridrich & Goljan, 1999; Van Schyndel et al., 1994; Wolfgang & Delp, 1996). Any modification to the image would disturb the fragile watermark, thereby indicating tampering. These methods inserted the watermarks into the least significant bits of image pixels or by adding pseudorandom noise patterns to the image in the spatial or transform domains. While effective at detecting tampering, they could not precisely localize the tampered regions.

To enable localized tampering detection, block-wise fragile watermarking techniques were later developed (Fridrich, 1998; Chuang et al., 2013; Yang & Huang, 2004). These divided the image into blocks and embedded quantized transform coefficients or robust watermarks into each block to authenticate it. Upon detection of a tampered block, the embedded data could be used to attempt recovery of the original content.

Qi et al. (2009) generated complementary edge-based and invariance watermarks from the image wavelet coefficients to distinguish malicious and non-malicious modifications. While providing localization, the fragile nature of these schemes limited their robustness against compression or incidental processing.

As an alternative to authentication data embedding, many techniques relied on extracting intrinsic image features and applying machine learning for tampering detection (Zhang, 2009; Shrishail Math & Tripathi, 2011). Common features included Markov statistics, wavelet transform coefficients, image quality metrics like blur/noise, and more. Both universal classifiers trained on image datasets as well as one-class techniques modeling the statistics of an individual image have been explored.

Dirik (2009) detected tampering from inconsistencies in demosaicing artifact patterns caused by the color filter arrays in digital cameras. With the widespread use of JPEG compression, many forensic techniques targeted detection of double compression artifacts as evidence of tampering in JPEG images (Lin et al., 2009). These analyzed the quantization coefficients and statistical distributions to identify periodic artifacts indicating a compression before the final encoding. Such artifacts could only exist if portions of the image had been pasted or spliced from another source prior to re-compression.

Geometric and photometric inconsistencies across regions in an image provided another avenue for tampering detection.



A Peer Reviewed Journal

Techniques extracted invariant feature points/descriptors (Beis & Lowe, 1997; Yu, 2009) or computed local noise/blur estimates to identify spliced regions with inconsistent properties (Fridrich et al., 2003; Chen et al., 2008). Face detection and illuminant color estimation were leveraged to expose splicing of human subjects across different lighting conditions (Farid, 2009; Kavitha & Priyatha).

Several robust image hashing schemes computed compact signatures from the image content to enable integrity verification and tampering detection (Venkatesan et al., 2000; Chetty & Singh, 2010; Krawczyk et al., 2007). These used wavelet-based feature extraction, random feature projection, or models like bag-of-features to generate representations invariant to allowable transformations like compression or filtering. Changes in the hash indicated content tampering. Yang (2012) matched clustered high-variation feature points between images to detect and localize tampering.

More recent work has focused on developing hybrid schemes combining multiple tampering cues and models. Zhou et al. (2018) fused results from pixel co-occurrence, frame departure, and benford's law features using a Markov random field. Cozzolino et al. (2015) combined camera-based, physics-based, and coding-based component forensic analyzers with a machine learning merger. Such ensemble approaches leveraging complementary tampering traces have shown improved detection accuracy over individual component techniques.

Some key challenges that persist include reliably distinguishing malicious tampering from regular image processing and compression, enabling precise localization of tampered regions, providing robust authentication against geometric transformations and transfer of image content across different devices, and designing effective fragile/semi-fragile watermarking balancing imperceptibility with recoverability. Future directions may involve exploring deep learning for feature learning and tampering detection, integration of richer semantic understanding about visual contents, secret sharing-based authentication, and establishing theoretical security analysis frameworks for image forensic schemes.

Overall, the literature demonstrates impressive progress in image forensics, with a diverse array of techniques spanning from watermarking and intrinsic fingerprinting to statistical modeling and machine learning. However, the continuous advancement of image editing capabilities and sophistication of attacks necessitates ongoing research to develop increasingly effective, generalized, and robust tamper detection and localization algorithms. This survey aimed to comprehensively cover the major technical approaches and their evolution thus far to serve as a foundation for further innovations in this critical field.

III. Method

The paper [1] describes the methods to detect tampering in images. It is a new anti-tampering technique that can be used to retrieve the original content rather than just indicate which pixels or blocks have been tampered with. The image is divided into 8×8 blocks and each



A Peer Reviewed Journal

block is DCT transformed, quantized and carefully encoded into the LSBs of other distant 8×8 blocks. It recovers portions of images that have been cropped or replaced or severely modified. It is designed with the intent to maximize the quality of the recovered image. The method in the paper [2] verifies the authenticity of image using the image quality features like Markov and moment-based features. This method extracts Image Quality Metrics (IQMs) by dividing test image into 4 regions, extract features from every region and extract moment-based features by applying wavelet transform to this image and obtain all the sub-bands including the test image itself. The DFT to the histogram of each sub-band is used to obtain its characteristic function. The moments are calculated and 2-D histograms are obtained for the test image. The remaining images are predicted using SVM model. The accuracy of this method is about 79% to 83% for different dataset. The method [3] detects Image Tampering Using Feature Fusion. In this paper, the feature statistics are used for training a one-class classifier to get the feature pattern and sliding segmentation is done to testing images. The images with low percentage of matched blocks are classified as tampered ones. This method achieves a high accuracy in detecting the tampered images. This method does not work for the image is tampered by directly splicing the parts of two images taken by the same camera without any post-processing. The paper [4] detects tampering based on artifacts created by Color Filter Array (CFA) processing in most digital cameras. The techniques are based on computing a single feature and a simple threshold-based classifier. In this paper, two different feature-based tamper detection methods are introduced. The proposed features are used with empirically determined linear thresholds to determine whether given images are tampered or not. This method is successful for tamper detection problem with very low error rates and is not for images acquired with X3 Foveon sensors. The method given in paper [5] presents an innovative algorithm for image tampering detection based on forgery suspect generated by the claimant. The image at the sender side is shielded with security parameters generated from cumulative visual word from unique color features of the image. The recipient checks for the match with secret parameters shared commonly. A mismatch helps in generation of suspicion parameters which serves as a testament in generation of bag of features. The Euclidean distance is used as a metric in localization of tampered regions. The scheme shows 98% true positives detection and the false positive detection rate is nearly negligible. The method [6] describes a new technique that inserts robust watermarks into small disjoint blocks. The pattern is generated by modulating the middle frequencies of the blocks' DCT with a spread spectrum noise-like signal. The watermark is embedded in a robust manner and cannot be removed without introducing visible distortions into the image. This method embeds checksums in LSB M sequences, spread spectrum signal. It enables us to distinguish visible non-malicious changes due to common image processing operations from malicious modifications, such as replacing or adding features. The method [7] focuses on JPEG images and detects tampered images by analyzing the double quantization effect hidden among the discrete cosine transform (DCT) coefficients. This method detects at the scale of 8×8 DCT block and insensitivity to different kinds of forgery methods. The advantages of this method are automatic tampered region determination, resistant to different kinds of forgery techniques in the tampered region, ability to work without full decompression and fast detection speed. The paper [8] detects image slicing using enhanced face extraction techniques and universal classifiers. It detects forged



A Peer Reviewed Journal

images of people using the illuminant color. The illuminant color is estimated using a statistical gray edge method and a physics-based method which exploits chromaticity color space. An efficient face extraction method called successive means quantization transform algorithm has been proposed. Human faces are extracted from the illuminant maps and extract feature using both edge based and gradient based algorithms. And then combine these complementary cues using machine learning late fusion SVM classifier that helps in classification of forged image. The paper [9] presents tamper detection and image recovery method for digital images. It employs the vector quantization scheme to generate the authentication data. The indices for image blocks are generated and multi-copies of the indices are embedded into the selected blocks by the pseudo random number generator. The forward detection strategy and the backward detection mechanism is used to find out the image modifications. The watermarks were generated from the host image by the VQ encoding and multi-copies of watermarks were embedded into the last ebb bits of image pixels of the selected blocks. This method effectively improves the detection accuracy. The paper [10] hides logo information into an image by tuning block pixels based on a bitmap parity checking approach. A secure key and a random number generator are used to hide the logo information in a secret, undetectable, and unambiguous way. The characteristics of the mean gray value and the bitmap in a block are exploited for performing the embedding work efficiently and for hiding a logo into an image imperceptibly. The logo can be extracted without referencing the original image. The proposed method is a fragile watermarking technique; each logo bit can be multiply embedded into the watermarked image. The proposed method is useful for authentication of original digital products. The extracted logo not only can be used to identify tampered locations in digital images but also can resist JPEG compression to a certain degree. This method is feasible and effective. The paper [11] presents image residue features for detecting tampering or forgery in video sequences. It uses feature selection techniques in conjunction with fuzzy fusion approach. It examines different feature selection techniques, the independent component analysis (ICA), and the canonical correlation analysis (CCA) for extracting tamper signatures from quantization and noise residue features. The evaluation of proposed fuzzy fusion technique along with different feature selection techniques for copy-move tampering emulated on low bandwidth Internet video sequences, show a significant improvement in tamper detection accuracy with fuzzy fusion. The method [12] uses Replicated Image detector (RIME). RIME checks if near-replicas of the image exist on the Internet and returns a list of suspect URLs. The core technologies that the RIME project develops are effective image characterization for copy detection, and efficient image indexing for finding images with similar characteristics. The VACH feature appears to be tamper resistant and is good for image copy detection. The experimental results indicate that a color histogram returns the correct closest match less than 20% of the time whereas the conjoined VACH has better than 80% accuracy. The paper [13] introduces a fully affine invariant image comparison method, Affine-SIFT (ASIFT). SIFT is fully invariant with respect to only four parameters namely zoom, rotation and translation, the new method treats the two left over parameters : the angles defining the camera axis orientation. This method identifies features that have undergone very large affine distortions measured by a new parameter, the transition tilt. This paper [14] presents a novel semi-fragile watermarking scheme for image authentication and



A Peer Reviewed Journal

tamper detection. It extracts content-based image features from the approximation sub band in the wavelet domain to generate two complementary watermarks. It generates an edge-based watermark sequence to detect any changes after manipulations and encodes the invariant relationship between quantized wavelet coefficients after incidental distortions. The watermarks are embedded into the high frequency wavelet domain to ensure the watermark invisibility. The method successfully distinguishes malicious attacks from non-malicious tampering of image content. It also accurately localizes maliciously tampered regions. The paper [15] describes an efficient and automatic techniques to identify and verify the content of digital multimedia. It is a perceptual image authentication technique based on clustering and matching of feature points of images. Feature points are first extracted from images with the k-largest local total variations, and clustered using Fuzzy C-mean clustering algorithm. Then feature points in the query image and the anchor image are matched into pairs in zigzag ordering along the diagonals of the image cluster by cluster. The authenticity of images is determined by the majority vote of whether three types of distance between matched feature point pairs are larger than their respective thresholds. The three types of distance include 1) histogram weighted distance, which is proposed in this paper, 2) normalized Euclidean distance, and 3) Hausdorff distance. The geometric transform between the query image and the anchor image is estimated and the query image is registered. The possible tampered image blocks are detected and the percentage of the tampered area is roughly estimated. The results show the effective and robust image authentication system. The paper presents [16] is an efficient image tamper detection method using 3 LSB watermarking technique to authenticate the digital image and detect the tamper locations accurately. A 12-bit watermark key is created from each block of host image; embed to last three significant bits of each block. Different types of tampering attacks have been experimented in order to evaluate the proposed method. It gives high tamper detection rate. The method [17] addresses the problems in the authentication of the image received in a communication. To localize the tampering aligned image should be first registered at the sender by making use of the information provided by a specific component of the forensic hash associated to the image. The image hash component is based on the Bag of Features paradigm. The signature is attached to the image before transmission and then analysed at destination to recover the geometric transformations. The image hash encodes the spatial distribution of the image features to deal with highly textured and contrasted tampering patterns. A block-wise tampering detection uses histograms of oriented gradients (HOG) representation.

III. Conclusions

This comprehensive survey has reviewed the state-of-the-art techniques for detecting and precisely localizing image tampering and forgeries. As digital images play an increasingly vital role across many domains - from news media and journalism to legal evidence, intelligence gathering, medical imaging and more - ensuring their authenticity and integrity has become a critical research challenge. The



A Peer Reviewed Journal

widespread availability of sophisticated photo editing tools has enabled the creation of highly realistic tampered and manipulated images that can go undetected by the naked eye. Developing effective image forensic methods to expose such tampering is essential for maintaining trust and credibility in digital image sources.

The techniques covered in this survey span a diverse range of approaches - from fragile watermarking schemes that embed imperceptible authentication data within the image file itself, to blind methods extracting intrinsic fingerprints and statistical models of image properties to detect inconsistencies caused by tampering. Early strategies relied on simple checksums or signatures inserted into the image pixels or transform coefficients. While capable of detecting tampering, they could not localize the modified regions. This limitation was addressed by block wise fragile watermarking methods that authenticated individual image blocks to isolate tampered areas. Robust watermarks resilient to allowable manipulations like compression were also explored.

As an alternative to watermarking, many techniques took a features-based approach - extracting characteristic image descriptors like quality metrics, Markov statistics, wavelet coefficients etc. and training machine learning classifiers to identify anomalies indicating tampering. With JPEG's widespread adoption, analysis of double compression artifacts indicative of splicing became a popular JPEG image forensic tool. Geometric and photometric inconsistencies across image regions also provided tampering cues leveraged by various algorithms.

Perceptual hashing schemes computed compact image digests or signatures to enable integrity verification - any modification would alter the hash value. These used dimensionality reduction, randomized feature embedding and bag-of-words models to obtain robustness against acceptable transformations like compression while remaining sensitive to malicious tampering. Features from high-contrast regions, clustered feature point patterns and gradient histograms were also utilized for tamper localization.

More recent work explored fusing multiple cues like coding artifacts, camera fingerprints, physics-based illumination inconsistencies etc. using machine learning ensembles to boost detection accuracy. There has also been growing interest in applying deep learning for automatic feature learning and end-to-end tampering detection. Despite the impressive body of work, several key challenges persist that provide avenues for future research:

Distinguishing malicious tampering from regular image processing/compression: Many existing methods are limited in their ability to differentiate between intentional malicious modification and innocuous alterations introduced by common processing operations like resizing, compression, filtering etc. More sophisticated semantic analysis may be needed to disambiguate based on tampering intent.

Precise localization of tampered regions: While some techniques aim to localize modified areas, there is still room for improvement in enhancing localization granularity and accurately delineating tampered regions, especially for geometrically transformed spliced regions.

Robustness against geometric transformations: Most current forensics schemes are fragile to geometric transformations like rotation, scaling, skewing etc. Developing geometric transform-invariant authentication approaches is an important challenge.



A Peer Reviewed Journal

Integration of semantic understanding: Most current methods use low-level signal characteristics like compression artifacts, filter residues etc. Integrating higher-level semantic understanding about the visual contents of the image could potentially enhance forgery detection capability.

Scalable analysis of large image datasets: With the explosion of imagery data from social media, surveillance and other sources, there is a growing need for computationally efficient image forensic techniques that can conduct rapid screening and analysis at scale across massive image collections.

Moving forward, continued innovation leveraging latest technologies like deep learning, advancement of theoretical foundations providing security guarantees, and development of hybrid multi-modal approaches that can synergistically combine different classes of tampering detection cues, hold significant promise to push the boundaries of robust image authentication and forgery detection. As image manipulation capabilities continue evolving, so must image forensic countermeasures to ensure the credibility of digital imagery sources.

This survey aimed to comprehensively capture the evolution of image tampering localization techniques thus far, highlighting their core principles, merits and limitations. By dissecting state-of-the-art methods and elucidating open challenges, it provides a foundation for further research advancing this critical field. Only through continuous advancement of image forensic capabilities can we effectively combat the growing threat of deceptive multimedia forgeries and uphold trust in digital image sources across critical applications.

IV. References

- [1] Jiri Fridrich and Miroslav Goljan (1999) Images with Self-Correcting Capabilities, International conference on image processing, 588.
- [2] Shrishail Math1 and R.C.Tripathi (2011) Image Quality Feature Based Detection Algorithm for Forgery in Images, International journal of computer graphics and Animation, 1(1), 2231 - 3591.
- [3] Pin Zhang (2009) Detecting Image Tampering Using Feature Fusion, International Conference on Availability, Reliability and Security, 335-340.
- [4] Ahmet Emir Dirik (2009) Image Tamper Detection Based On Demosaicing Artifacts, IEEE International Conference on Image Processing, 1497-1500.
- [5] Deepali N. Pande, A.R. Bhagat Patil and Antara S. Bhattacharya (2014) Generic Algorithm for Image Tampering Detection Based on Claimant Suspect Decision Rule, International Journal of Computer Science Engineering & Technology, 4(4), 121.
- [6] Jiri Fridrich (1998) Methods For Detecting Changes In Digital Images, Proc. of The 6th IEEE International Workshop on Intelligent Signal Processing and Communication Systems (ISPACS'98), 4x6.
- [7] Zhouchen Lina, Junfeng Heb, Xiaou Tanga and Chi-Keung Tang (2009) Fast, automatic and fine-grained tampered JPEG image



A Peer Reviewed Journal

detection via DCT coefficient analysis, Elsevier Journal of pattern recognition, 42(11),2492–2501.

[8] V.P.Kavitha and M.Priyatha, A Novel Digital Image Forgery Detection Method Using SVM Classifier,International Journal of Advanced Research in Electrical, Electronics and Instrumentation Energy, ISSNONLINE(2278-8875) PRINT (2320-3765).

[9] Jun-Chou Chuang, Yu-Chen Hu, Chun-Chi Lo and Wu-Lin Chen (2013) Grayscale Image Tamper Detectionand Recovery Based on Vector Quantization, International Journal of Security and Its Applications 7(6), 209-228.

[10] Chen-Kuei Yang and Chang-Sheng Huang (2004) A Novel Watermarking Technique for TamperingDetection in Digital Images, Electronic letters on computer vision and image analysis, Barcelona, Spain, 3(1), 1-12.

[11] Girija Chetty and Monica Singh (2010) Nonintrusive Image Tamper Detection Based on Fuzzy Fusion,IJCSNS International Journal of Computer Science and Network Security, 10(9), 86-90.

[12] Beis, J. and D.G. Lowe (1997) Shape indexing using approximate nearest-neighbor search in highdimensional spaces, International Conference on Computer Vision and Pattern Recognition, Puerto Rico, 1000-1006.

[13] Guoshen Yu (2009) A Fully Affine Invariant Image Comparison Method Cmap, IEEE InternationalConference on Acoustics, Speech and Signal Processing, 1597-1600.

[14] Xiaojun Qi, Xing Xin, and Ran Chang (2009) Image Authentication And Tamper Detection Using TwoComplementary Watermarks, 16th IEEE International Conference on Image Processing (ICIP), 4257-4260.

[15] Lei Yang (2012) Content Based Image Authentication by Feature Point Clustering and Matching, ACMJournal of Security and Communication Networks, 5(6), 636-647.

[16] Sajjad Dadkhah, Azizah Abd Manaf and Somayeh Sadeghi (2012) Efficient Digital Image Authenticationand Tamper Localization Technique Using 3Lsb Watermarking, IJCSI International Journal of ComputerScience Issues, 9(1(2)), 1-8.

[17] Jyoti Rao (2013) Robust Image Alignment and Tampering Detection Using SVM Clustering, ASM'sInternational E-Journal of ongoing research in management and IT, 1-6.