Transactions on Computer Science and Intelligent Systems Research
ISSN: 2960-1800, eISSN: 2960-2238 | Volume 5
AIDML 2024

Analysis of Digital Media Image Tampering and Forgery Control

Bin Tang [1, *], Mengyao Sun [2], Haowen Zheng [3], Haiyun Han [2], Yu Jin [1]

[1] School of Intelligence Policing, China People's Police University, Beijing, China

[2] School of fire protection Engineering, China People's Police University, Beijing, China

[3] Xiamen East border inspection, Xiamen, China

* Corresponding Author Email: toynbinbin@163.com

Abstract. In recent years, the continuous innovation of computer vision processing technology has greatly improved the quality of people's production and life, while the accompanying tampering risk has also gradually increased. There are frequent incidents of people with ulterior motives using computer image processing technology to tamper with digital media and falsifying and cheating others. This paper mainly analyzes the technology and method of digital media image tampering and forgery, the social harm caused by the resulting tampering risk, and puts forward the corresponding management countermeasures.

Keywords: Digital media; Falsification; Social harm; Prevention and treatment.

1. Introduction

In the information age with the rapid development of high-speed networks and intelligent terminals, the high real-time performance of media network platforms makes the transmission of digital media information unprecedentedly rapid. People often believe that "hearing is false, seeing is believing", and the popular "picture (video) with truth" in network media has always been a convincing way of communication. But is "a picture (video)" the "truth"? The reality is that fake news is all over the Internet, and today's technology allows even inexperienced counterfeiters to alter digital media without leaving a visible trace with the help of editing tools, such as Photoshop, CorelDraw, Deepfake, the "ZAO" App, and more. "There is' picture (video) 'and there is' truth'", which is very misleading to the public and even causes hidden dangers of social unrest [1].

2. Overview of Digital Media Tampering and Counterfeiting

Digital media tampering forgery forms are generally divided into traditional tampering forgery and deep tampering forgery two kinds.

2.1. Traditional Tampering Forgery

The traditional tampering and forgery technology is relatively mature, and several common techniques are:

2.1.1. Copy and Paste.

Remove a local area in the image, fill it with the pixel value of the surrounding area, or copy and paste a local area in the same image from one location to another;

2.1.2. Compositing.

The local areas of different images are spliced together, or multiple images are copied by copying a region of their own images to another image to add the image that does not exist;

2.1.3. Embellishment and Beautification.

Retouching is mainly to modify the digital image to make it more beautiful and better effect, its main operation is to modify and render some images in the image which cannot reach the psychological expectation;

2.1.4. Image Carrying.

It is through the digital media carrier, so that the transmitter or witnesses cannot judge whether there is any hidden information except the carrier itself, to achieve the need to transfer or need to hide the information in the carrier.

2.2. Deep Tampering Forgery

The computer-generated image is different from the previous tampering methods, of which the most commonly used and most mature technique is deep falsification. Deep forgery is a technology that uses artificial intelligence (AI), machine learning, and neural networks to falsify content such as images and videos. For example, the current most popular application Sora can generate up to 60 seconds of high-quality video based on simple text description. The new technology not only benefits the field of video creation, but also makes video forgery unprecedentedly simple and efficient.

3. Research Status at Home and Abroad

Digital media image tampering detection technology can be divided into two categories. According to whether the digital image is pre-embedded with anti-counterfeiting information, tampering detection methods can be divided into active detection and passive detection (blind detection). Active detection technology mainly adopts digital signature technology and digital watermarking technology. The common point of these two methods is that the original party of the image needs to cooperate with the extraction of abstract information or the recognition of watermark during the actual detection. This condition has great limitations in actual operation, and is generally applied in important occasions such as judicial authentication. In practical application, the vast majority of digital images are not pre-processed for anti-counterfeiting, which greatly limits the scope of use of this detection technology. Blind detection of digital images refers to judging whether they have been tampered with or forged according to the inherent attributes of the digital image to be identified (such as illumination, artifacts, background noise, statistical characteristics, etc.). It does not rely on any pre-signature abstract information extraction, and has no special requirements for image shooting equipment. It is a convenient and effective means of tampering detection, so it has become a major research hotspot in the field of digital media tampering detection.

4. The Management Strategy of Digital Media Tampering and Forgery

4.1. Classification of Identification Methods

Digital media image can be divided into digital image and digital video, because digital video can be decomposed into a certain number of digital images according to the number of frames, if the identification of each separated frame image, you can complete the identification of the entire video, so the identification method of digital video can be reduced to a group of digital image identification method.

As mentioned earlier, there are two types of feasibility detection measures to deal with digital media tampering and forgery, one is active detection and the other is passive detection. Active detection belongs to pre-emptive defense, which can be understood as "anti-counterfeiting", that is, to establish a trusted mechanism, so that falsified images cannot be spread under the mechanism. The idea of active defense is to add small information summaries or watermarks before the original author releases the images, and the added protection information can be verified for authenticity. Active

defense technology is divided into two categories: digital signature technology and digital watermarking technology.

The digital signature technology combines the digital signature formed by the original image and the encrypted abstract in the publishing end, extracts the abstract content of the image as the identification signature at the user end, in order to verify the authenticity of the image, and recognizes the original state of the received image with the help of the content summary signature information. Digital watermarking technology is the copyright information embedded in the digital work itself, the use of digital works in the universal redundant data space, in order to achieve the purpose of protecting the copyright or authenticity of digital products.

Passive detection belongs to post-detection, that is, to detect the digital media that has been produced and disseminated, and judge whether the suspected image is tampered with and forged. In the post-event detection method, tampering detection can be carried out from two directions. One is from the technical point of view to detect whether the image has the characteristics of tampering and forgery; The second is to trace the source from the perspective of open source information retrieval, to collect and analyze the material on the network or other open sources, to restore the "truth" of the image, which is the "open source verification" method. "Open source verification" provides a new idea for the detection of tampering and forgery.

4.1.1. Classification of Passive Detection Technology for Image Tampering.

a. Passive detection of image reproduction

The basic principle is that if the same copied content can be detected in the digital image, it can be identified as tampering, and similar image content areas can be found using publicly available computer vision algorithms.

b. Passive detection of image stitching

The basic principle is to process the doubtful image through the filter to highlight the edge features of the image, so as to find out the edge difference between the real area and the tampered area.

c. Passive detection of background noise

Each kind of shooting equipment has its own unique characteristics: the background noise generated by the photosensitive material of the imaging sensor, these characteristics are called "device fingerprint", this noise will be present in the captured image, but because its amplitude is very small, the naked eye can not detect. The background noise of different images is different, and it can be judged whether the image has been tampered with according to its distribution law.

d. Passive detection of illumination consistency

Each shooting scene has its own unique lighting environment, the direction of each scene is often different, when two different scenes appear in the same picture, the scene carries the light trace for detection.

e. Computer-generated passive detection

With the successful application of artificial intelligence technology in various computer vision and image processing tasks, the deep learning image detection method combines traditional filtering methods with deep learning models, and can extract key statistical features of digital images for automatic detection and analysis, and has achieved good results.

4.1.2. Image Hidden Information Detection.

The purpose of image secret is to embed the secret information into the carrier image in a way that is difficult to detect through the key and a specific algorithm, and then the information receiver extracts the secret information through the key and extraction algorithm. One of the application directions of image secret technology is the digital watermarking technology mentioned earlier, which is used for content authentication and copyright protection.

The common secret image detection methods can be divided into blind attack detection and non-blind attack detection according to whether the secret algorithm or key is mastered in advance. Blind attack detection does not need to know the specific implementation details of the secret algorithm, but only extracts the secret information by analyzing the carrier's statistical characteristics or other clues. The non-blind attack detection needs to obtain some additional information, such as the algorithm details of steganography or the embedded key.

4.2. Governance Strategy

Due to the increasing challenge of digital media image tampering and forgery to social credibility, in recent years, many countries and regions, including China, the United States and the European Union, have realized the serious security threat, and relevant departments and academic groups in various countries have formulated coping strategies. This includes but is not limited to amending legislation for control, strengthening supervision of information release platforms, researching and developing new early warning and detection technologies, and strengthening publicity and education actions to deal with technology abuse, tampering and forgery and their security risks.

4.2.1. Formulate Regulatory Regulations.

China has issued a series of regulations on tampering and forgery. The CAC issued the Regulations on the Administration of Online Audio and Video Information Services in 2019, which stipulates that media publishers should conduct security assessments in accordance with relevant national regulations, identify inauthentic audio and video information in a prominent way, and refrain from using AI technology to publish fake news or deploy AI forgery identification technology to improve rumor-refuting mechanisms. In 2021, the CAC, the Ministry of Public Security, the General Administration of Market Regulation, the State Administration of Radio, Film and Television and other relevant departments promulgated the Regulations on the Administration of Algorithm Recommendation for Internet Information Services, the Regulations on the Administration of In-depth Synthesis of Internet Information Services, and the Interim Measures for the Administration of Generative Artificial Intelligence Services, among other laws and regulations. It requires providers of algorithm recommendation services and practitioners providing Internet news and information services to carry out prominent signs of algorithmic generation of synthetic information to highlight the audience, and must not generate synthetic false news information. The newly released Interim Measures for the Management of Generative AI Services further standardizes the development and use of generative AI, and strictly supervises computer-generated content.

4.2.2. Strengthen Platform Review.

In recent years, China's digital media industry has developed rapidly, and many social media platforms such as wechat, Weibo, QQ, Douyin, Beizhong station, Doudou, Kuaishou and XiaoHongshu have emerged as The Times require. These "we media" platforms, which spread through mobile terminals, have fast dissemination speed, large audience and wide influence. Many "we-media" accounts will use the platform to produce creative works, maliciously forge false images and spread rumors in order to gain traffic and attention. In order to prevent the "wild growth" of the "we media" industry, the relevant authorities require the operators of social media platforms to formulate industry self-discipline norms and develop detection tools, strengthen the behavior management of "we media" content creation, and take necessary crisis response mechanisms when malicious tampering of media information is found to be spread wantonly [2].

4.2.3. Technology to Serve the Public.

With the continuous development of tampering and forgery technology, it is difficult for the general public to distinguish the authenticity of digital images. Relevant departments may authorize competent companies to develop technical service platforms (such as mobile phone apps, websites, etc.) capable of identifying falsified images for the public, where users can self-check the authenticity of videos, images and other digital media images and generate appraisal reports. Companies can

charge a certain amount of fees as the cost of platform maintenance and for the sustainable development of technological innovation (the relevant departments can bear a certain proportion), so as to form a win-win benign pattern [3].

4.2.4. Strengthen Law Popularization Education.

Relevant departments can provide AI safety education to the public through a variety of channels, such as movies, TV propaganda videos, public interviews, short videos from the media, wireless broadcasting, Internet pop-up advertisements, mass text messages on mobile phones, and propaganda in places with display screens such as banks and shopping plazas. It can also print and distribute some free leaflets and pamphlets to industrial and mining enterprises, enterprises and institutions, community streets, schools and villages to disseminate publicity, remind the public not to easily believe "seeing is believing", improve identification ability, enhance vigilance awareness, and effectively safeguard their own life and property safety. Publicity should focus on introducing the basic methods and effective tools for the identification of tampering and forgery, as well as the means and measures usually taken to confirm the true and false of tampering and forgery.

5. Summary

Deep forgery can bring the public a new surreal entertainment experience, with a strong sense of reality and confusion, in the dissemination process can also use the high radiation of digital media network platform, further exacerbate the rapid spread of false information. Relevant departments at home and abroad have paid great attention to digital media tampering and forgery technology, and promulgated governance measures from different ways. In the future, the supervision of digital media tampering and forgery should focus on prevention and control, and use image processing technology to develop more advanced detection tools to prevent the wantonly spread of virtual information.

Acknowledgements

The authors gratefully acknowledge the financial support from the fund: Science and Technology Project of Ministry of Public Security in 2022, Project number :2022JSYJC24.

References

[1] WU Shuang. The problem of "Deep forgery" under AI craze cannot be ignored [N]. People's Posts and Telecommunications, 2023 - 06 - 22 (6).

[2] Zhang Yun, Sun Rongxin, Zhang Xu. The status quo, problems and approaches of Network short video content governance in the Digital Intelligence Era [J]. Publication and Distribution Research, 2023 (06): 13 - 19.

[3] Long Kun, Ma Yue, Zhu Qichao. The challenge of deep forgery to national security and its countermeasures [J]. Information Security and Communication Security, 2019 (10): 21 - 34.

[4] Lei Jinjin. Research on image hashing method for content forensics [D]. Xian: Xian University of Technology, 2022.

[5] Chen Qiansheng. Digital image forensics technology and its development [J], Silicon Valley, 2011 (10): 26+17.