IJCRT.ORG ISSN : 2320-2882

INTERNATIONAL JOURNAL OF CREATIVE RESEARCH THOUGHTS (IJCRT)

An International Open Access, Peer-reviewed, Refereed Journal

IMAGESHIELD: ML-POWERED IMAGE TAMPER DETECTION AND RECOVERY

[1] I Rajesh, [2] K Dharun, [3] C Balakrishnan, [4] S Karthikeyan, [5] R Venkatapathy

[1]Associate Professor, [2,3,4,5]UG Student
[1]Computer Science and Engineering,
[1]Knowledge Institute of Technology, Salem, Tamil Nadu, India

***Abstract* :** The widespread use of digital images in online social networking sites has sparked concerns about the susceptibility to image tampering, which can compromise security and privacy by distorting content. ImageShield, a deep learning solution aimed at enhancing resistance to image tampering and enabling lossless recovery. ImageShield integrates multiple modules through multitask learning, ensuring image integrity and authenticity. It applies a vaccine to uploaded images,maintaining alignment with the original. Forward pass transformation generates an immunized version, crucial for consistency. In cases of tampering, a localizer module identifies altered areas using a tamper mask. The backward pass, utilizing Run-Length Encoding, deciphers hidden perturbations, enabling lossless recovery of the original image. Real-world experiments confirm ImageShield's effectiveness in tamper localization and content restoration, offering a promising solution for social networking image tampering, thus reshaping image security and privacy in the digital era.

Keywords: Image Tampering Detection, Deep Learning,Lossless Recovery, Multitask Learning, Content Restoration

I. INTRODUCTION

In today's digital age, the pervasive use of digital imagery across online social networking platforms has prompted increasing apprehension about the vulnerability to image tampering. This susceptibility poses substantial risks to both security and privacy by distorting content, potentially leading to the dissemination of false information, manipulation of identities, or unauthorized access to personal data. As individuals rely more on digital platforms to communicate, share information, and interact with others, the integrity of digital images becomes paramount in preserving trust and safeguarding online interactions. The demand for robust image security measures becomes more evident. Conventional tampering detection methods often prove inadequate, especially when dealing with compressed or low-resolution images. This paper introduces ImageShield, an advanced deep learning solution, aiming to strengthen resistance against image tampering and facilitate the seamless restoration of original content by integrating various modules through multitask learning to ensure the credibility and genuineness of images.At its core, ImageShield introduces a approach to image protection. By applying a protective "vaccine" to uploaded images and employing forward pass transformation to produce immunized versions, ImageShield maintains alignment with the original content, thereby enhancing consistency and preemptively addressing potential tampering instances. Furthermore, in instances of detected tampering, ImageShield utilizes a localizer module to pinpoint altered areas using a tamper mask, enabling precise identification of tampered regions and subsequent mitigation measures. A pivotal aspect of ImageShield is its backward pass mechanism, which leverages Run-Length Encoding to uncover concealed perturbations and facilitate lossless recovery of the original image. This innovative strategy ensures that even in the presence of tampering, the authenticity and integrity of the original content can be accurately restored. The efficacy of ImageShield has been confirmed through extensive real-world experiments, validating its effectiveness in both tamper localization and content restoration. These results underscore the potential of ImageShield as a promising solution to combat the challenges associated with image tampering on social networking platforms. This paper will delve into the role of ML in comprehensive tamper detection and recovery, demonstrating its potential in reshaping the landscape of image security and privacy in the digital era. ImageShield offers a crucial defense against the escalating threats of image tampering, thereby fostering a safer and more secure online environment.

II. EXISTING SYSTEM

Digital forensics techniques involve the thorough examination of digital evidence to detect and analyze patterns of forgery. This process often includes scrutinizing metadata, compression artifacts, and other forensic traces present within the image. Signature-based methods rely on known patterns or signatures of forgery to identify manipulated images. These signatures may encompass noise patterns, repeated patterns, or specific features associated with common tampering techniques.Watermarking entails embedding invisible or visible marks within an image to establish its origin or ownership. These watermarks serve to track and authenticate the authenticity of images.Image hashing generates a unique hash or fingerprint for an image, which drastically changes with any alterations to the image, no matter how minor. This allows for the detection of forgery through hash value comparison.Image processing techniques, such as error level analysis (ELA) and noise analysis, are employed to identify inconsistencies in pixel values or compression artifacts that may indicate tampering. Steganalysis techniques are focused on uncovering hidden information within images, particularly alterations made through steganography. Steganography involves concealing additional information within the image, which steganalysis aims to detect and analyze.

III. PROPOSED SOLUTION

The proposed system, called ImageShield, is a middleware for Online Social Networks (OSN) utilizing Invertible Neural Network (INN), aiming to bolster image security and integrity on social media platforms. It comprises several modules and functionalities, including pre-processing, mid-processing, and post-processing steps. Landmark detection algorithms create binary masks to distinguish object contours, while imperceptible perturbations are introduced to ensure visual consistency and embed crucial information. ImageShield features a Vaccine Validator module specific to OSN, distinguishing between secured and potentially tampered media to validate image integrity. It includes an adversary to simulate threats like deepfake attempts, ensuring robustness within the social network context. Transforming the original image and metadata into an immunized version using INN constitutes the forward pass, while the backward pass facilitates image recovery. An adversarial simulation strategy during training exposes the network to OSN-specific threats, including deepfakes and manipulations. Designed for seamless integration with existing OSN architectures, the middleware ensures widespread adoption and compatibility among OSN users.

3.1 VACCINATION

In Vaccination, the Invertible Neural Network (INN) ensures media authenticity preservation through pre-processing, mid-processing, and post-processing stages. Landmark detection creates a binary mask initially, followed by the INN generating a raw output from the original portrait and mask. Post-processing replaces the object region in the raw output with the original portrait, guided by the mask. Imperceptible perturbations are strategically introduced only to the non-object region. The Vaccine Validator module distinguishes between vaccinated and unvaccinated media, monitoring social network activities for image sharing and initiating notifications when a shared vaccinated image is detected without any attack by the ImageShield.

3.2 PREDICT TAMPER MASK

The forward pass incorporates an Invertible Neural Network to transform an original image and its edge map into an immunized version. In the face of an attacked image, a localizer comes into play, determining tampered areas by predicting the tamper mask and attack. By training a system to identify the mask region in an image, then analyze its shape, texture, and consistency for anomalies that might indicate tampering, this vigilance is instrumental in identifying and localizing digital tampering, fortifying the system against a spectrum of potential attacks.

3.3 AUTO RECOVERY

In the backward pass, the hidden perturbation, transformed by the Invertible Neural Network, is converted into information. This transformative process serves the dual purpose of recovering the original image and its edge map while fostering image self-recovery. By encouraging the recovered image to closely resemble the original, this module ensures a seamless restoration process.

3.4 QUALITY ASSURANCE

Lossless image recovery through Run-Length Encoding (RLE) effectively maintains the original data after identifying and eliminating tampering. RLE compression efficiently represents sequences by encoding consecutive identical pixel values. Combining tamper detection, removal, and RLE-based recovery enhances the robustness of the system. The Peak Signal-to-Noise Ratio (PSNR) provides a quantitative assessment of image quality relative to the original. This metric is computed using the formula PSNR = 10 * log10((MAX^2) / MSE), where MAX represents the maximum pixel value (255 for an 8-bit image), and MSE denotes the Mean Squared Error.

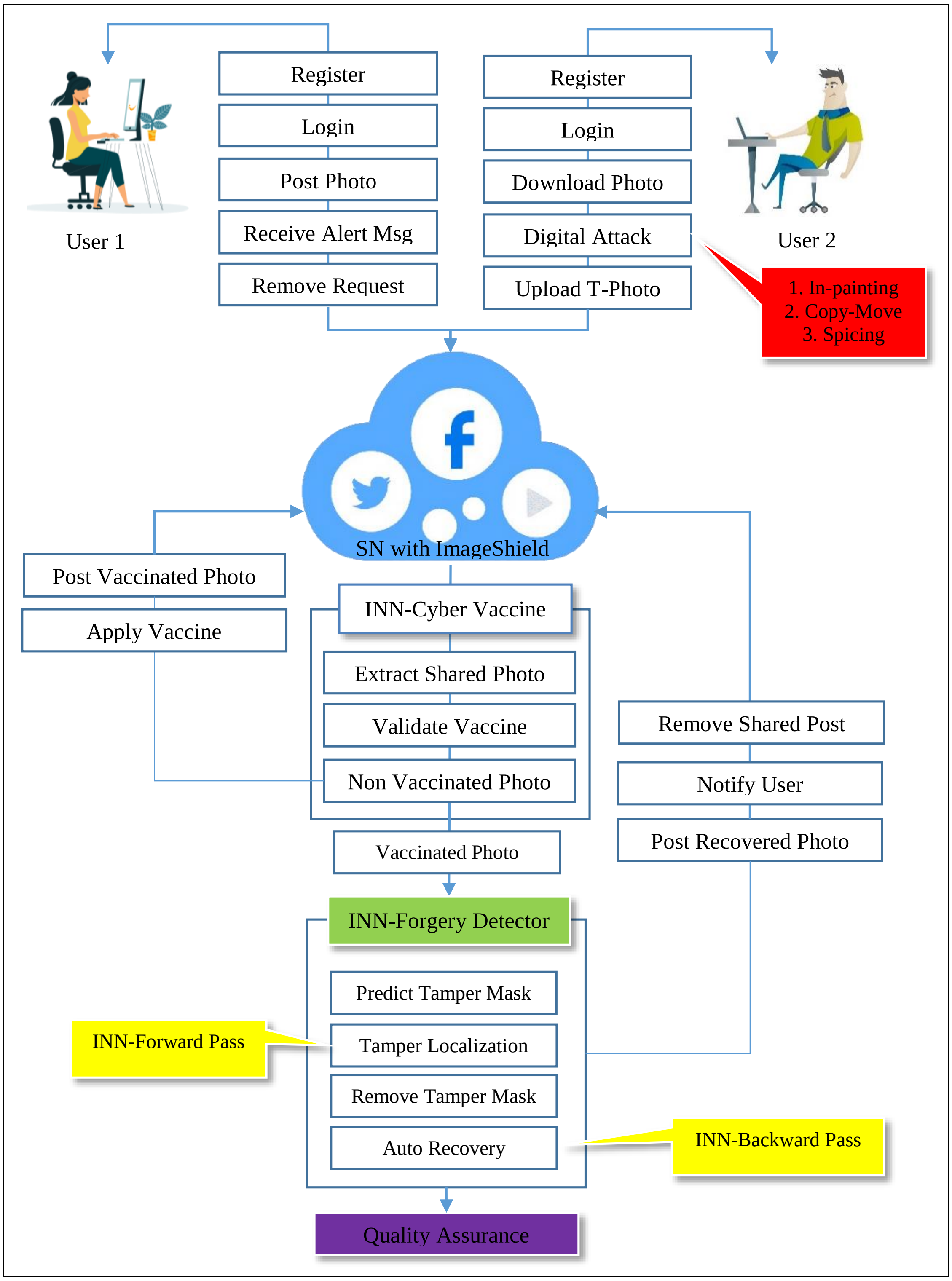


Fig 1: Workflow Of The System

IV. RESULT AND DISCUSSIONS

After modelled and trained, the result Obtained from our system. The model is developed using Invertible Neural Network (INN), a classification algorithm whose accuracy is predicted to be 96%. It enhances tamper resilience, ensuring resistance against unauthorized alterations and facilitating recovery without loss of original image information. By improving the security of digital images against tampering threats, the solution is applicable across domains such as forensics and digital communication, maintaining the authenticity of recovered images.

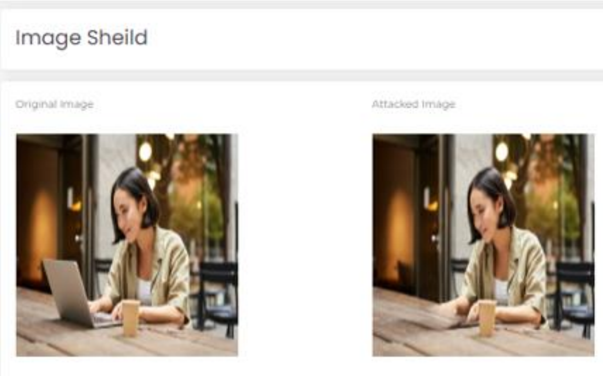
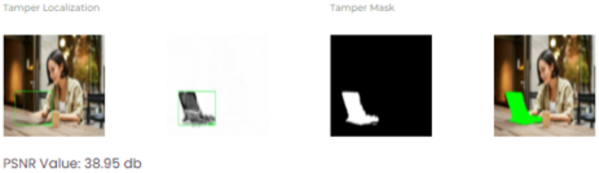


Fig.2 Original Image & Modified Image



PSNR Value: 38.95 db

Fig.3 Detection & Recovery Process

In addition to its tamper resilience, the Invertible Neural Network (INN) also offers scalability, allowing for efficient processing of large volumes of image data without compromising accuracy. Its robustness against adversarial attacks further solidifies its reliability in safeguarding digital assets. Moreover, the adaptability of INN facilitates its integration into various systems and platforms, ensuring seamless implementation across diverse technological environments.

V. Conclusion

The proposed ImageShield Middleware for Online Social Networks offers an innovative solution to combat digital image tampering. Leveraging Invertible Neural Network technology and adversarial simulation, the system ensures the authenticity and integrity of images on social platforms. It preprocesses, vaccinates, and post-processes images, introducing imperceptible perturbations to safeguard against tampering. The Vaccine Validator distinguishes between vaccinated and unvaccinated media, enhancing security. Utilizing the Forward Pass with INN and the subsequent Backward Pass enables tampered area identification and restoration, reinforcing shared media reliability. Adversarial simulation during training enhances resilience against diverse manipulation forms. Seamless integration with existing OSN architectures ensures compatibility and widespread adoption. Additionally, the system notifies users about image status and restores tampered images, fostering a secure social media landscape.

References

[1] Abdul Shaik, Ram Karsh. “Content Authentication And Tampered Localization Using Ring Partition And Cslbp-Based Image Hashing” IEEE(2023),DOI: 10.1109/ACCESS.2023.3330969.

[2] Xin-Yu Zhang, Kai Xie, Mei-Ran Li. “Generative Facial Prior And Semantic Guidance For Iterative Face Inpainting” IEEE(2022),DOI: 10.1109/ACCESS.2022.3185210.

[3] Sunpreet Sharma. “A Novel Multipurpose Watermarking Scheme Capable Of Protecting And Authenticating Images With Tamper Detection And Localisation Abilities” IEEE(2022),DOI: 10.1109/ACCESS.2022.3198963.

[4] Caipng Yang, Shuyuan, Hong Li. “Transu2-Net: A Hybrid Transformer Architecture For Image Splicing Forgery Detection” IEEE(2023),DOI:10.1109/ACCESS.2023.3264014.

[5] Cheol Woo Park, Yong Ho Moon (2021) “Image Tampering Localization Using Demosaicing Patterns And Singular Value-Based Prediction Residue” IEEE,DOI: 10.1109/ACCESS.2021.3091161.

[6] Yuan Rao, Jiangqun Ni And Huimin Zhao “Deep Learning Local Descriptor For Image Splicing Detection And Localization” IEEE,DOI: 10.1109/ACCESS.2020.2970735.

[7] Zhen Liu. “Enhanced Tamper Detection Algorithm Using Yolov5s With Cbam Attention And Eiou Loss” IEEE(2023),DOI: 10.1109/ACCESS.2023.3311474.

[8] Kejun Zhang, Yu Liang “No One Can Escape: A General Approach To Detect Tampered And Generated Images” IEEE(2019),DOI: 10.1109/ACCESS.2019.2939812.