International Journal of Scientific Research and Engineering Development-– Volume X Issue X, Year

Available at www.ijsred.com

RESEARCH ARTICLE OPEN ACCESS

IMAGE FALSIFICATION DETECTION USING ADAPTIVE OVERSEGEMENTATION AND KEY POINT MATCHING

Mrs. N Deepashree *, Jatin Singh**, Mohammed Zayed Pasha ***, Second Harsh****,

*(Dept of Computer Science and Engineering, KSSEM, Bangalore.
Email: deepashreen@kssem.edu.in)
** (Dept of Computer Science and Engineering, KSSEM, Bangalore.
Email: jatinsingh1655@gmail.com)

*** (Dept of Computer Science and Engineering, KSSEM, Bangalore.
Email: mohammedzayed602@gmail.com)

**** (Dept of Computer Science and Engineering, KSSEM, Bangalore.
Email: harzh54325@gmail.com)

----------**********----------

Abstract:

This paper outlines a novel strategy for detecting image falsification by employing adaptive super segment fusion and feature point matching. The proposed method utilizes a copy-and-move approach to identify forged regions within images. Through super-pixel analysis, The images are divided into irregular... blocks, enabling the detection of characteristic points that indicate manipulation. Subsequent application of morphological operations enhances the accuracy of identifying false regions. The effectiveness of the approach is demonstrated through 16 experiments, highlighting its superior performance, especially in challenging scenarios, when compared to established methodologies. This research contributes to advancing image forensics by providing a robust solution for detecting image forgeries with enhanced efficiency and accuracy.

Keywords — **Forgery, Adaptive oversegmentation ,key point feature matching.**

----------**********----------

1. INTRODUCTION

Copy-move falsification represents a prevalent method of image manipulation, wherein a chosen segment of an image is duplicated and inserted elsewhere within the same image. This technique is used to either disguise or replicate specific content within the image, potentially misleading viewers by creating the illusion of identical objects or scenes.

In the realm of cyber investigation and image processing, experts have actively devised methodologies to detect such alterations. They employ algorithms and techniques aimed at unveiling instances of copy-move falsification by scrutinizing inconsistencies in duplicated regions, identifying patterns, or analyzing discrepancies in pixel values, indicative of potential tampering.

Numerous methods have been put forward, including block matching, which involves comparing image blocks to identify similarities, utilizing key segments to pinpoint replicated regions, analyzing the frequency domain through methods like Discrete Wavelet Transform (DWT), and leveraging machine learning and deep learning models trained on extensive datasets to recognize patterns of manipulation.

However, despite these endeavors, detecting and thwarting image tampering remains an ongoing challenge, owing to the continual evolution of editing tools and techniques. Thus, ongoing research aimed at developing robust falsification detection methods is imperative to uphold the credibility and integrity of digital images. One such method, proposed by Fridrich et al. [1], entails dividing the image into overlapping blocks and matching quantized DCT coefficients to identify tampered areas. Additionally, keypoint-based methods, utilizing image keypoints for resilient duplication detection, have been suggested.

.

2 Terminologies

2.1 Machine Learning

Machine learning Machine learning is a small section of ARTIFICIAL INTELLIGENCE that designs algorithms that allow machines to learn from data and improve performance. It uses methods of computer-science, mathematics and statistics to analyze and understand patterns in data. This depends on the idea of statistical learning. The usage of machine-based learning techniques can benefit many tasks, including NATURAL LANGUAGE-PROCESSING, predictive modeling and image recognition. They are useful for predicting and Extracting insights from intricate data is also said as "data analytics" or "data interpretation." because they are learned by big data and adapt to new information. Machine learning has many uses across industries such as manufacturing, healthcare, finance and entertainment.

2.2 OPEN CV

OpenCV, the Open-Source Computer Vision Library, is a central player in computer vision and image processing. Since its creation, OpenCV has become a flexible and reliable framework, offering a wide range of capabilities from simple image editing to advanced machine learning-based tasks. It can use various coding based language like Python and C++, making it suitable for a diverse set of applications. OpenCV isn't just for academic purposes; it's also widely used in industry where real-time processing and accuracy are essential.

A key feature of OpenCV is its unique and modular design, that enables developers to focus on the parts they need, which helps beginners get started and provides more experienced users with additional depth. The extensive documentation and strong

community support also make it convenient to use and encourage innovation. OpenCV's compatibility with other popular frameworks and libraries, like TensorFlow and PyTorch, enables seamless development of complex computer vision applications.

OpenCV has made a substantial impact in various fields, including robotics, augmented reality, and autonomous vehicles, influencing the development and implementation of these technologies. Its main job in object detection and recognition has driven advancements in security systems, surveillance, and even medical imaging, where accuracy is crucial.

OpenCV's popularity goes beyond its technical features. Being open-source, it fosters a collaborative environment, allowing developers and researchers worldwide to contribute and share their knowledge. This collaborative spirit has driven innovation and encouraged exploration within the compluter feild.

Overall, OpenCV is more than just a combination of tools; it's a vibrant community that pushes the boundaries of technology. As computer vision is rapidly growing, OpenCV is gonna remain at the forefront, providing a platform for both experienced professionals and newcomers to experiment, innovate, and ultimately shape the future of how we interact with visual data.

2.3 CNN(Convolutional Neural Network)

Convolutional Neural Networks (CNNs) are a cornerstone of deep learning, especially in the realm of image processing and pattern recognition. These networks of multiple layers that progressively learning to identify increasingly complex features in images through the usage of convolutional filters. This ability to detect edges, shapes, and textures gives CNNs their strength in recognizing and understanding visual patterns. Max pooling, which reduces the spatial dimensions, enhances computational efficiency without the significant information.

A critical component of CNNs is the fully connected layer, which synthesizes the features learned by earlier layers to make predictions or classifications. The use of shared weights in convolutional layers endows CNNs with translation-invariant capabilities, allowing them to perform consistently across varying spatial locations.

CNNs are inspired by the human visual cortex, mimicking the way neurons respond to specific areas of visual stimuli. This biological basis is a key reason for their remarkable success in tasks, and facial recognition. The versatility of CNNs has been instrumental in the advancement of computer vision and artificial intelligence, leading to significant breakthroughs in analyzing and interpreting visual Data.

CNNs have become essential in the modern machine learning landscape. Their ability to foster innovation in computer vision has led to advancements in multiple domains, from autonomous vehicles and medical imaging to security systems and entertainment technology. The capacity of CNNs to process complex visual information Their capacity to apply broadly across diverse contexts ensures their enduring importance in shaping the future landscape of ML and AI.

3. Related Work

[1]Analysis of Digital Image Forgery Detection using Adaptive Over-Segmentation Based on Feature Point Extraction and Matching.

The literature review on image fraud detection highlights progress in identifying image tampering. Traditional techniques like principal component analysis (PCA), fast copy-motion detection, and the Fourier-Mellin transform are commonly used, while feature-based methods like Speeded Up Robust Features (SURF) and (SIFT) often face computational efficiency challenges.This study proposes a novel approach that combines feature score matching with adaptive oversegmentation, dividing images into irregular parts to increase processing efficiency. Extracted feature points are analyzed for potential tampering, with additional accuracy provided by a new false area extraction method using advanced morphological techniques and superpixel-based replacement.This innovative approach addresses computational limitations while improving accuracy, offering a significant advancement in image fraud detection and opening new avenues for further research.

[2] Image Falsification Detection using Adaptive Over-Segmentation and Feature Point Matching.

The literature introduces a new strategy for detecting copy-and-move forgeries in images by combining keypoint-based methods with adaptive oversegmentation. This approach dynamically divides the source image into irregular blocks, enabling the detection of falsification through feature point matching. By examining these points within each block, this technique can reveal duplicated regions that might indicate tampering. This method takes it a step further by replacing key points with superpixels, allowing connected blocks to be grouped into larger regions. Moreover, it uses morphological processes to identify suspect areas, thereby enhancing the ability to detect manipulated regions. A deletion algorithm then restores the original image by removing falsified sections. Experiments shows that this new approach outperforms existing methods in several scenarios. This comprehensive method significantly improves the accuracy and flexibility of falsification detection in digital image forensics. It offers a valuable solution for identifying and correcting image tampering in digital copies and

transmissions, representing a meaningful advancement in the field.

[3]A Study of Copy-Move Falsification Detection Based on Segmentation.

Block-based methods are heavily used for copy-move forgery detection (CMFD), where images are split into smaller sections to extract distinctive features for analysis. However, this technique requires considerable searching, leading to significant time consumption. CMFD can be further complicated by transformations like scaling, rotation, and translation, which can mask signs of manipulation. Traditional methods typically rely on shape and color properties, while others utilize frequency-based analysis, such as the Fourier transform. These approaches, however, can struggle when images are edited with noise or blurring, making it challenging to identify forgery. If strong transformations occur, such as rotation or scaling, it becomes nearly impossible to recognize manipulated regions, especially when affine transformations are involved. To overpower these limitations, some different methods, like Self-Adaptive Transform Selection (SATS), focus on adapting to these changes, while others employ texture analysis with wavelet transforms (Discrete Wavelet Transform, DWT) and cosine transforms. Although these approaches offer improved accuracy, detecting forgeries in complex scenarios remains a challenging task.

[4] Image forgery detection using adaptive oversegmentation and feature score matching.

The presented work outlines an innovative program aimed at detecting digital image fraud and correcting copy-move forgeries. This approach uses adaptive super-segmentation and comparative analysis to improve the effiecency of fraud detection. To further isolate manipulated regions, the false region extraction algorithm replaces signals with small super-pixels, acting as uniform shape blocks. Local image blocks with similar color features are then merged into composite regions, enabling more precise identification of image tampering. The work also suggests extending this methodolgy to other types of manipulation, like spatial distortions, and exploring its application in multiple media formats, including video and audio. This expansion could significantly enhance forensic capabilities, allowing for a comprehensive approach to detecting digital manipulation across a wide range of contexts. By incorporating these advanced techniques, the study could have a substantial impact on the field of digital forensics, promoting stronger methods for detecting and addressing digital fraud.

4. Dataset

The dataset described in this above paper obtained from an open-source image forgery detection project, designed to promote research

and developed in the field of digital forensics. This dataset contains manipulated and authentic images for use in training and testing image forgery detection algorithms.

- Each subject was scanned on multiple iteration.
- There are both living and non-living pictures on which model is trained.
- 35 pictures were trained and characterized as 'original' and 'forged'..

5. Architecture

An architecture for image falsification detection typically involves several critical steps and components to effectively identify manipulated images. The process starts with data ingestion, where images are loaded for analysis, including a mix of authentic and manipulated examples used for training and validation.

The preprocessing stage involves normalizing image sizes and formats to ensure consistency. Data augmentation might also be applied, using transformations like rotation, scaling, or flipping to create a more varied dataset.

Next comes feature extraction, where the image is split into smaller blocks or segments. Keypoint detection algorithms, such as SIFT, SURF, or ORB, are used to find distinctive points within each block or across the whole image. Once detected, feature descriptors are extracted, employing techniques like gradient analysis, edge detection, color histograms, or frequency domain analysis using the Fourier transform.

Forgery detection follows, starting with similarity analysis, where feature descriptors are compared to identify duplicacy regions, a common sign of copy-move forgery. Machine learning models, such as (SVM), Random Forest, or Convolutional Neural Networks (CNN), are then employed to classify images as original or forged based on these features. In some cases, anomaly detection is used to detect region in the image that deviate significantly from the norm, suggesting potential manipulation.

In the post-processing stage, morphological operations like dilation and erosion help clean up detected regions, improving accuracy. Further analysis is conducted to determine which parts of the image were flagged as forged and the significance of these findings.

The output and visualization phase creates forgery masks to highlight the detected forged regions and generates a report summarizing the forgery detection results, detailing the type of forgery and the areas affected. Optionally, a user interface can be built to visualize results and

allow user interaction with the detection system.

Finally, performance evaluation ensures the robustness of the system, using metrics such as precision, recall, F1-score, and ROC-AUC. Cross-validation is employed to test the system with different data splits and configurations, providing a comprehensive assessment of the system's accuracy and reliability. This architecture offers a flexible framework that can be adapted to suit specific requirements and incorporates the latest advances in image falsification detection.

6. Algorithm

1. Start

2. Import necessary libraries: Flask, tf learn, request, tqdm, shutil, pandas, sqlite3.

3. Initialize the Python application.

4. Load the trained model from the training data set.

5. Define a function to preprocess input data.

6. Define routes for different pages:

a. Route for the home page

b. Route for the team page

c. Route for the recognize page

7. Define the function for the home page

a. If request method is POST:

i. Get input from user

ii. Preprocess the user input.

iii. Feature extraction of image

iv. segmentation of image

v If the given image is forged analayze the forgery percentage

vi. otherwise the given image is original with accuracy

b. If request method is GET, render the 'index.html' template.

8. Run the python application.

9. Stop

7. CONCLUSION

In conclusion, this literature review presented a novel and accurate technique for detecting fake images by merging point key matching with adaptive over segmentation. This viewpoint has demonstrated superior accuracy in pinpointing falsified areas in images. Future studies could aim to refine this method and investigate its practical applications to enhance the reliablity and integrity of digital images. By further developing these techniques, we can better safeguard against image manipulation and ensure the reliability of digital media.

8. REFERENCES

[1] Analysis of Digital Image Falsification Detection using Adaptive Over Segmentation Based on Feature Point Extraction and Matching
Ch. SUDARSHAN, U. SATHISH KUMAR

[2] Image Forgery Detection using Adaptive Over-Segmentation and Feature Point Matching

RAVI BABU KANCHARLA, NAGI HYMAVATHI

[3] A Study of Copy-Move Forgery Detection Scheme Based on Segmentation
Mohammed Ikhlayel , Mochamad Hariadi and Ketut Eddy Pumama

[4] Image Forgere-Detection Using Adaptive Oversegmentation and Feature Score Matching
Chi-Man Pun, Senior Member, IEEE, Xiao-Chen Yuan, Member, IEEE, and Xiu-Li Bi

