INTERNATIONAL JOURNAL
FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Digital Image Forgery Detection using Deep Learning

Harsha Jain[1], Prathmesh Jamale[2], Jankee Patil[3], Surbhi Kate[4], Sakshi Loni[5]

Dept. of Computer, KJ College of Engineering and Management Research Pune, India

Abstract: In recent years, the rise in accessibility and simplicity of image editing tools has led to a surge in the production and dissemination of fake or manipulated images across media and the internet. Numerous techniques have been introduced to verify the originality of images, for certain cases, to pinpoint the specific regions that have been altered or forged. This paper offers a review of several of the latest image forgery detection methods, particularly those based on Deep Learning (DL) frameworks, with a focus on prevalent attacks such as copy-move and splicing.

The growing phenomenon of Deepfake content is also discussed, especially in its application to images, which results in effects similar to splicing.

The relevance of this survey is highlighted by the fact that deep learning-based methods currently offer some of the best performance on standard datasets. We explore the essential features of these methods and describe the datasets used for their training and validation.

Furthermore, we assess and compare their performance where applicable. Based on this examination, we suggest potential future directions for research, particularly in the areas of deep learning architectures, evaluation methodologies, and the development of datasets to facilitate easier comparison of these techniques.

Keywords: Fake Images, Altered Images, Image Authentication, Forgery Localization, Deep Learning Techniques, Copy-Move Attack, Splicing Attack, Image Editing Tools, Deep Learning Architectures.

I. INTRODUCTION

The rapid adoption of smart devices globally, combined with the integration of high-quality cameras and advanced image processing tools, has significantly transformed how digital visual data is handled. Modern smartphones, tablets, and computers, now nearly always connected to the internet and cloud-based servers, have empowered individuals to capture, store, and process vast amounts of digital images and videos—something that would have seemed unimaginable just a few years ago. This seamless connection between devices and remote servers through the internet enables people to manage and exchange massive volumes of visual content effortlessly.

As a result, images and videos are increasingly used as primary sources of information across numerous fields, from casual social interactions to professional contexts. Whether it's a news story, a scientific study, or everyday events, smartphones and other portable devices have made it easy for people, including professionals, to document what's happening around them.

The explosion of visual content sharing is fuelled by various digital technologies, including efficient data compression techniques, high-speed internet connections, and user-friendly applications designed specifically for managing and distributing media. In particular, platforms such as social media networks (e.g., Instagram) and online forums (e.g., Reddit) have revolutionized the way visual content is shared, allowing users to instantly upload and spread images and videos across the world.

The rapid dissemination of such content has become almost second nature to users, enabled by these highly specialized digital tools. On the other hand, the widespread availability of sophisticated yet easy-to-use image editing software—both commercial applications like Adobe Photoshop and free, open-source alternatives like GIMP—means that practically anyone can alter digital images.

Additionally, smartphone-based applications provide users with on-the-go editing capabilities, allowing basic image manipulation to be performed in just a few taps. This combination of high-quality visual data capture, advanced processing tools, and instantaneous sharing capabilities has fundamentally changed the way society interacts with digital media, making images and videos a central part of how information is consumed and disseminated.
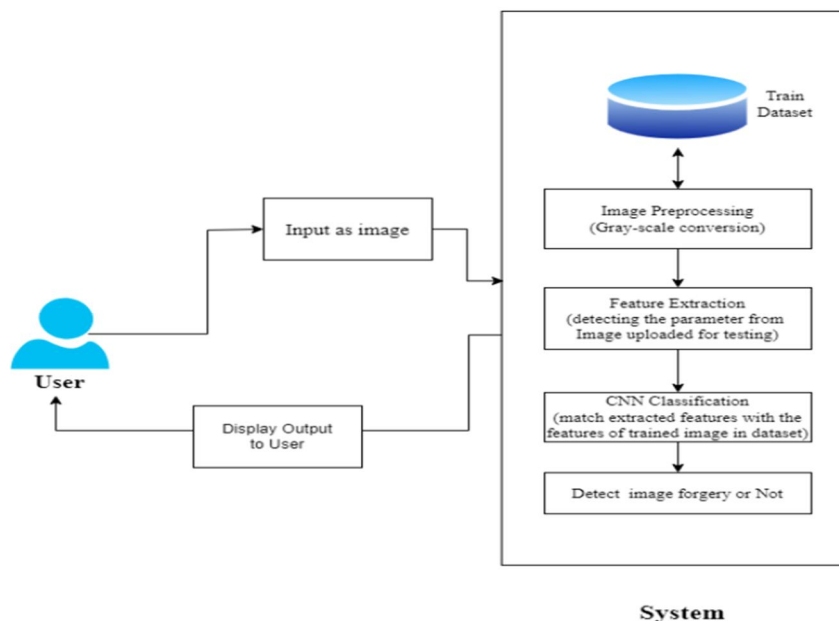
Fig. 1.System Architecture

II. LITERATURE SURVEY

A. An Assessment of Image Forgery Detection Techniques

As digital editing tools become more accessible, image forgery is on the rise, making reliable detection methods crucial for verifying the authenticity of visual content. In their paper published in 2019 by Chandandeep Kaur and Navdeep Kanwal examine various forms of forgery, including copy-move, splicing, and retouching. They categorize detection techniques into two domains: Transform Domain methods, such as Discrete Cosine Transform (DCT) and Discrete Wavelet Transform (DWT), and Spatial Domain methods like Scale Invariant Feature Transform (SIFT) and Speeded Up Robust Features (SURF). These techniques employ classifiers such as Support Vector Machines (SVM) to distinguish between genuine and altered images. The authors evaluate these techniques using datasets such as MICC-F220, MICC-F600, and CASIA v2.0, with accuracy and false positive rates as key metrics. While techniques like SIFT and DWT demonstrate high accuracy, they sometimes struggle with false positives or detecting subtle manipulations. The paper calls for unified detection methods and highlights deep learning's potential to improve adaptability and performance across different types of image manipulation. In the future, the focus should be on using advanced machine and deep learning techniques to create more reliable detection systems with better accuracy and fewer false positives.

B. An Efficient CNN Model to Detect Copy-Move Image Forgery

In 2022, Khalid M. Hosny, Akram M. Mortda, Mostafa M. Fouda, and Nabil A. Lashin introduced a computationally efficient Convolutional Neural Network (CNN) architecture aimed at detecting copy-move image forgery, a manipulation technique where sections of an image are duplicated and pasted within the same image. The model relies on convolutional and max-pooling layers for feature extraction, which identifies patterns indicative of tampering. The CNN serves as the classifier, determining whether images are original or forged. The process consists of three main stages: preprocessing, feature extraction, and classification. During preprocessing, the input image is resized without cropping. Feature extraction involves three convolution layers, each followed by max-pooling to reduce feature dimensions. Finally, a fully connected layer links the extracted features to the classifier, which categorizes the image. The model was tested on MICC-F220, MICC-F600, and MICC-F2000 datasets, achieving 100% accuracy in detecting forgeries, measured through accuracy, precision, recall, and F1-score metrics. This model offers key advantages, including high accuracy, low computational cost, and a lightweight design, making it suitable for practical use. However, its focus is limited to copy-move forgery, and it does not extensively address other forms of manipulation like image splicing or retouching. Future enhancements could expand the model's capabilities to cover more types of forgeries, enhancing its versatility across diverse datasets. Overall, this work significantly advances image forgery detection by offering an efficient deep learning solution.

C. Detection and Localization of Multiple Image Splicing Using MobileNet V1

In 2021, Kalyani Kadam, Swati Ahirrao, Ketan Kotecha, and Sayan Sahu introduced an innovative approach for detecting and localizing multiple image splicing forgeries, where parts from different images are combined into one. Their method employs Mask R-CNN with a MobileNet V1 backbone, creating a lightweight and computationally efficient model. The primary aim of the research is to enhance the accuracy and speed of detecting multiple spliced regions within images, while also calculating the percentage of the manipulated area.

The model was trained and evaluated on several datasets, including MISD (Multiple Image Splicing Dataset), CASIA 1.0, WildWeb, and Columbia Gray. It demonstrated superior performance compared to other methods that use ResNet-based architectures. By incorporating MobileNet V1, the model reduces complexity through depthwise separable convolutions, which help maintain efficiency and speed without sacrificing detection accuracy. Using region proposal networks (RPN) and ROIAlign, the model successfully identifies and segments forged regions at the pixel level.

The results showed notable precision, with an average of 82% on MISD, 74% on CASIA 1.0, 81% on WildWeb, and 86% on Columbia Gray datasets. The key strengths of the model include its computational efficiency, quick processing times, and accurate localization of spliced areas. However, a limitation of the study is the relatively small dataset used for training, which could impact the model's effectiveness on more complex forgery scenarios. Future research may focus on adapting the model to detect other post-processing operations and exploring newer architectures like MobileNet V2 and GANs to further improve detection capabilities.

D. A Full-Image Full-Resolution End-to-End-Trainable CNN Framework for Image Forgery Detection

In their 2020 study, Marra et al. propose a CNN-based technique for detecting image forgeries, offering comprehensive analysis at both full-image and full-resolution levels. The approach utilizes sophisticated deep networks, including Xception, and integrates noiseprints to enhance feature extraction. Classification is performed through two fully connected layers, while the methodology includes extracting patch-level features, aggregating them, and making global decisions. To optimize memory, gradient checkpointing is employed. The model was tested on both synthetic datasets (Vision, UCID, Dresden, FAU) and realistic datasets (DSO-1, Korus, NC2017, MFC2018), with the Area Under the ROC Curve (AUC) serving as the key evaluation metric. The proposed method surpasses existing state-of-the-art models, maintaining high-frequency detail and supporting end-to-end training. Despite challenges related to computational demands and memory usage, the framework demonstrates strong potential for detecting localized forgeries, such as splicing, copy-move, and inpainting. Future research could aim at improving its efficiency, highlighting its valuable contribution to the field of digital image forensics.

E. CNN-Key point Based Two-Stage Hybrid Approach for Copy-Move Forgery Detection

In their 2024 study, Anjali Diwan and Anil K. Roy introduce a hybrid solution for detecting copy-move forgery in digital pictures. This method combines traditional techniques with advanced algorithms to enhance accuracy in identifying duplicated regions within an image. By integrating both feature extraction and matching processes, the proposed system offers a more robust solution for detecting forgery, even in complex scenarios. The method has been tested on various datasets and shows promising results in terms of precision and recall, making it an effective tool for digital image forensics. The approach utilizes CenSurE keypoint detection combined with a CNN architecture for feature extraction. A fusion classifier integrates outputs from both the CenSurE keypoints and CNN features for more accurate forgery detection. The model is evaluated using several datasets like CMFD, GRIP, CoMoFoD, MICC-F600, MICC-F220, COVERAGE, and CASIA-II, with performance measured through accuracy, precision, recall, and F1-score. The system shows strong performance in detecting forgeries, even in cases with complex post-processing or geometric transformations. Key advantages include robustness against various attacks and the potential to manage diverse textures by combining the strengths of both keypoint and CNN-based approaches. However, the system's computational complexity poses a challenge, especially when dealing with extremely smooth or highly textured images. The paper suggests that future work could focus on improving detection accuracy, reducing computational costs, and addressing more sophisticated forgery techniques. This study contributes to the multimedia forensics domain by integrating keypoint detection with deep learning for enhanced forgery detection.

F. Copy-Move Forgery Detection: A State-of-the-Art Technical Review and Analysis

The paper above by Songpon Teerakanok and Tetsutaro Uehara (2019) provides a comprehensive review of recent advancements in copy-move forgery detection (CMFD). The study explores various feature extraction techniques, including transformation-based, hashing, LBP-based, keypoint-based, histogram-based, color/intensity-based, and statistical moments.

While no specific classifier is mentioned, the focus is on the detailed analysis of feature extraction and matching techniques. A new CMFD process pipeline is introduced, which categorizes techniques at each stage of the detection process. The paper references datasets like CoMoFoD and evaluates methods based on accuracy, speed, and robustness against attacks including noise, compression, resizing and rotation. The review highlights the trade-offs between speed and accuracy in different methods and the robustness of various techniques. While the advantages include a comprehensive review and the introduction of a new pipeline, some methods face challenges such as high computational complexity and limitations in real-time applications. The future scope suggests improving detection accuracy, reducing complexity, and enhancing robustness. This review is a significant contribution to the digital forensics and image processing domains.

G. Copy-Move Forgery Detection Based on Keypoint Clustering and Similar Neighborhood Search Algorithm

The paper above by Haipeng Chen, Xiwen Yang, and Yingda Lyu (2020) presents The detection of copy-move forgeries can be improved by employing keypoint clustering techniques coupled with neighborhood search algorithms. This method increases the precision of forgery localization. Although no specific classifier is mentioned, the methodology involves clustering SIFT keypoints based on scale and color, followed by matching them. The algorithm that is used novel localization to differentiate similar neighbourhoods. The approach is evaluated using datasets like GRIP, Dataset, and FAU, with performance metrics including True Positive Rate (TPR), False Positive Rate (FPR), and F1-score. The method shows superior performance in terms of matching time complexity, detection reliability, and forgery location accuracy, though it struggles with large-scale forgery detection. Key advantages include reduced matching time complexity and accurate localization of tampered regions, while the disadvantages lie in its poor performance with large-scale forgeries. The future scope suggests improving detection for large-scale and post-processed forgeries. This study contributes to the domain of digital image forensics by enhancing the efficiency and exactness of copy-move forgery detection.

H. Enhancing Digital Image Forgery Detection Using Transfer Learning

The above paper by Ashgan H. Khalil, Atef Z. Ghalwash, Hala Abdel-Galil Elsayed, Gouda I. Salama, and Haitham A. Ghalwash (2023) presents a technique that leverages transfer learning to boost the detection of image splicing and copy-move forgeries. The system uses the difference between the real image and its reduced version for feature extraction and employs pre-trained models like MobileNetV2, VGG16, and ResNet50 for binary classification. The method is tested on the CASIA 2.0 dataset, with performance measured through accuracy, recall, precision, F1 score, and AUC. Among the models, MobileNetV2 achieved the highest detection accuracy of approximately 95%. Key advantages include high detection accuracy, faster training times, and reduced computational costs. However, the method faces challenges like limited generalization to unseen data and computational complexity. The future scope suggests further improving detection accuracy, introducing forgery type detection and localization, and expanding capabilities to detect forged videos. This study contributes to the field of digital image forgery detection, offering a transfer learning-based approach that balances efficiency and accuracy.

I. Copy-Move Forgery Detection Technique Using Graph Convolutional Networks Feature Extraction

In this 2024 paper, Shinde et al. introduce a novel method for copy-move forgery detection by utilizing Graph Convolutional Networks (GCNs) for feature extraction. The approach involves converting images into graph representations to capture spatial and structural relationships, which are then processed by GCNs. Support Vector Machine (SVM) is employed for classification, providing a robust framework for detecting forgery regions. The methodology is evaluated using the MICC F220 and CoMoFoD datasets, with performance metrics including accuracy and F1 score. The results demonstrate high accuracy and robustness in identifying copy-move forgery, highlighting the method's advantages in handling complex forgery patterns. However, the paper notes challenges related to computational cost and complexity. Future research could explore further advancements in GCNs and other neural network architectures, emphasizing the technique's potential in the field of digital image analysis.

J. Modelling of Reptile Search Algorithm With Deep Learning Approach for Copy Move Image Forgery Detection

In their 2023 study, Maashi et al. introduce an creative method for finding copy-move forgery in images. Their approach combines Neural Architecture Search Network (NASNet) for feature extraction with the Reptile Search Algorithm (RSA) for optimizing hyperparameters, and Extreme Gradient Boosting (XGBoost) for classification. This hybrid model significantly improves the detection of forged areas within images by optimizing both feature selection and classification processes, leading to enhanced accuracy and efficiency. The methodology is evaluated on benchmark datasets such as MNIST and CIFAR-10, utilizing a variety of

evaluation metrics, including precision, recall, accuracy, F-score, training and validation accuracy, training and validation loss, and the ROC curve. The results indicate that the proposed model achieves high precision, recall, accuracy, and F-score, demonstrating superior performance compared to recent techniques. While the paper does not specify any disadvantages, it highlights the advantages of effective feature extraction and optimized hyperparameters. Future research directions include enhancing the model with hybrid deep learning methods and testing it on real-time data, emphasizing its potential applications in cybersecurity and image forensics.

K. Unveiling Copy-Move Forgeries: Enhancing Detection with SuperPoint Keypoint Architecture

In their 2023 paper, Diwan et al. introduce an innovative approach to detecting copy-move forgeries in digital images using the SuperPoint keypoint model. The methodology begins with extracting keypoints from the image and calculating descriptors, which are distinctive features that help identify specific areas in the image. Following this, the keypoints are matched, and outlier data is eliminated by applying clustering techniques such as Fuzzy C-Means and RANSAC. This approach effectively handles the removal of outliers, improving the accuracy of the forgery detection process. The technique is evaluated on several well-known datasets, including CMFD, CASIA V2.0, MICC-F220, MICC-F600, CoMoFoD, COVERAGE, and GRIP. To measure the method's effectiveness, performance metrics like precision, recall, and F1-score are used. The results show that the proposed method excels in detecting copy-move forgeries, demonstrating both high precision and recall, which highlights its strong ability to identify manipulated regions in images. Furthermore, the approach proves to be efficient, making it fit for real-time applications in multimedia forensics and digital image forgery detection. While the paper does not explicitly mention any drawbacks, it suggests that future work could focus on improving the method's robustness and efficiency, as well as exploring the potential of other deep learning models to further enhance detection accuracy.

L. An Analysis of Image Forgery Detection Techniques

In their 2019 study, Kaur and Kanwal provide an in-depth review of image forgery detection techniques, highlighting the increasing demand for effective methods as digital image manipulation becomes more prevalent. The paper explores various types of image forgery, such as copy-move forgeries, image splicing, and image retouching, and examines blind detection methods that do not require access to the original image. The review also covers a variety of feature extraction techniques, including Discrete Cosine Transform (DCT), Principal Component Analysis (PCA), Scale-Invariant Feature Transform (SIFT), Speeded-Up Robust Features (SURF), and Local Binary Patterns (LBP). Several classifiers are discussed, such as Support Vector Machine (SVM), Recursive Feature Elimination (RFE), PCA, and Kernel PCA (KPCA). The study evaluates different datasets like MICC-F2000, MICC-F220, MICC-F600, CoMoFoD, CISDE, CUISDE, and CASIA v1.0 and v2.0, measuring performance in terms of accuracy, computational efficiency, and robustness against various transformations. The analysis provides a comprehensive comparison of techniques, emphasizing their strengths, including accurate detection, low computational cost, and reliability. However, some methods face challenges when dealing with noisy images, small tampered regions, or specific transformations. The authors propose future research directions, such as the creation of more unified and robust methods, extending forgery detection to audio and video formats, and further advancing deep learning-based approaches, making this paper a significant contribution to the field of digital image forensics..

M. Framework for Image Forgery Detection and Classification Using ML

In their 2021 paper, Nande et al. introduce a detailed framework for detecting and classifying digitally altered images through machine learning methods. The authors focus on pattern recognition using Gray Level Co-occurrence Matrix (GLCM) features and Convolutional Neural Networks (CNNs). The process involves essential preprocessing steps, including histogram equalization and noise removal, followed by image segmentation using k-means clustering to prepare the data for further analysis. The dataset used in the study contains 120 pairs of original and manipulated images, sourced from the internet and edited with Adobe Photoshop CC 2017. The proposed method's performance is evaluated based on accuracy, achieving an impressive result of 96.4%. While the CNN-based approach demonstrated exceptional accuracy and effectiveness in detecting and classifying forged images, the initial classifier, Support Vector Machine (SVM), showed a lower accuracy of 87.6%. The authors highlight the potential for further improvements by exploring the use of more advanced image editing tools, suggesting directions for future research to enhance forgery detection capabilities. This study contributes significantly to the field by integrating traditional machine learning techniques with modern deep learning methods, providing a more robust and accurate approach to image analysis.

III. METHODOLOGY

A. Image Acquisition

Objective: Collect a varied dataset that includes both genuine and potentially forged images to ensure thorough training and testing. This dataset will provide a broad spectrum of examples, enabling the model to learn to differentiate between real and manipulated pictures effectively. A diverse set of images will enhance the robustness of the model and improve its ability to generalize across different types of forgeries.

Approach: Utilize publicly available datasets (e.g., CASIA, CoMoFoD, MICC) and user-contributed images to create a robust dataset representative of real-world scenarios.

B. Preprocessing

Once the dataset is established, the next phase is preprocessing. This step is vital for cleaning and preparing images for analysis, which may include resizing images to standard dimensions, applying noise reduction techniques to minimize artifacts, and enhancing image quality through methods like histogram equalization. These preprocessing techniques are instrumental in ensuring that the images are in an optimal state for the subsequent steps, as they help to eliminate any potential distortions that could affect the accuracy of the detection process.

C. Feature Extraction

Following preprocessing, the methodology transitions into feature extraction. This stage focuses on identifying and extracting relevant features within the images that can aid in forgery detection. Various techniques are employed, such as analysing pixel values for inconsistencies, utilizing Local Binary Patterns (LBP) or Gabor filters to capture texture variations, and extracting metadata like EXIF data to gain insights into the manipulation history. By thoroughly analysing these features, the model can better distinguish between authentic and forged images, Genre Selection.

D. Model Development

The next component involves model development, where a machine or deep learning model is created to classify the images. This process typically utilizes Convolutional Neural Networks (CNNs), which excel in automatic feature extraction and classification tasks. Additionally, transfer learning techniques can be utilized by leveraging pre-trained models such as VGG16 to improve detection performance. By using models that have already been trained on large datasets, these approaches capitalize on existing knowledge, significantly boosting the model's ability to recognize features relevant to forgery detection. This not only enhances the efficiency of the training process but also accelerates convergence, reducing the time needed to achieve optimal performance.

E. Training

After the model is developed, the training phase commences. During this stage, the model is taught on labelled dataset to identify patterns associated with forgery. Supervised learning techniques are applied, with hyperparameters adjusted to optimize performance. Additionally, cross-validation methods are implemented to enhance the model's ability to generalize, ensuring it can accurately classify unseen images and maintain reliable performance across diverse datasets.

F. Forgery Detection Techniques

The system employs distinct techniques for detecting different types of image forgeries. It identifies copy-move forgeries by locating areas within the image that have been duplicated. It also detects image splicing by examining the inconsistencies at the borders where images are merged. Furthermore, the system enhances its detection abilities by analysing artifacts of manipulation, including noise patterns and compression traces.

G. Deployment

Once the model is verified, the next step involves deployment. A detection system is integrated into practical applications, which may include social media platforms, news organizations, and digital forensics labs. Deploying the system in real-world environments allows it to aid in the detection of manipulated content, addressing the growing concern over image authenticity in the digital age.

H. Maintenance and Updates

The methodology highlights the critical role of continuous maintenance and updates in sustaining the effectiveness of the forgery detection system. As new forgery techniques emerge, regular updates allow the system to evolve, ensuring it can accurately detect the latest manipulation methods. By integrating updated datasets and refining detection algorithms, the system's effectiveness is continually enhanced, ensuring it remains reliable and accurate in identifying forgeries over time. This ongoing development is vital for maintaining the system's performance and relevance in a rapidly changing digital landscape.

IV. CONCLUSION

The conclusion of image forgery detection can be briefed as follows: Image forgery detection is a crucial field in digital forensics and computer vision. It involves the identification of manipulated or counterfeit images to maintain the virtue of digital content. Various techniques and algorithms are developed to identify various types of image forgeries, such as copy-move, splicing, and retouching. In recent years, deep learning-based approaches, including convolutional neural networks (CNNs) and generative adversarial networks (GANs), have shown promising results in image forgery detection. These methods can effectively identify subtle alterations in images and provide reliable forgery detection. Despite the advancements in image forgery detection, challenges remain, such as the emergence of more sophisticated forgery techniques and the need for big and distinct datasets for training. Researchers and practitioners in this field continue to work on improving the correctness and strength of forgery detection techniques. In conclusion, image forgery detection plays a vital role in ensuring the realness of digital images. It is a dynamic field that continues to evolve with the advancement of technology, and ongoing research is essential to stay ahead of increasingly sophisticated forgery techniques

REFERENCES

[1] Abdalla Y, Iqbal T, Shehata M (2019) "Copy-move forgery detection and localization using a generative adversarial network and convolutional neural-network. Information" 10(09):286. https://doi.org/10. 3390/info10090286 Multimedia Tools and Applications (2023) 82:17521–17566 17561

[2] Achanta R, Shaji A, Smith K, Lucchi A, Fua P, S¨usstrunk S (2010) Slic superpixels. Technical report, EPFL

[3] AdobePhotoshop. https://www.adobe.com/it/products/photoshop.html. Accessed 16 Mar 2022

[4] Agarwal R, Verma O (2020) An efficient copy move forgery detection using deep learning feature extraction and matching algorithm. Multimed Tools Appl 79. https://doi.org/10.1007/s11042-019- 08495-z

[5] Amerini I, Ballan L, Caldelli R, Del Bimbo A, Serra G (2011) A SIFT-based forensic method for copy-move attack detection and transformation recovery. IEEE Trans Inf Forensics Secur:1099–1110. https://doi.org/10.1109/TIFS.2011.2129512

[6] 6.X. Lin and C. T. Li, "Large-scale image clustering based on camera fingerprints," IEEE Transactions on Information Forensics and Security, vol. 12, no. 4, pp. 793–808, April 2017.

[7] M. Goljan, M. Chen, P. Comesana, and J. Fridrich, "Effect of compression on sensor-fingerprint based camera identification," vol. 2016, pp. 1–10, 02 2016.

[8] I. Amerini, T. Uricchio, and R. Caldelli, "Tracing images back to their social network of origin: A cnn-based approach," in 2017 IEEE Workshop on Information Forensics and Security (WIFS), Dec 2017, pp. 1–6

INTERNATIONAL JOURNAL
FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Call : 08813907089   (24*7 Support on Whatsapp)