Library Progress International *Print version ISSN 0970 1052*
Vol.44 No.3, Jul-Dec 2024: P.4491-4497 *Online version ISSN 2320 317X*

Original Article *Available online at www.bpasjournals.com*

Copy Move Image Forgery Detection Using Keypoint Based Approach

[1] Smruti Dilip Dabhole[*], [2]G.G Rajput, [3]Prashantha

[1,2] Department of computer science, Karnataka State Akkamahadevi Women University Vijayapura, Karnataka INDIA - 586108
[3] Rani Channamma University, Belagavi India.

How to cite this article: Smruti Dilip Dabhole, G.G Rajput, Prashantha (2024). Copy Move Image Forgery Detection Using Keypoint Based Approach. *Library Progress International*, 44(3), 4491-4497.

ABSTRACT
Identifying and detecting portion that has been manipulated in an image is a challenging research area. Therefore, in the given paper, we propose a fusion for copy-move forgery area detection based on identifying Scale Invariant Features in an image without using any reference image. Here, features are extracted using SIFT algorithm and matched using Brut force matcher. The identical portions are clustered using BIRCH clustering to detect the forged portions in an image. In our study, we considered natural images which are tampered using image manipulating tools. The experiments are performed on publicly available datasets Viz., MICC-F220, MICC-F600 and results obtained are compared with other existing methods in the literature along with provided ground truth images.

KEYWORDS
BIRCH, Brut Force Matching, Image Tampering, SIFT

INTRODUCTION

Today millions of images are captured using digital tools such as cameras, CCTV, smart phones, and scanners. These images are stored in a database and are utilized for different applications in the field of journalism, forensics, multimedia security, invoices, scientific publications, and in document verification. The advancement in image manipulation techniques has made it possible to tamper the image contents without leaving any visible clues in the images and thus delivering forged information. Various image editing tools like Photoshop, Paint, and Picasa are available; there are many chances of damaging the shreds of evidence. Hence, automatic computer assisted tools are essential for image verification [1,2] and document fraudulent detection to find the authentication of a document or identity card [3,4] as the manual task is cumbersome.

Among several approaches [5] used in image tampering, "copy-move" refers to creation of new content in an image by taking portion from the image itself. That is, a forgery duplicates a region within an image. Copy-move is generally used to increase the number of existing objects by covering it with a part of the image background. Since qualities in the image, such as illumination, proportion, and focus are not affected, such an image has a higher likelihood of leaving no evidence of tampering [6]. There are two commonly used approaches in "copy-move" forgery detection [7]; block-based approach and forgery and a keypoint-based approach [8,9]. Various techniques are found in literature to detect the forgery from an image like Block-based DCT for feature extraction [10], and local binary pattern for detection [11]. For clustering and matching various algorithms are used such as K-Means clustering, hybrid clustering, and FANN [12,13].These methods attempt to detect the forged content without using any reference image for detection and hence are termed as single image forgery detection methods

1. Related Work

Azra Parveen et al. [10] proposed "Block-based copy-move forgery detection using DCT", here, gray image is divided into 8X8 overlapping blocks, features were extracted using DCT based on various feature sets, and blocks were grouped

using K-means clustering algorithm then features were matched using radix sort, the clustering method is used to speed up the matching in block matching, but it increases the time to detect forged parts from an image and model fails on other sets of images. Osamah M. Al-Qershi et.al.[21], proposed a method using overlapping blocks which are clustered using the K-Means algorithm and RANSAC used to remove outliers. The results were detected using a binary detection map. Hesham A. Alberry, et al [22] presented a model for "copy-move detection using Fast SIFT techniques for forensic images". Features were extracted using "Scale Invariant Feature Transform" with Fuzzy C Means clustering. The work has been implemented on dataset MICC-220 and given better results in case of a decrease in execution time. Badal Soni et.al.[17] proposed method highlights various block-based techniques employed in copy-move forgery detection such as SVD, PCA, FFT and many more implemented on various datasets with scaling and rotational attacks.

Priyanka et.al. [20] addresses the growing need to authenticate digital images by proposing a novel technique for detecting copy-move forgery. This approach combines Discrete Cosine Transformation (DCT) and Singular Value Decomposition (SVD) to enhance robustness against compression, transformations, and noise. A Support Vector Machine (SVM) classifies images as authentic or forged, while K-means clustering localizes forged regions. The method surpasses other state-of-the-art techniques in accuracy, precision, recall, and F1-score for Copy-Move Forgery Detection.

Ankit Kumar Jaiswal et.al. [25] proposed "Shift Invariant SWT and Block Division Mean feature vector" method using overlapping blocks for feature extraction in YCbCr color space. The blocks were subdivided into 4 rectangular and 2 triangular blocks. Even though method detects forgery but computational time increases. Kunj Bihari Meena et.al [27] presented hybrid method which uses fourier mellin for block-based computation and SIFT for keypoint based approach, as SIFT may not extract keypoints from smooth region, image has been divided into smooth and texture part. SIFT applied in texture region and FMT is applied for smooth region. G2NN and patchmatch algorithms were used for matching keypoints and blocks respectively.

Chengyou Wang et.al [16] introduces a novel approach for detecting image copy-move forgery using a combination of accelerated-KAZE (A-KAZE) and speeded-up robust features (SURF). A challenge in many keypoint-based forgery detection methods is acquiring enough points in smoother regions. To address this limitation, the proposed method sets low response thresholds for the A-KAZE and SURF feature detection steps. The innovation continues with the introduction of a correlation coefficient map that delineates duplicated regions through a fusion of filtering and mathematical morphology operations. Rigorous experiments substantiate the efficacy of this method in identifying duplicated areas and its resilience against distortions and post-processing techniques—such as noise injection, rotation, scaling, image blurring, JPEG compression, and hybrid image manipulation. Notably, the experimental outcomes underscore the superiority of the proposed approach over other tested copy-move forgery detection methods.

The methods mentioned above perform effectively in restricted constraints such duplication of the object limited to one or two, dynamic range of intensity values in the image is limited, presence of outliers leads to wrong results and in certain cases computational time is high.

Instead of block-based approach for feature extraction that increases the computation cost [24], in this paper, we present a keypoint based single image forgery detection method. Keypoints are extracted from the input image using SIFT approach. Next, features matched using brute force algorithm. Then, clusters are generated using BIRCH clustering of matching points thereby it correctly identifying forgery portion in an image.

2. Methodology

The proposed approach uses a key-point-based method to identify the forged region inside the same image without utilizing a reference image. The stages involved in CMFD are as follows: feature extraction, feature matching, clustering. (Figure 1)

Gray images include minimal computations to recognize an object compared to color images and it has been noted that grayscale images produce findings with higher accuracy when compared to RGB images [14,15]. Therefore, an RGB image is transfigured to gray using frequent method by taking weighted average of color channels which preserves luminance information. The weights are often chosen to mirror how people perceive color intensity, with green being perceived as being the strongest, followed by red and blue. The formula for this weighted average is:

$$GrayValue = 0.2989 * R + 0.5870 * G + 0.1140*B \quad (1)$$

Where, R, G, and B are the respective red, green, and blue color channel values of the pixel, each ranging from 0 to 255.

3.1 Feature Extraction

Now, SIFT is used to fetch features from a gray image. SIFT was introduced by David Lowe in 1999 [19] to describe local features from an image. SIFT can detect interest points in an image that are invariant to scale changes and are also invariant to rotation, translation, and minor affine transformations. SIFT descriptors are designed to be distinctive and can handle significant viewpoint changes of an object, enabling it to recognize objects from different angles.

The steps involved in extracting features are Difference of Gaussian (DoG) space generation, Keypoints detection, and

Feature description [18]. The scale space of an image is described as function $L(m, n, \sigma)$ which is the convolution of Gaussian kernel $G(m, n, \sigma)$ at various scales with input image $I(m, n, \sigma)$, where σ is a constant factor for true scale invariance. DoG is calculated from eq.(2)

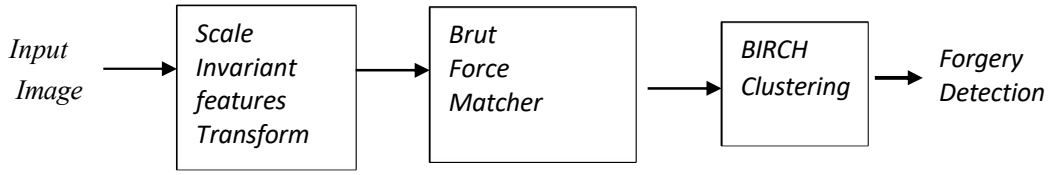

Figure 1 : Proposed Block Diagram

$$D(m,n,\sigma)=(G(m,n.k\sigma)-G(m,n,\sigma))*I(m,n)$$

(2) Followed by, keypoint finding and localization, For a pixel at coordinates (*x*,*y*), its neighbors within a 9×9 region are represented by intensity values $N(x',y')$ for (x',y') within the region. Now in extremum check, a total of 26 comparisons are performed where, V represents intensity of central pixel and $N(x',y')$ represents the intensity of a neighboring pixel at (x',y') therefore, $V > N(x',y')$ *or* $V < N(x',y')$. If either condition is met for any of the neighbors, the pixel is considered a potential keypoint. As keypoints with low contrast and keypoints on edges are not useful they are eliminated. Afterward, assign orientations to the keypoints. Next, created keypoints are with the same location and scale with differences in direction.

A 128-dimensional feature vector is used to represent the extremely unique features. Consider, the set of extracted keypoints and their descriptors denoted as, F = {k1,,….,kn}, and D = {D1,….,Dn}, respectively [13],Now similar descriptor vectors can be found by comparing the descriptor vectors within D from different regions of the given image. This aids in recognizing and matching features.

3.2 Feature Matching

The keypoints and descriptors extracted from an image are further taken for matching. Brute-Force matcher algorithm is applied to extracted features from an image. The method finds correspondences between features in images. For each feature in the image, the algorithm compares its descriptor with all the other features in the image. The goal is to find the best matches based on a similarity metric such as Euclidean distance, Hamming distance, etc. between the feature descriptors based on the keypoints [28,30]

Now, apply algorithm for single image without having any reference image. For each keypoint '*ki'* in the image, '*Di'* be the descriptor associated with '*ki'*. Now compare '*Di'* with the descriptors of all other keypoints in the same image and calculate the similarity metric $S(Di,Dj)$ for each comparison, where '*Dj'* is the descriptor of another keypoint.

To filter out potential false matches, a threshold on the similarity metric is applied. Like Apply a threshold T to the similarity scores, $S(Di,Dj)<T$. Here, only matches with $S(Di,Dj)<T$ are considered valid correspondences. The output of this process is a list of matched keypoint pairs {(k1,k2),(k3,k4),...,(kn−1,kn)}, where each pair contains two keypoints from the same image that are considered to be corresponding based on their similarity in terms of their descriptors.

3.3 Clustering

From previous step the feature matching process extracted matching data points are subject to further analysis. Here, the matching points clustered together to identify the forged area within an image. And, Balanced Iterative Reducing and Clustering using Hierarchies (BIRCH) clustering algorithm is applied.

The Birch clustering algorithm used to cluster two sets of points. It combines `points_1` and `points_2` into a single array called `points`. And then creates a Birch clustering model with the given `threshold` and `branching_factor`[23]. Next Fit the Birch model to the data (points) to get cluster labels. Then filter out noise points (cluster label -1) and get valid points and labels. Divide the valid points array into two sets (`valid_points[:n]` and `valid_points[n:]`), where `n` is half of the total number of points. And this step will provides required output as detecting forgery from tampered image.

3. Result and Discussion

The experiments are performed on the publically available MICC-F220 dataset consisting of images of the "author's [2] personal collection" and a set of images from the "Columbia photographic image repository". MICC-F220 is a balanced dataset of 220 images in that 110 are original and 110 are forged images. The resolution of images ranges from 722 × 480 to 800 × 600 pixels. The tampered area is 1.2% of the image where copied portion from that image is either

rectangular or square randomly in the image and forgery attacks on the image are in way of either rotational or scaling. Further, MICC-F600 dataset has been taken for analysis which includes 440 original images and 160 tampered images. The resolution of images ranges from 800 × 532 to 3888 × 2592 pixels. The tampered region size varies from one image to another [26] and dataset contains ground truth images along with it includes images which have multiple copied contents.

Figure 2: Result of proposed method for MICC-F220 DatasetFigure 2 shows sample images from MICC-F220 dataset row1 show tampered images and row2 shows the corresponding forgery-detected images. The proposed method works good for MICC-F220 dataset. All 110 tampered images are detected as forged with accurate forged region similarly, non-forged images are detected as non-forged. It returns false value for not detecting the forgery. And due to pixel variation very few images from dataset detected as forged which are not forged. The accuracy has been calculated as,

Accuracy = (True Positive + True Negative)/ Total no of images

where, True Positive - The images which are forged and detected as forged.

True Negative - The images which are not forged and detected as not forged.

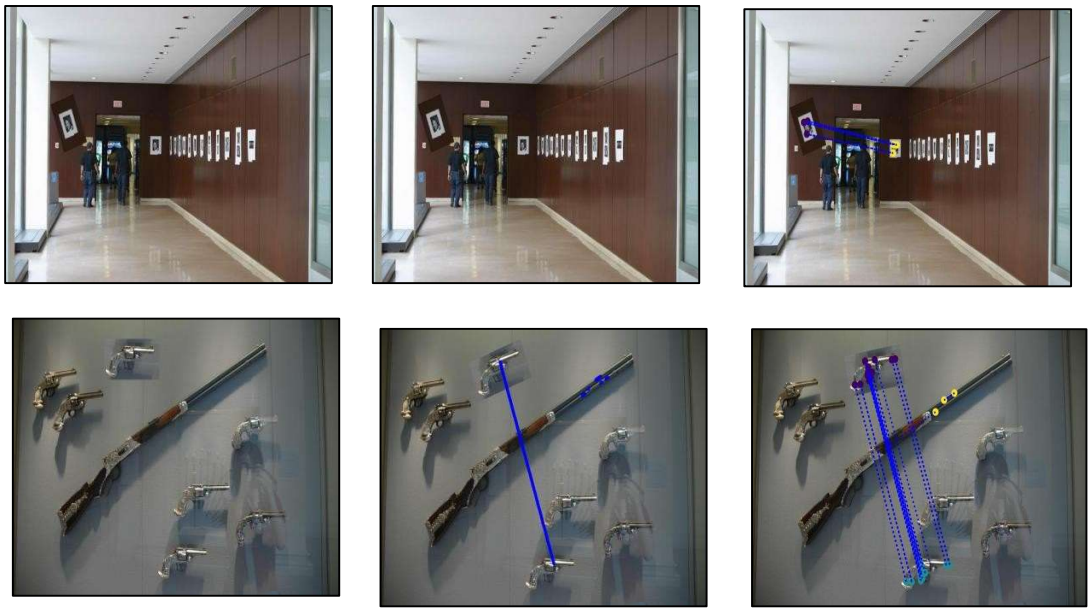

Figure 3: Comparison of Proposed method with SIFT+DBSCAN approach (column 1.Sample input from MICC-F220, column 2. Result of SIFT+DBSCAN and column 3. proposed result

Figure 3 represents a comparative analysis for a proposed method which highlights the robustness and accuracy of

proposed method comparing with SIFT+DBSCAN which shows, very few tampered images are excluded or detected with lower density by this method are detected as forged with proper matching. In Table1. The results of proposed method are compared with existing method for MICC-F220 dataset.

Table 1. Comparative Analysis for MICC-F220 dataset

Author	Method	accuracy
Umair A. Khanet.al. (2018) [11]	Hybrid with Sift	81%
	Hybrid with SURF	89%
	Hybrid with MinEigen	90%
D. Vaishnavi et. al(2019) [29]	Symmetry features	83.64%
Mohamed A. Elaskily et.al [26]	CNN (15epochs)	92.18
Proposed Method	SIFT+DBSCAN **SIFT+BrutForce+BIRCH**	**92%** **96%**

Figure 4. Result of proposed method for Rotation and Scaling on MICC-F220 dataset

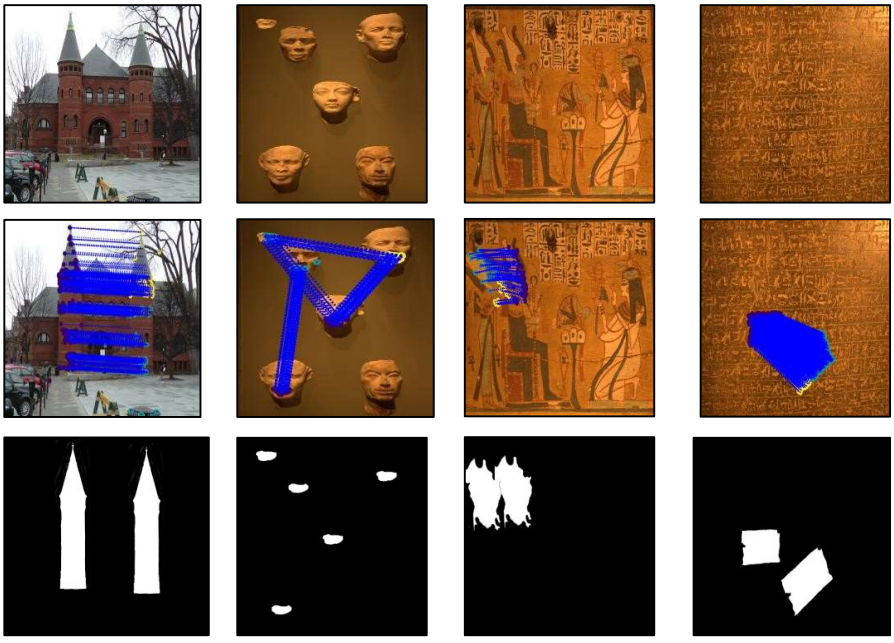

Figure 5: Result of proposed method for MICC-F600 dataset

Figure 4 shows that proposed method yields good result for various attacks like rotation and scaling without generating

false results. Figure 5. shows images from MICC-F600. First row focus on tampered images, second row displays tampered area detection in the images and third row focuses on corresponding groundtruth images which shows proposed method detects accurate forgery. But in some images proposed method generate high false positive rate due to sudden pixel variation along with noisy background of image.

4. Conclusion

The use of keypoints is recommended as an effective and trustworthy technique for copy move forgery detection. Keypoints and descriptors are generated using SIFT algorithm and matched the descriptors using brut force matching method then those points are clustered using BIRCH where detection of forgery depends on the similarity of valid clusters. The proposed method provides good accuracy on MICC-F220 dataset and even though it works fine for MICC-F600 dataset, the images having sudden pixel variations and it cause improper detection for very few numbers of images. Therefore, in the future, it has scope to work on such images which is having non uniform background for forgery detection.

Reference

[1] Vincent Christlein, Christian Riess and Elli Angelopoulou, "*On Rotation Invariance In Copy-Move Forgery Detection*", **2010 IEEE International Workshop on Information Forensics and Security,** 12-15 Dec. 2010

[2] Irene Amerini; Lamberto Ballan; Roberto Caldelli; Alberto Del Bimbo; Giuseppe Serra , *"A SIFT-Based Forensic Method for Copy–Move Attack Detection and Transformation Recovery",* IEEE Transactions on Information Forensics and Security (Volume: 6, Issue: 3, September 2011)

[3] Prasetyo Adi Wibowo Putro, *"An Authentic and Secure Printed Document from Forgery Attack by Combining Perceptual Hash and Optical Character Recognition",* 2019 International Conference on Informatics, Multimedia, Cyber and Information System (ICIMCIS), 978-1-7281-2930-3/19/$31.00 ©2019 IEEE

[4] Lingzhi Li, Jianmin Bao, Ting Zhang, Hao Yang, Dong Chen, Fang Wen, Baining Guo, *Face X-ray for More General Face Forgery Detection*, 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 2575-7075/20/$31.00 ©2020 IEEE DOI 10.1109/CVPR42600.2020.00505

[5] Lilei Zheng, Ying Zhang Vrizlynn L.L.Thing, "A survey on image tampering and its detection in real-world photos", Journal of Visual Communication and Image Representation Volume 58**,** January 2019, Pages 380-399

[6] Hailing Huang, Weiqiang Guo, Yu Zhang, *"Detection of Copy-Move Forgery in Digital Images Using SIFT Algorithm",* IEEE Pacific-Asia Workshop on Computational Intelligence and Industrial Application,2008

[7] Vincent Christlein, Christian Riess, Johannes Jordan, Corinna Riess, Elli Angelopoulou, "An Evaluation of Popular Copy-Move Forgery Detection Approaches", IEEE Transactions on Information Forensics and Security, volume 7, number 6, 2012, pp. 1841-1854

[9] Babak Mahdian, Stanislav Saic, *"Detection of copy–move forgery using a method based on blur moment invariants",* Forensic Science International 171 (2007) 180–189 © 2006 Elsevier Ireland Ltd.

[10] Azra Parveen, Zishan Husain Khan, Syed Naseem Ahmad, *"Block-based copy–move image forgery detection using DCT",* Iran Journal of Computer Science https://doi.org/10.1007/s42044-019-00029-y, Received: 22 June 2018 / Accepted: 8 January 2019 © Springer Nature Switzerland AG 2019

[11] Umair A. Khan, Mumtaz A. Kaloi , Zuhaib A. Shaikh, Adnan A. Arain, *"A Hybrid Technique for Copy-Move Image Forgery Detection",* 3rd International Conference on Computer and Communication Systems 2018

[12] Sunitha K, Krishna A N, "*Efficient Keypoint based Copy Move Forgery Detection Method using Hybrid Feature Extraction",* Proceedings of the Second International Conference on Innovative Mechanism for Industry Applications ICIMIA 2020 IEEE Xplore Part Number: CFP20K58-ART; ISBN: 978-1-7281-4167-1

[13] Aya Hegazi, Ahmed Taha, MazenSelim, *"An Improved Copy-Move Forgery Detection Based on Density-Based Clustering and Guaranteed Outlier Removal",* Journal of King Saud University - Computer and Information Sciences, 2019

[14] Kanan, C., Cottrell, G.W, "Color-to-grayscale: does the method matter in image recognition", PLoS ONE **7**(1), 1–7 (2012)

[15] Saravanan, C, "*Color image to grayscale image conversion*", In: Proceedings of 2nd IEEE International Conference on computer engineering and applications, pp. 196–199 (2010)

[16] Chengyou Wang , Zhi Zhang and Xiao Zhou, *"An Image Copy-Move Forgery Detection Scheme Based on A-KAZE and SURF Features*", Symmetry 2018, 10, 706; doi:10.3390/sym10120706

[17] Badal Soni, Debalina Biswas, *"Image Forensic using Block-based Copy-move Forgery Detection",* 978-1-5386- 3045-7/18/$31.00 ©2018 IEEE

[18] R. Rizal Isnanto, Ajub Ajulian Zahra, Imam Santoso, and Muhammad Salman Lubis, *"Determination of the Optimal Threshold Value and Number of Keypoints in Scale Invariant Feature Transform-based Copy-Move Forgery Detection",* INTL JOURNAL OF ELECTRONICS AND TELECOMMUNICATIONS, 2020, VOL. 66, NO. 3, PP.

561-569 Manuscript received January 12, 2020; revised July, 2020. DOI: 10.24425/ijet.2020.134013

[19] Lowe, David G. 2004. "Distinctive Image Features from Scale-Invariant Keypoints." *International Journalof Computer Vision* 60 (2): 91–110. https://doi.org/10.1023/B:VISI.0000029664.99615.94.

[20] Priyanka & Gurinder Singh & Kulbir Singh, *"An improved block based copy-move forgery detection technique",* @Springer Science+Business Media, LLC, part of Springer Nature 2020

[21] Osamah M. Al-Qershi,Bee Ee Khoo, *"Enhanced block-based copy-move forgery detection using k-means clustering",* Multidimensional Systems and Signal Processing, © Springer Science+Business Media, LLC, part of Springer Nature 2018

[22] Hesham A. Alberry, et al *"A fast SIFT based method for copy move forgery detection",* Future Computing and Informatics Journal 3 (2018) 159-165

[23] G.Nirmala and K.K.Thyagharajan, "A Modern Approach for Image Forgery Detection using BRICH Clustering based on Normalised Mean and Standard Deviation", International Conference on Communication and Signal Processing, April 4-6, 2019, India

[24] Jiangbin Zhenget. al. "*Fusion of block and keypoints based approaches for effective copy-move image forgery detection,"* MultidimSyst Sign Process DOI 10.1007/s11045-016-0416-1, © Springer Science+Business Media New York 2016

[25] Ankit Kumar Jaiswal and Rajeev Srivastava, *"Copy Move Forgery Detection Using Shift Invariant SWT and Block Division Mean Features",* Proceedings of IC3E 2018

[26] Mohamed A. Elaskily & Heba A. Elnemr & Ahmed Sedik & Mohamed M. Dessouky & Ghada M. El Banby & Osama A. Elshakankiry & Ashraf A. M. Khalaf & Heba K. Aslan& Osama S. Faragallah & Fathi E. Abd El-Samie, "A novel deep learning framework for copy-move forgery detection in images", @Springer Science+Business Media, LLC, part of Springer Nature 2020

[27] Kunj Bihari Meena & Vipin Tyagi, "*A hybrid copy-move image forgery detection technique based on Fourier-Mellin and scale invariant feature transforms",* Multimedia Tools and Applications, https://doi.org/10.1007/s11042-019-08343-0, Springer Science+Business Media, LLC, part of Springer Nature 2020

[28] Neema Antony, Binet Rose Devassy, *"IMPLEMENTATION OF IMAGE/VIDEO COPY- MOVE FORGERY DETECTION USING BRUTE-FORCE MATCHING",* Proceedings of the 2nd International Conference on Trends in Electronics and Informatics (ICOEI 2018) IEEE Conference Record: # 42666; IEEE Xplore ISBN:978-1-5386-3570-4

[29] D. Vaishnavi, T.S. Subashini*, "Application of local invariant symmetry features to detect and localize image copy move forgeries",* Journal of Information Security and Applications https://doi.org/10.1016/j.jisa.2018.11.001 2214-2126/©2018PublishedbyElsevierLtd.

[30] Lokesh Sharma, Bhawana Sharma , D P Sharma, *"Implementation of compressed Brute-Force Pattern Search algorithm using VHDL",*