International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 12 Issue 04, April-2023

Critical Study of the Copy-Move Forgery Datasets

Dhanya R[1], R Kalaiselvi[2]
Research Scholar, Noorul Islam Centre for Higher Education.
Associate Professor/CSE, RMK College Of Engineering And Technology,
Puduvoyal, Gummidipoondi, TamilNadu, India.

***Abstract*: The sophisticated softwares and platforms available now-a-days have made the tampering of images very easy. Researchers in this area have proposed and proved many algorithms to solve this problem with its own merits and demerits. Main two challenges any researcher in this field faces is the availability of standard datasets and the right usage of evaluation metrics. This paper makes a critical assessment of the datasets available for copy-move forgery detection evaluating their merits and demerits. By this way, we would like to suggest specification standards of an appropriate CMFD dataset that will help the researchers in this area to make an evaluation of their algorithms in an effective manner.**

Keywords : - Copy-Move Forgery; Forensic Image Processing; Datasets.

1. INTRODUCTION

In this information packed world, digital images occupy a major role as is said 'A picture is worth a thousand words'. But, the availability of various photo editing softwares like Adobe Photoshop has made this saying unacceptable. The ease in the usage of these softwares enabled even an average computer user to make changes to an available image without leaving a trace. The popularity of the social media has also influenced the sharing of forged images. Such a situation has given a negative effect on even considering an image as scientific evidence before law (Mahdian and Saic, 2010). There are active and passive methods to identify a forged image. A passive identification method differs from an active method in in case that the former does not need any additional information whereas the latter uses some kind of additional information to find the forgery (Birajdar and Mankar, 2013).

In passive authentication techniques, Copy-Move Forgery Detection (CMFD) and splicing forms two popular methods, of which CMFD is the most popular (Redi et.al, 2011). CMFD is employed in enhancing or hiding something in the image by simple copy – paste of one part from it into some alternate portion of that image itself. As the cloning is done on the same image, some crucial features like noise, color palette etc. matches with rest image part. The initial work in this area was done by Fridrich et al(2003) and from then on a substantial amount of work has been done and is attracting many researchers in this area. Many algorithms were proposed to tackle this problem and they have challenges and drawbacks [Bakiah et al.,2016; Al-Quershi et. al.,2013]. The criteria on which an apt benchmark dataset is selected and the selection of evaluation metrics plays a very critical role in determining the efficiency of these algorithms.

The widely used common datasets in copy move forgery detection will make clear all the difficulties included. If one dataset comprises of easier copy move forgeries and dataset includes much harder forgeries, the methods using one algorithm will show better result with the first dataset than when it is implemented using the second dataset. Hence it is rather awkward to get into conclusions about the two selected algorithms if they are not tested on the same dataset especially with respect to accuracy (Torralba and Efors,2011). This necessitates the need of a benchmark dataset for comparing and reproducing the results of the prevailing algorithms. This also necessitates for the development of new algorithms.

This paper covers a critical analysis of the existing datasets discussing their merits and demerits. The paper is structured in the following way. Analysing the merits and demerits of the available datasets is done in Section 2. It is followed by the Conclusion in Section 3.

2. BENCHMARK DATASETS

Datasets forms one of the primary reasons for the progress of research in any field and especially in computer vision and signal processing. In these areas, copy move forgery detection is one with committed datasets. Only finite datasets are available online for CMFD and hence are being used by a number of researchers in their work. The authors of the published papers do not provide a valid reason for choosing one particular dataset in their work. These datasets are used for measuring and comparing of the algorithms. Some of the researchers develop their own datasets, but they have to use the same dataset on different algorithms to compare their efficiency. But this comparison becomes difficult as the original source code of most of the work is not available. The exact implementation of the most popular research done in this field is also a time taking process, which is dependent on several factors comprising the amount of images used, image size, programming language and platform used for experimentation etc. The effort need to be put in by the researchers in terms of time and work in the comparison of the existing algorithms with their research work is one of the major difficulties faced. The use of a common dataset as used by the other researchers saves the time and effort. In this paper we describe the characteristics, benefits and drawbacks of some of the online available datasets in this area and will conclude by presenting some considerable suggestions for the selection and design of an ideal dataset.

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 12 Issue 04, April-2023

2.1 CMFD Datasets

The commonly available online datasets in the field of CMFD are MICC (Amerini et.al, 2011), IMD (Christlein et al., 2012), COVERAGE (Wen et al, 2016), CoMoFoD (Tralic et al, 2013), Copy Move Hard (C M H) (Silva et al, 2015), Copy Move Forgery Dataset (Ardizzone et al, 2015), CMFDdb_grip (Cozzolino et al, 2014). CASIA V1 AND V2 (Dong et al., 2013), Columbia (Hsu et al, 2006; Ng et.al, 2004)

2.1.1 MICC dataset

MICC is the most widely accepted datasets in this area. This dataset is subdivided into 4 subsets with initial two subsets created by selecting a rectangular portion from an image and putting above another part of the image followed by a number of affine transformations. Out of the four subsets, MICC –F220 is the most commonly used one (Chihaoui et al.,2014; Hashmi et al., 2014; He et al., 2013; Hsu et al., 2012; Koul, 2014; Uliyan et al., 2015,2016). Some of the images from the subset MICC – F2000 were used by (Kulkarni and Chavan, 2014) and entire dataset found it's way in (Amerini et al, 2013; Yu and Niu, 2016). MICC- F600 was also used by Amerini et al. (2013).

The major demerit of this dataset is that only types of transformations used in this dataset for tampering is rotation and scaling. It does not include noise addition or JPEG compression. Rather than this, ground truth images are missing in three subset of this dataset and thereby the detection accuracy is measured in terms of number of images detected as tampered than pixel level evaluation. The accuracy is considered as cent percent even though only a small portion is detected. The images possess large size which makes it not suitable for block based detection resulting in large computational time. But it was analyzed to be best for key-point based algorithms. Samples from the dataset is given in Fig. 1. The advantage of the dataset is that of gradual attack levels used for thorough examination of dependency among different attack levels and the detection accuracy. MICC dataset is available at http://lci.micc.unifi.it/labd/2015/01/copy-move-forgery-detection-and-localization/.

Subset	Total Images	Orig images	Tamp images	Size of images	Description
MICC – F 220	220	110	110	722 x 480 pixels - 800 x 600 pixels	The threshold was set by the authors themselves
MICC – F2000	2000	1300	700	2048 x 1536 pixels	Created for complete evaluation of every type of modifications
MICC –F8multi	8		8	800 x 532 to 2048 x 1536 pixels	Prepared to identify realistic multiple cloning
MICC – F600	600	440	160+ 160 ground truth images	800 x 532 to 3888 to 2592 pixels	Only deals with block based algorithms

Table 1 – MICC Dataset

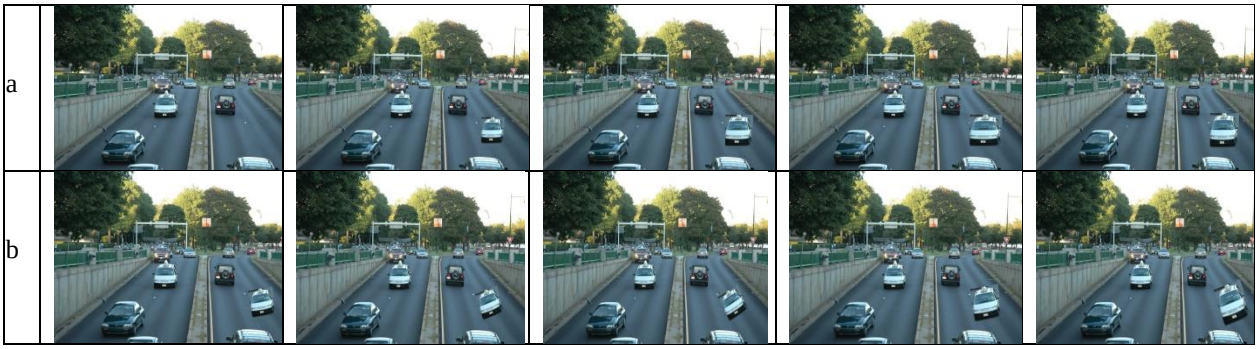

Fig. 1 Images from MICC-F220; First image - original image, All the images in first row are tampered images with scaling. Second row, images with rotation done at various levels combined with scaling

2.1.2 Image Manipulation Dataset (IMD)

Christlein et al (2012) developed Image Manipulation Dataset with the realistic images collected from cameras. The dataset consisted of ground truth images in which the pixels represented as background or elsewise processed the pixels as painted etc. It also includes contour pixels in which the data copied is represented as partially transparent that forms smooth transitions within the adjacent copied pixels and the original pixels, Fig 2. This dataset is used for both key-point based algorithms and block-based algorithms. It is also suitable for detailed study of an algorithm with respect to rotation with varying angles.

Large sized images in the dataset are its major disadvantage. MATLAB like platforms require a large processing time if uses the images in this dataset. The use of programming languages like Python or C++ decreases the processing time, but it is still slower when need to process a large number of images at a stretch. The difficulty in fixing the threshold, which requires a lot of repeated experiments in block based methods, is another disadvantage. The researchers using this dataset require double the effort when he is trying to compare the developed algorithm with prevailing algorithms that did not use this dataset. Because of these demerits, only few researchers made use of this dataset. Pun et.al (2015), Cozzolino et al (2015), Yu et al (2016) used the entire dataset in their work other than the creator themselves. Many of the research works used only a small subset of this dataset approximately equal to 2 % only of the entire dataset, in their work like (Emam et al, 2015; Kumar et al.,2013; Lee et

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 12 Issue 04, April-2023

al.,2014; Maind et al., 2014; Ryu et al., 2013; Satapathy et al., 2014; Silva et al., 2015; Sunil et al., 2014; Uliyan et al., 2015; Yu et al., 2014). This dataset didn't consider some type of transformations like blur and combination attacks including rotation, scaling and additive noise. Another disadvantage is the smooth edges of the ground truth images which does not form part of the copy move area or original images is confusing when considering the accuracy of the algorithm. Dataset is available at https://www5.cs.fau.de/research/data/image-manipulation/.

Subset	Tampering type	Description
orig	48 camera captured high resolution original images	Original images [3000 x 2300]
nul	simple copy move with no transformations	Simple copy with no transformations
Inoise	Images with added noise	Each image added with 5 noise levels
jpeg	Images with added JPEG compression	Each image with 9 quality factors
rot	Images with slight rotation	Rotation of the copied areas with a increment of 2º [2 to 10]
rotExtra	Images with high rotation	Increment of 20º
rotExtra2	Images with extra high rotation	Rotation at a degree of 60 and 180
Scales	Images with slight scaling	Scaling with 91 to 109 %
scaleExtra	Images with high scaling	Scaling with 80 to 120 %
scaleExtra2	Images with extra high scaling	Scaling with 50 to 20%
cmb_easy	Combined attacks – Rotation, Scaling and JPEG compression	Rotation – 2º, Scaling – 110% and compression -80%
cmbExtra	"	Rotation – 4º- 8º, Scaling – 103 – 107 % and compression -65 – 75%
cmbExtra2	"	Rotation – 20º- 60º, Scaling – 120 – 140 % and compression -50 – 60%
multi_paste	The same copied part pasted at multiple sites	Placed at random locations
nul_sd	Downscaled version of nul subset	Scaled down to 50% of the original size
jpeg_sd	Downscaled version of jpeg subset	1500 x 1200
rot_sd	Downscaled version of rot subset	
scale_sd	Downscaled version of scales subset	
scale_down	Downscaled version of scales subset with simple copy paste in the scale of 0.9 to 0.1 of the original size	Image size in the range 2880 x 2000 upto 300 x 230

Table 2 – Image Manipulation Dataset

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 12 Issue 04, April-2023

Fig. 2 Images from IMD: (a) – Original Images, (b) – Rotation by 6º, (c) – Scaled to 1200%, (d)- Combination of rotation and scaling, (e)- Ground truth of (d)

2.1.3 CoMoFoD

This dataset was developed by Tralic et. al(2014), which includes some missing intermediate attacks in the previous datasets, like distortion and color reduction. Two datasets, one with small sized images and another with large sized images were created. These two datasets contain a total of 10,400 and 3120 images respectively that includes 5 categories of transformations and only post processing attacks in 6 categories. The relatively large size and hence the high computational time has made it difficult for any researcher so far to use the entire dataset for evaluating their algorithm. Only some researchers have used this dataset that too a small subset of the whole (Liu and Feng, 2014; Liu et al, 2014; Thajeel and Sulong, 2015; Usubioglu et al., 2016). In the case of intermediate attacks, various angles of rotation and scales were used and it also included distortion and combinations of transformations. For post processing operations, JPEG compression with 9 levels of quality factors were used. Addictive Gaussian White Noise with zero mean and diverse values of variance, Image blurring by convolving the image with 3 x 3, 5 x 5 averaging filters. Change in brightness was attained by original image intensity values mapping. Uniform quantization intensity values for original images resulted in the post processing operation of color reduction. The number of various levels of intensity in the actual image is deducted from 256 to 32, 74 or 128 levels for every single color channel. Color adjustments were also done on the images with mapping of entire input image intensity values range to new bounds of interval.
Wide range of attacks used in this dataset with post processing attacks done at various levels for the in-depth analysis is one of its great advantages. The availability of small images in the dataset makes it readily useful for programming languages like MATLAB. The large number of images possesses another drawback which requires extra computation time and space for the researcher to use the dataset in its entirety. Threshold setting is really difficult as the copy-move area size range extends from extremely small to very large. This dataset is publicly accessible at: http://www.vcl.fer.hr/comofod.

2.1.4 CMFDdb_grip

This dataset was developed by Cossolino et. al (2014) for their own algorithm evaluation. And later they made the dataset able to download . Different parameters and levels were used in this dataset depicted in Table 4.
The advantage of this dataset it the size of the images used in this dataset when matched to the other datasets. The low variances in the size of the pictures are another merit of this dataset. The non-consideration of some attacks like blurring, combinations etc gives the demerits of this dataset. Experiments shows no clear variance in accuracy with the transformations like rotation and scaling done and small angles and scaling factors. Hence the levels can be easily reduced without compromising the accuracy and efficiency when dealing with rotation and scaling in detail. Another problem is that the whole dataset is not directly downloadable, but it facilitates the generation of the dataset using the available code. The change in the values of parameters when done in the code will generate a different dataset with variations that makes it difficult for the evaluational comparability of different algorithms. Dataset is available at http://www.grip.unina.it/index.php?option=com_content&view=article&id=79&Itemid=489&jsmallfib=1&dir=JSROOT/CMFD.

Attack type	**No. of images**	**Description**
Rotation	1120	14 levels – Used different angles of rotation ranging from 0 – 8 with an incremental step of 2º 10-30 with an incremental step of 10º 45 – 105 with an incremental step of 15º 180º
Scaling	1120	14 levels ranging from 0.5 to 2
Noise	480	Addition of Guassian white noise of 0 - 0.10 with normalized standard deviation – 6 levels
JPEG compression	720	9 levels - with quality factor ranging from 20 -100 %

Table 4 – CMFDdb_grip Dataset

2.1.5 Copy Move Hard (C M H)

Silva et.al developed this dataset with not so easy to detect operations. A wide range of parameters and levels are used for creating this dataset. This is a small dataset with only 108 images which makes it less selected for the evaluation of an algorithm by many of the researchers. This datasets is divided into 4 subsets (CMH_{P1}(23 plain images with simple copy-move operations), CMH_{P2}(25 images with rotation within a range of -90 to 180º), CMH_{P3} (26 images with scaling with a factor between 80 to 154%), CMH_{P4}(34 images with combined attacks)). The irregular level of scaling and rotation attacks, the non-availability of post processing attacks, large size of images, the large variance in the size of the copied portion etc are some of

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 12 Issue 04, April-2023

the demerits of this dataset. And all these demerits have led the researchers not to select this dataset for the evaluation and comparison of their algorithms. This dataset is downloadable at : http://dx.doi.org/10. 6084/m9.figshare.978736.

2.1.6 Copy Move Forgery Dataset

The dataset for the evaluation of their own algorithm, by Ardizzone et al (2015), contains a total of 1020 medium sized images, which is also its advantage. The dataset consists of 4 subsets D0 (50 plain images with simple 1 to 1 copy move), D1(600 images with 30 levels of rotation with step of 5, 30 and 1),D2 (320 images in 16 levels with scaling with a step of 0.25 and 0.05), D4(50 original images). The regular levels at which the attacks were created for each image also makes it convenient for the detailed analysis of its efficiency. The 30 levels within a range of -5 to 5 degree of rotation accounts for redundancy. The experiments by the author reveals that there is no clear difference in the precision and recall values with a rotation ranges of -5 to 5 degrees. The non-occurrence of post processing attacks and combination attacks are the demerits of this dataset other than the large size difference in the copy move regions. Manu et al. (2016) used this dataset in their research work for the evaluation. This dataset is available at http://www.diid.unipa.it/cvip/?page_id=48#CMFD.

2.1.7 COVERAGE

The COpy-moVe ForgERy dAtabase with similar but Genuine objEcts (COVERAGE), was developed by Wen et al.(2016) to comeover demerits of prevailing datasets. It consisted of convincing and perplexing forgery with several similar but candid objects. It was developed keeping in mind the six intermediate attacks, but it lacks any additional information about the scaling factor or the rotation angle used for the generation of the dataset. This dataset consists of 100 images with 16 translated images, 16 scaled images, 16 rotated images, 16 distorted images, 16 images with lighting effect modification and 20 images with combination attacks.

The considerably small dataset is one of the demerits accompanied with no trace of post processing attacks. Missing details of factors upon which scaling and rotating is done also possesses the challenge. The larger copy move region size in the images makes it less puzzling. The ground truth is shown separately for the duplicated and pasted region in this dataset that gives an extra overhead to the researchers to combine them prior to calculating the accuracy. This dataset is available at https://github.com/wenbihan/coverage.

2.1.8 CASIA V1 and V2

Dong et al (2013) introduced a new database for CMFD names as CASIA Image tampering Detection Evaluation Database. The dataset consisted for color images developed using Adobe Photoshop. In CASIA Version 1.0 named as CASIA V1, there are 1721 images (800 authentic and 921 tampered) considering only splicing as the one and only tampering and the image size is fixed at 381 x 256 in JPEG format.

CASIA Version 2.0 named as CASIA V2 contains 12,323 images (7200 authentic images and 5123 tampered images) with more comprehensive images sizes and formats. This dataset was developed from Corel dataset (9), public websites and captured. It consisted of compressed and uncompressed images collected under 9 different categories. Considering its advantages, it is one of the widely selected datasets for the evaluation of CMFD algorithms. CASIA dataset is available at: http://forensics.idealtest.org/.

2.1.9 Other datasets

Columbia gray datasets (Ng et al., 2004) included 1845 bitmap images with 128 pixels x 128 pixels fixed sized images. The extracted images are from CalPhotos collection, some extra images captures using professional cameras. The dataset contains equal number of original and tampered images categorized into different subsets as textured, arbitrary object boundary, smooth, straight boundary.

Columbia color dataset (Hsu et al., 2006) consisted of 363 images in the size range of 757 x 568 – 1152 x 768 pixels. The dataset comprises of equal number of uncompressed real images and forged images in TIFF or BMP formats. The tampering is done in a simple manner with no post processing operations. The images are shot indoor and on cloudy days to include the outdoor illumination identical to inside conditions. Columbia and CASIA datasets were used for comparison of their datasets by Islam et al(2020). Dataset is readily accessible at https://www.ee.columbia.edu/~dvmmweb/dvmm/downloads/PIM_PRCG_dataset/dataset-download.htm.

Federation – Berkeley Deep Drive Forgery (FBDDF) Dataset (Islam et al, 2020) was developed from the existing Berkeley Deep Drive (BDD) Dataset (Yu et al, 2018), which is an open source driving dataset of videos and images captured from those videos including different times of a day, weather and road conditions. It is mainly used in Internet of Things (IoT) based researches. The FBDDF included 200 real images from BDD database with another 200 forged images created using Photoshop. It included JPEG compressed and uncompressed TIFF images. It is an online available dataset. Dataset is available online at https://bit.ly/2ZrKe8Q.

2.1 Factors that characterizes the dataset

From the understanding of the various datasets available for copy move forgery detection, the researchers should consider certain details like creation of realistic forgeries with numerous types and levels of attacks, the contents of the image, the selection of copy-move regions. Comparing the datasets, they have some demerits that prevent the researchers for using these datasets and instead to create and use their own. Most of the researchers in this field of study have used only considerable subset of the datasets instead of the entire dataset in their work.

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 12 Issue 04, April-2023

2.1.1 Dimension of the Image

Image size has an important role as there exist a direct relationship between the image size and the computation time. Large images are included in the datasets to consider the real life images captured through cameras and the small sized images to decrease the computational time. Studies have shown that most of the experiments used the small images to reduce the processing time. Images shared through social media sites are downscaled to limit the server storage and taking into consideration the mobile phone users. Millions of images are shared daily though internet which make it difficult for the storage and transmission of big size images. Social media websites itself provide some guidance to the ideal size of the images to be included. These sites such as Facebook, Linkedin, Pinterest, Google+ etc., limits the average image size from a larger 757 x 474 to smaller 249 x 137 and allows the users to upload the images of dimensions upto a limited file size. Many types of forgeries are prevailing in the blogging or social media websites and the forged images also pave the way for mass opinion. Studies show that the ideal size of the image would be nearly 1024 x768 pixels, that is most prevalent in the internet and that does not compromise on the computational complexity. Moreover, the datasets should include realistic and high resolution image as a subset to meet the research requirements.

2.1.2 Types and levels of copy-move attacks

As the dataset is a sample of the real life images, it should contain all possible intermediate and post-processing attacks. The difficulty in building such datasets have made the developers compromise on the complex transformations and give priority to common attacks only. According to the experiments conducted the datasets b and c are characterized by several types of attacks. But they also ignored some of the common attacks and only some combinations were taken into account. A good dataset should contain all possible attacks including the combination attacks to evaluate the efficiency of an algorithm against every type of attacks.

Some of the datasets discussed in this article uses certain levels of attacks for evaluation. The controlled finite levels of transformations help in the in-depth analysis of an algorithm. But studies have shown that in some datasets the levels of transformations is leading to redundancy as we say in Copy Move Forgery Dataset and CMFDdb_grip dataset. Limiting the number of levels of a particular transformation below 10 can help in reducing the redundancy.

2.1.3 Copy-move region size

The accuracy of detection is directly correlated to the extent of the forged region in this type of forgery and hence quantity and quality of features extracted from the region should be considered when determining the copy-move region size. Although researchers who developed these datasets created images taking into consideration quality factors like image size, types and levels of attacks, none of them have given due importance to the copy-move region size. Almost all the algorithms is using a threshold value, or the block size as the parameter, which gives the researcher difficulty in setting up the optimum threshold for varying copy-move regions sizes. Analysis shows that the prevailing datasets include images with smaller copy-move regions to larger copy move regions. Hence it would have been good, if the dataset would have classification of the subsets according to the copy-move region sizes also which will enable the researchers to set the apt threshold values and parameters. Further studies can also be conducted on the dependency between the size of the copied region and the performance of an algorithm.

2.1.4 Realistic forgery

Most of the dataset developers claim their datasets to be realistic which in turn depends on many parameters like the image size, the copied region size, the transformation level and type done on the images, the actions applied on the contours, the contents of the copied region etc. The main factor which needs to be considered for a realistic forgery is the contents of the entire image rather than the contents of the copied region. If the images contents have natural similarities, then the image will look more realistic for the human and the machine finds it difficult to detect. If the copy pated regions are unique when equaled with the rest of the image, then it will be unrealistic in the sense that a human naked eye or the machine will find it easy to identify the forgery.

2.1.5 Number of Images

The size of dataset determined by the number of images in it depends on many factors like the variety of image sizes, the types and levels of transformations, size of the copied regions. Some of the datasets includes only smaller number of images while some other include a larger number of images in thousands. As the computation time can be decreased, the researchers will tend to choose a dataset with lesser number of images. But these smaller datasets will fail to give all the possible combinations or types of attacks needed for the evaluation of the algorithms. The larger datasets can also be not used entirely without compromising the computational complexity and studies show that the researchers will chose a subset of these larger dataset for their evaluation purposes. As a result, an ideal dataset would need to have 1000 to 1200 images that satisfy all the requirements for the comparative evaluation of an algorithm.

3. CONCLUSION

An analytical survey is done on the prevailing benchmark datasets for CMFD algorithms. Seven datasets are compared taking into account the merits and demerits of each. Some of the datasets are found much efficient for block based methods while some are found effective for block-based algorithms. The datasets evaluated varies widely with respect to the number of images, the image size, the types and levels of attacks and the copy move region size. The databases were created to be more realistic than serves the purpose for evaluation. Some of the datasets are extensively used for its advantages, while some datasets are seldom used for its disadvantages. This necessitates the need for an ideal and reliable dataset that includes all possible combinations of attacks, which can be used for the evaluation and comparison of various algorithms used for CMFD.

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 12 Issue 04, April-2023

Some practical suggestions are put forward in this article which can lead to the development of an ideal and realistic dataset for CMFD algorithms.

4. REFERENCES

[1] Al-Qershi OM, Khoo BE (2013) 'Passive detection of copy-move forgery in digital images: State-of the-art', Forensic Science International, Vol.231(1–3), pp. 284–295
[2] Amerini I, Ballan L, Caldelli R, Del Bimbo A, Del Tongo L, Serra G (2013) 'Copy-move forgery detection and localization by means of robust clustering with J-linkage'. Signal Processing & Image Communications Vol.28(6), pp.659–669
[3] 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 Transactions on Information Forensics Security, Vol. 6(3), pp.1099–1110
[4] Ardizzone E, Bruno A, Mazzola G (2015) 'Copy-move forgery detection by matching triangles of Keypoints'. IEEE Transactions on Information Forensics Security Vol.10(10), pp.2084–2094
[5] Bakiah N, Wahid A, Yamani M et al (2016) Copy-move forgery detection : Survey, challenges and future directions', Journal of Networking Computational Application , Vol.75, pp.259–278
[6] Birajdar GK, Mankar VH (2013) 'Digital image forgery detection using passive techniques: a survey', Digital Investigations Vol. 10, pp.226–245
[7] Chihaoui T, Bourouis S, Hamrouni K (2014) 'Copy move image forgery detection based on SIFT descriptors and SVD matching'. Proceedings of the International Conference on Advanced Technologies for Signal and Image Processing, pp. 125–129
[8] Christlein V, Riess C, Jordan J, Angelopoulou E (2012) 'An evaluation of popular copy-move forgery detection approaches'. IEEE Transactions on Information Forensics Security ,Vol.7(6), pp.1841–1854
[9] Corel Database, "http://corel.digitalriver.com/,"
[10] Cozzolino D, Poggi G, Verdoliva L (2014) 'Copy-move forgery detection based on patchmatch'. The International Conference on Image Processing (ICIP), pp 5247–5251
[11] Cozzolino D, Poggi G, Verdoliva L (2015) 'Efficient dense-field copy – move forgery detection', IEEE Transactions on Information Forensics Security Vol.10(11), pp.2284–2297
[12] Dong, J., Wang, W., & Tan, T. (2013). 'CASIA Image Tampering Detection Evaluation Database', 2013 IEEE China Summit and International Conference on Signal and Information Processing. doi:10.1109/chinasip.2013.6625374
[13] Emam M, Han Q, Niu X (2015)' PCET based copy-move forgery detection in images under geometric transforms'. Multimedia Tools Applications, Vol.75(18), pp.11513–11527
[14] Fridrich J, Soukal D, Lukáš J (2003) 'Detection of copy-move forgery in digital images', Proceedings of Digital Forensic Research Workshop 2003, Cleveland, pp. 19–23
[15] Hashmi MF, Hambarde AR, Keskar AG (2014) 'Copy move forgery detectionusing DWT and SIFT features'. Proceedings of the International Conference on Intelligent Systems Design and Applications, ISDA, pp.188–193
[16] HashmiMF, AnandV, KeskarAG(2014) 'Copy-move image forgery detection using an efficient and robust method combining un-decimated wavelet transform and scale invariant feature transform'. AASRI Procedia Vol. 9, pp. 84–91
[17] He H, Huang X, Kuang J (2013) 'Exposing copy-move forgeries based on a dimension-reduced SIFT method'. Information Technologies Journal, Vol. 12(14), pp.2975–2979
[18] Hsu H-C, Wang M-S (2012) 'Detection of copy-move forgery image using Gabor descriptor', Anti counterfeiting, security and identification (ASID), 2012 international conference, pp.1–4
[19] Hsu, Y.F.; Chang, S.F(2006)' Detecting image splicing using geometry invariants and camera characteristics consistency'. Proceedings of the 2006 IEEE International Conference on Multimedia and Expo, Toronto, ON, Canada, 9–12 July 2006, pp. 549–552
[20] Islam, M.M.; Karmakar, G.; Kamruzzaman, J.; Murshed, M(2020) 'A Robust Forgery Detection Method for Copy–Move and Splicing Attacks in Images'. *Electronics* 2020, *9*, 1500. https://doi.org/10.3390/electronics9091500
[21] J. Lee, K. Chang, C. Chang, and C. Cheng (2014) 'Image copy-move forgery detection based on HOG'. Proceedings of the 27th IPPR conference on computer vision, graphics, and image processing', pp 1–5
[22] Kulkarni VS, Chavan YV (2014) 'Comparison of methods for detection of copy-move forgery in digital images'. Spvryan's International Journal of Engineering Science and Technology Vol.1(1), pp.1–6
[23] Kumar S, Desai J,Mukherjee S (2013) 'A fast DCT based method for copy move forgery detection', IEEE Second International Conference on Image Information Processing, pp 649–654
[24] Liu F, Feng H (2014) 'An efficient algorithm for image copy-move forgery detection based on DWT and SVD'. International Journal of Security and its Applications, Vol.8(5), pp.377–390
[25] Liu L, Ni R, Zhao Y, Li S (2014) 'Improved SIFT-based copy-move detection using BFSN clustering and CFA features', Proceedings of the 10th International Conference on Intelligent Information Hiding and Multimedia Signal Processing; pp. 626–629
[26] Mahdian B, Saic S (2010) 'A bibliography on blind methods for identifying image forgery', Signal Process Image Communications, Vol. 25(6), pp. 389–399
[27] Maind RA, Khade A, Chitre DK (2014) 'Image copy move forgery detection using block representing method'. International Journal of Soft Computing Engineering Vol. 4(2), pp.49–53
[28] Manu VT, Mehtre BM (2016) 'Detection of Copy-Move Forgery in Images Using Segmentation and SURF', Proceedings of the 2nd International Symposium on Signal Processing and Intelligent Recognition Systems (SIRS), Vol.425, pp.645–654
[29] Ng, T.T.; Chang, S.F.(2004) 'A model for image splicing'. In Proceedings of the 2004 International Conference on Image Processing, ICIP'04, Singapore, 24–27 October 2004; Vol 2, pp. 1169–1172
[30] Pun C, Member S, Yuan X, Bi X (2015) 'Image forgery detection using adaptive over segmentation and feature point matching', IEEE Transactions on Information Forensics Security , Vol.10(8),pp.1–12
[31] Redi JA, Taktak W, Dugelay JL (2011) 'Digital image forensics: a booklet for beginners', Multimedia Tools Applications Vol. 51(1), pp.133–162
[32] Ryu S-J, Kirchner M, Lee M-J, Lee H-K (2013) 'Rotation invariant localization of duplicated image regions based on zernike moments'. IEEE Transactions on Information Forensics Security, Vol. 8(8), pp.1355–1370
[33] S, Koul DA (2014) 'Passive image forensic method to detect copy move forgery in digital images'. IOSR Journal of Computational Enggineering, Vol.16(2), pp. 96–104
[34] Satapathy SC, Biswal BN, Udgata SK, Mandal JK (2014) 'Passive copy move forgery detection using SURF, HOG and SIFT Features'. Proceedings of the 3rd International Conference on Frontiers of Intelligent Computing: Theory and Applications (FICTA), Vol. 327, pp. 659–666
[35] Silva E, Carvalho T, Ferreira A, Rocha A (2015) 'Going deeper into copy-move forgery detection: Exploring Image telltales via multi-scale analysis and voting processes', Journal of Visual Communication & Image Representation, Vol.29, pp.16–32
[36] Sunil K, Jagan D, Shaktidev M(2014) 'DCT-PCA based method for copy-move forgery detection', Advances in Intelligent Systems and Computing Vol.249, pp.577–583
[37] Thajeel SA, Sulong G (2015) 'A Novel Approach for detection of copy move forgery using completed robust local binary pattern'. Journal of Information Hiding and Multimedia Signal Processing, Vol. 6(2), pp.351–364

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 12 Issue 04, April-2023

[38] Torralba A, Efros AA (2011) 'Unbiased look at dataset bias'. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 1521–1528
[39] Tralic D, Rosin PL, Sun X, Grgic S (2014) 'Detection of duplicated image regions using cellular automata', Proceedings of the International Conference on systems, signals and image processing (IWSSIP), pp. 167–170
[40] Tralic D, Zupancic I, Grgic S, Grgic M (2013) 'CoMoFoD - New database for copy-move forgery detection'. Proceedings of 55th International Symposium ELMAR, pp. 49–54
[41] Uliyan D, Jalab H, Abdul Wahab A, Sadeghi S (2016) 'Image region duplication forgery detection based on angular radial partitioning and Harris key-points'. Symmetry (Basel) Vol.8(7), pp. 62
[42] Uliyan DM, Jalab H, Abdul Wahab AW (2015) 'Copy move image forgery detection using hessian and center symmetric local binary pattern'. Proceedings of the IEEE Conference on Open Systems ICOS, pp. 7–11
[43] Ustubioglu B, Ulutas G, Ulutas M, Nabiyev V (2016) 'A new copy move forgery detection technique with automatic threshold determination'. AEU International Journal of Electronics Communications Vol.70(8), pp.1076–1087
[44] Wen B, Ye Z, Ng RST-T, Shen X, Winkler S (2016) 'COVERAGE – A novel database for copy-move forgery detection', Proceedings of the IEEE International Conference on Image Processing (ICIP), pp. 161–165
[45] Yu L, Han Q, Niu X (2014) 'Copy-rotation-move forgery detection using the MROGH descriptor'. Proceedings of the IEEE International conference on cloud engineering, IC2E, pp. 510–513
[46] Yu L, Han Q, Niu X (2016) 'Feature point-based copy-move forgery detection: covering the non-textured areas' Multimedia Tools Applications Vol.75(2),pp.1159–1176
[47] Yu, F.; Xian, W.; Chen, Y.; Liu, F.; Liao, M.; Madhavan, V.; Darrell, T. Bdd (2018) '100k: A diverse driving video database with scalable annotation tooling. arXiv 2018, arXiv:1805.04687.

AUTHORS PROFILE

Dhanya R, is a Research Scholar in the Department of Computer Applications at Noorul Islam Centre for Higher Education, Kumaracoil, Thuckalay, Tamil Nadu. She is currently working as an Assistant Professor in the Department of Computer Applications, St. Teresa's College, Ernakulam, Kerala. She is an MCA graduate with research interest in Digital Image Processing and Digital Electronics and with 14 years of teaching experience. She has presented and published papers in International Conferences and Journals

R. Kalaiselvi, is currently working as an Associate Professor in the Department of Computer Science and Engineering, RMK College Of Engineering And Technology, Puduvoyal, Gummidipoondi, TamilNadu, India. She has more than eighteen years of teaching experience. She obtained her B.E (Electronics and Communication Engineering) from M.S University, Tirunelveli, and her ME (Computer Science and Engineering) from Anna University, Chennai, India.and her Ph.D from Anna University Chennai. Her research interests include Network Security, Cloud Computing and Distributed Computing. She has published many research papers in National/International Conferences and Journals. She has attended several seminars and workshops in the past ten years. She is an IEEE member.