Forgery Detection in Documents

[1]Balaji V , [2]Ajith Kumar P, [3]Kiren Aananth A, [4]Gunasekar N, [5]Ciyamala Kushbu S
[1234]*UG STUDENT ECE , RMK College Of Engieering And Technology,Puduvoyal, India*
[5]*Assocoiate Professor ECE , RMK College Of Engineering And Technology,Puduvoyal, India*

***Abstract*:**
Paper is extremely important and has been a carrier of information for decades. Just as every coin has two sides, paper has been put to use for various malicious purposes. This has been possible by the use of technology as modification of documents are possible. All kinds of details from authority seal, signature and texting are being manipulated. As a result, both public and private companies are facing enormous losses in time and economic resources. Apart from forensic techniques, no other method has been successful in the detecting forgery when it comes to documents. There are various methods in which forgery can be performed and at the same tome there are various methods for detecting forgery. Various kinds of printers are used in manipulating documents namely inkjet printers and electrostatic copiers. These devices leave a distinct and specific signature and these are exploited. Apart from that signatures and seals from one document to another are being forged by means of copy move forgery. This problem is being handled by repeated segmentation and feature extraction techniques.

***Keywords*—**Forgery Detection, SURF, Document Forgery, Copy Move forgery detection.

I. INTRODUCTION

Digital imaging technology has been used in many different ways due to its increasing availability. Consequently, it has become possible to store all forms of data in digital format. Advancement in digital image technology has led to the manipulation of these kind of data which includes tampering the image, image retouching, image splicing, morphing and copy-move. Apart from signature forgery, textual forgery is also possible. Here the texture and alignment of various texts can play an important role in detection as forged texts cannot have same texture and alignment as that of the original document. By continuous segmentation and feature extraction these fake signatures can be detected. Documents of extreme importance which will include central government documents, wills, etc. Nowadays announcements from the central government are also made through authorized statement letters with the signature of the respective authority in order for swift message spread. The use of these announcements raised due to the fake information spread via various news channels.

UG STUDENT ECE
RMK College Of Engineering And Technology
Puduvoyal, India

II. CHARACTERISTICS OF THE FORGED DOCUMENTS

Here the documents that are forged will have distinct characteristics and features. When words are copied from one source to another there will be a lot of irregularities in layout, exposure, line width, roughness, noise etc. These irregularities can be spotted out or differentiated by various image processing steps. Similarly, when it comes to detection of fake signatures which is done via copy move forgery, block by block segmentation of the image is done in order to extract key features and compression of the image will also reveal a lot of differences. The major disadvantage with this procedure is it suits only images with JPEG compression while there are a lot of digital formats to store an image. Apart from fake signatures, logos, stamps and many other forms of pictorial documentation which are unique with that particular organization are also subjected to forgery. There are many cases where an original document can be used in forgery. The name of the particulars, the logos, the signatures, seals etc., can be forged into another document. In a wholescale forged document features like noise, line format, edge roughness will be completely different from the normal original document.

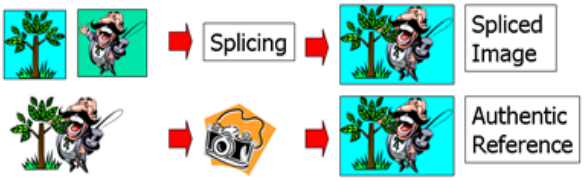
III. TYPES OF FORGERY IN TEXTS AND IMAGES

A. Image Retouching:

Image Retouching is less harmful than other types present. In case of image retouching original image does not significantly change. There is enhancement or reduction in certain features of original image. This is present in almost all-magazine covers so that it is more attractive. The same concept is used in facial enhancement and has also been used as applications in today's smartphones. These facial enhancements include basic "fixes", i.e. removal of pimple and smoothening a ruddy complexion.

B. *Image splicing or photomontage:*

Image splicing is fundamentally simple process and can be done as crops and pastes regions from the same or separate sources. In Image Splicing technique there is composition of two or more images, which are combined to create a fake image. Examples include several infamous news reporting cases involving the use of faked images.



C. *Copy-Move Attack:*

The copy move forgery is popular as one of the difficult and most commonly used kind of image tampering technique. One needs to cover a part of the image in order to add or remove information. The intention is to hide something in the original image with some other part of the same image.

IV. RELATED WORK

Francisco Cruz, Nicolas Sidere, Mickael Coustaty have come up with the idea of making use of the Local Binary Patterns of the image in order to detect irregularities, roughness, noise levels at each and every patch [1]. Their method starts with pre-processing the particular image by conversion of image from RGB to grey scale and the application of filters in order to remove the additional and unwanted noise. This is further followed by the extraction of the set of the patches using various features as extraction parameters. Based on the characteristics of the documents produced by different printing devices Shize Shang, Nasir Memon and Xiangwei Kong came up with the method of analyzing the document texture, noise around the edges, noise in the text part and the roughness in the text along with the line layout in order to detect the irregularities [2]. Here the text document undergoes segmentation and as a result there will be visible differentiation between text, edges and background. To handle forgery in images and pictures Resmi M.R and Vishnukumar S came up with a segmentation-based detection algorithm [3]. Here two kinds of detection where analyzed in this paper. They were Key point detection where key features of the particular document are being tracked using double thresholding methodology after subjecting the image to various edge detection techniques. The other method is the block-based method where the 'conquer and divide' methodology is used in order to find clusters. S. Murali and Basavaraj introduced JPEG compression techniques and filtering methods to detect image tampering by realizing the possible tampering locations [4]. JPEG compression techniques have a major disadvantage where only certain types image formats can be supported. Snigdha K. Mankar and Prof. Dr. Ajay A. Gurjar have given a complete analysis of the various available forgery detection [5]. Mohammed N. Nazli came up with various existing forgery detection algorithm which involves the use over segmentation and in some cases, it also involves the use of DCT (Discrete Cosine Transform) [6].

V. PROPOSED ALGORITHM

Working

Various forgery detection algorithms and feature extraction procedures have been taken into account in this project. These feature detection procedures all have a common principle and working principle but there will be a lot of differences in the working intensity. The SURF (Speeded Up Robust Features) methodology is chosen for our project. This methodology is a derivative of the SIFT (Scale Invariant Feature Transform). SURF is several times faster than SIFT and claimed by its authors to be more robust against different image transformations than SIFT. The main reason behind using the SURF algorithm apart from its advantages over SIFT is due to ease of access and operating in hardware as well as software. In this algorithm there will be a database that stores a series of formats of the original documents. In the first step there will be a storage of formats of certain original documents. These documents will serve as a point of reference. Now in the first step there will be an input of the testing document particulars into the blanks of the original format. After that the document is further resized in a common threshold scale for easy evaluation. In the process an original version or exact copy of the original document will be created. This simulation result will be used for comparison of extracted

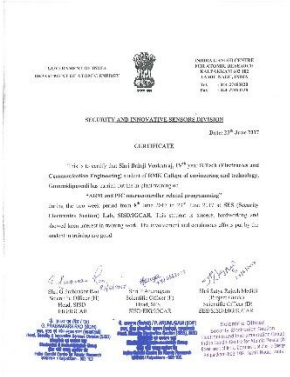
features. As soon as the features are extracted they are made to overlap on the original document itself which gives us a result regarding the position of the features. In the feature extraction process SURF (Speed-up Robust Features) methodology is used. Features like edges, corners, blobs, text roughness difference in shade between the text and the surrounding content. These features are highlighted by tiny circles. This step is followed by overlapping the feature labelled document over the same original document to give us an overall picture on how the result would be if an original document is the input. Now the testing document is taken in as the input and resized as the same scale as that of the original document. This same feature extraction technique is used to extract the features of the input document. These features are then labelled within the document. Now an overlapping of the two documents takes place. Here the features that are exact will match and the features that are not matching will be labelled with a different color. Furthermore, the part of the document which is forged will be extracted and displayed separately.

SURF (Speeded-Up Robust Feature Extraction) Algorithm and its advantages

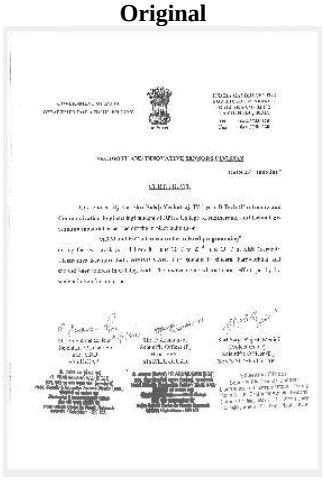
There has been a major conflict of interest on determining the efficiency between two feature extraction methodologies i.e., SIFT and SURF. The SURF algorithm was partly inspired/derived from the SIFT (Scale Invariant Feature Transform) algorithm. The SURF algorithm is similar to the SIFT algorithm in functions performed. They are both used in image registration, object recognition, 3D reconstruction, robotic mapping, navigation, image stitching, gesture recognition, video tracking, classification, etc. Many authors claim that SURF is advanced, faster and more robust than SIFT but researchers claim that SIFT may be a little advantageous. On the whole, via repeated experiments and trials it is found out that SURF usage is better and easy whereas implementation of SIFT is hard due its tedious algorithmic setup. This has been tested in various simulations and it is found to be simple, efficient and easy to understand. This local feature detector and descriptor was patented in 2006 in America by Herbert Bay. Like other feature detection algorithms, interest points of a given image are defined as salient features from a scale-invariant representation. For detection of these interest feature points such as edges, corners, noises etc. the SURF algorithm uses an integer approximation of the determinant of Hessian blob detector. In the case of SIFT objects are recognized in a new image by feature comparison of the new image with the database and finding candidate matching features based on Euclidean distance of their feature vectors. After evaluation of the entire set of matches, subsets of key points that agree on the object and its orientation, location, scale etc., in the new image are identified to filter out good matches. Calculation of Euclidian distance between the feature points are a tedious and time-consuming process. Implementation of the equation in real time systems and processors will be time consuming too. Whereas, many systems and simulation libraries along with processor directories includes SURF algorithm files making it easy and convenient. Initially Gaussian Smoothing is applied to the base image (integral image). It is faster than the normal cascade filters. Interest points are normally found at various scales. Search of correspondence often requires comparison images that are in different scales.

VI. RESULT AND ANALYSIS

The result is obtained based on a forged image rather than a signature for better explanation. Here the image is first preprocessed. Then the image is converted to greyscale.

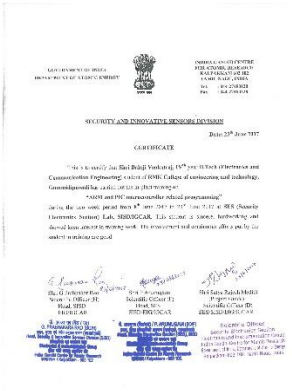


Original

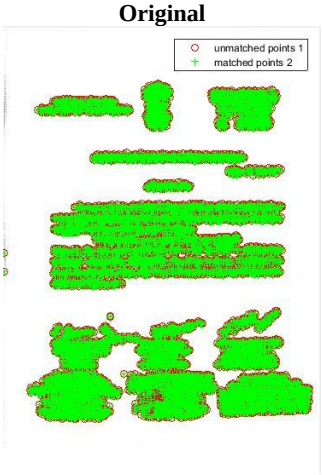


Gray Scale Image

Now the features are extracted and merged with the feature labelled image. If the input document is an original document then the output image will be like the one given below.

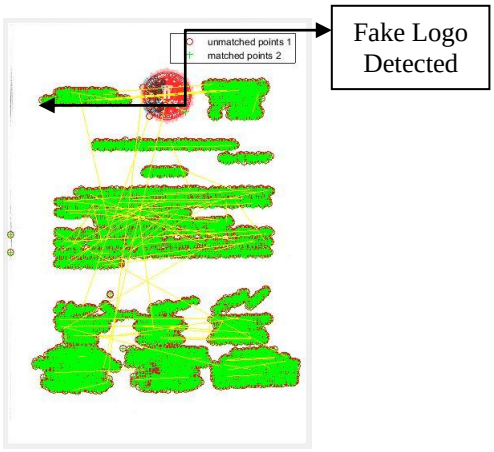


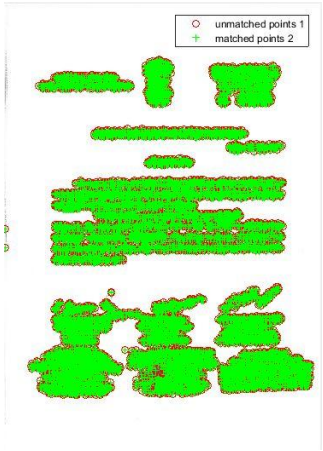
Original



Final Output Of The Original Image

If there is a forgery in the logo then the output turns out like the one given below where the logo part will be in red.





Similarly, if there is a copy move forgery in the text part then the forged part will be highlighted in yellow.

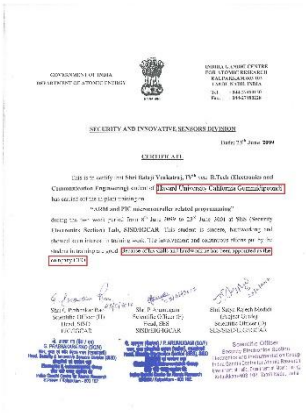


Fig 1

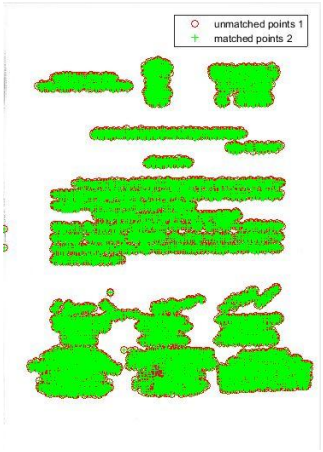


Fig 2

The highlighted part is fake in the fig 1 document. Fig 2 is the expected original document.

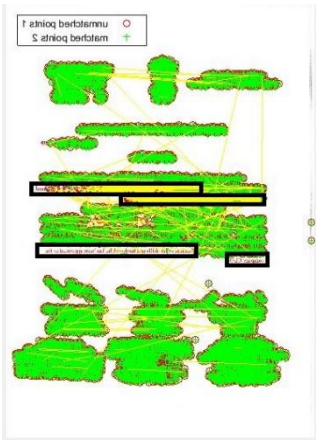


Fig 3

The highlighted text in fig 3 are the forged context of the document.

Another example of forgery in text is also given below.

Fig 4

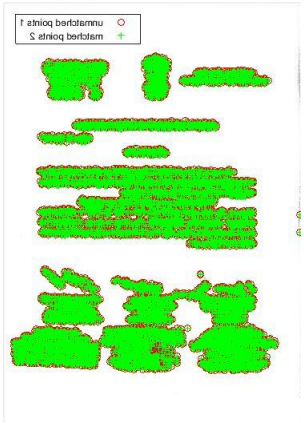


Fig 5

The highlighted part in fig 4 is the forged part. Fig 5 is the expected output for an original document.

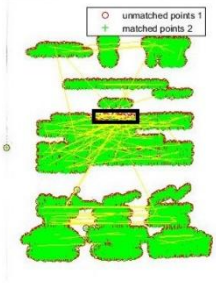


Fig 6

The highlighted part in fig 6 is the detected forged context.

In the end we were able to execute the algorithm on a full scale forged document and the results were successful.

VII. CONCLUSION

Full-scale document forgery is achieved by means segregating the textual and non-textual part of the document and performing forgery detection on each part separately. By doing so a lot of accuracy is possible especially while detecting forgery in texts.

REFERENCES

[1] Local binary patterns for document forgery detection. Francisco Cruz, Nicolas Sidere, Mickael Coustaty 2017 14th IAPR International Conference on Document Analysis and Recognition.

[2] Mehak, Tarun Gulati,"Detection of Digital Forgery Image using Different Techniques"International Journal of Engineering Trends and Technology (IJETT),Volume-46 Number-8,2017.

[3] Detecting documents forged by printing and Copying. Shize Shang, Nasir Memon and Xiangwei Kong *EURASIP* Journal on Advances in Signal Processing 2014, 2014:140
Dr.R.Surendiran,"Secure Software Framework for Process Improvement",International Journal of Computer Science and Engineerin g (SSRG-IJCSE),volume 3 Issue 12 2016,ISSN: 2348 – 8387,Page 19 - 25.

[4] A Novel Segmentation Based Copy-Move Forgery Detection in Digital Images. Resmi M.R. Vishnukumar S.

2017 International Conference on Networks & Advances in Computational Technologies (NetACT) |20-22 July 2017| Trivandrum

[5] R.Surendiran,Dr.K.Alagarsamy,"An Extensive Survey on Mobile Security and Issues",International Journal of Computer & Organization Trends(IJCOT) – Volume2 Issue1 2012,ISSN: 2249 - 2593, Page 39 - 46.

[6] Detection of Digital Photo Image Forgery. S.Murali, Basavaraj S. Anami, Govindraj B. Chittapur. 2012 IEEE International Conference on Advanced Communication Control and Computing Technologies (ICACCCT)

[7] Image Forgery Types and Their Detection: A Review. Snigdha K. Mankar, Prof. Dr. Ajay A. Gurjar. Volume 5, Issue 4, April 2015 ,ISSN: 2277 128X , International Journal of Advanced Research in Computer Science and Software Engineering

[8] COMPARISON BETWEEN IMAGE FORGERY DETECTION ALGORITHMS. Mohammed N. Nazli, Ashraf Y. A. Maghari 2017 8th International Conference on Information Technology (ICIT)

[9] Passive forensics for copy-move image forgery using a method based on DCT and SVD. Jie Zhao, Jichang Guo School of Electronic Information Engineering, Tianjin University, Tianjin 300072, China b School of Computer and Information Engineering, Tianjin Chengjian University, Tianjin 300384, China

[10] S.L.Jothilakshmi, K Valli, A.Vanitha, A. Selva nathiya, A.Shunmuga priya," Automatic Machine Learning Forgery Detection Based On Svm Classifier",International Journal of Mechanical Engineering (SSRG-IJME),Volume 1 Issue1- 2014.

[11] https://goo.gl/images/wbUAai example for image retouching

[12] http://www.ee.columbia.edu/ln/dvmm/trustfoto/projs/splicing/splicing.html for image splicing

[13] .https://www.sciencedirect.com/science/article/pii/S0045790613003169 for copy move forgery.

[14] Image Forgery Detection Using Adaptive Over-Segmentation and Feature Point Matching

[15] Chi-Man Pun, *Senior Member, IEEE,* Xiao-Chen Yuan, *Member, IEEE,* and Xiu-Li Bi Haritha Damarla,"Research Methodology on Offline and Online Signature Verification and Forgery Detection",International Journal of Computer Science and Engineering (SSRG-IJCSE),Volume-4 Issue-11 2017.

[16] Accurate and Efficient Image Forgery Detection Using Lateral Chromatic Aberration Owen Mayer, *Student Member, IEEE,* Matthew C. Stamm, *Member, IEEE* Dr.S.Kannan, Mr.T.Pushparaj,"An Analysis of Software Testing Security using Quality Assurance and Reliability",International Journal of Computer & Organization Trends (IJCOT),Volume - 7 Issue - 6 2017.

[17] A New Block-based Copy-Move Forgery Detection Method in Digital Images Hajar Moradi-Gharghani and Mehdi Nasri International Conference on Communication and Signal Processing, April 6-8, 2016, India

[18] Survey of Copy-Paste Forgery Detection in Digital Image Forensic. Anushree U. Tembe, Priya S. Thombre International Conference on Innovative Mechanisms for Industry Applications (ICIMIA 2017).