Journal of Basrah Researches ((Sciences)) Vol. (47). No. 1 (2021)

Available online at: https://www.iasj.net/iasj/journal/260/issues

ISSN –1817 –2695

Received date: 30-7-2020 Accepted date: 1-11-2020 Available online date: 31-12-2021

A new approach for finding duplicated words in scanned Arabic documents based on OCR and SURF

Asmaa Alsimry [1*], Khawla Hussein Ali [2] and Enas Wahab Abood [3]

[1] Information Technology Department, Financial Devision, Basrah Oil Company, Basrah, Iraq,*

[2]Computer Science Department, Education College for pure Science, University of Basrah, Basrah, Iraq, [3]College of Science, University of Basrah, Basrah, Iraq.

*E-mail:*asmamahdi44@gmail.com khawla.ali@uobasrah.edu.iq and enas.abood@uobasrah.edu.iq

Abstract

Copy move forgery detection (CMFD) is an important challenge nowadays especially with increasing electronic dealings for documents (e.g., images, pays lists, contracts, invoices, or any other information documents) and its availability in clouds exposes it to the danger of hackers and deceivers. Thus, many methods of detection images forgeries were presented to detect the forgery in the natural scenes and a few had scanned text documents. In this paper, an approach is presented to find duplicated words in Arabic scanned text documents as part of a system to detect forgery in official documents based on Optical character recognition (OCR) and Speeded Up Robust Feature (SURF). OCR was used for the purpose of segmenting the word out of the document while SURF was used to compute the feature descriptors of each segmented word for later matching which has two stages, the first stage is using the Euclidean equation for the purpose of matching features, followed by computing the ratio of the number of features corresponding to the reference word features. The two words are considered to be duplicated if the percentage increase or equal to a specific threshold (30%). The experimental results showed that the proposed approach was efficient and robust in finding duplicated words as it implemented for (image size <=90x90 pixel) the evaluation metrics were: TPR =50.54%, FPR=66.66%, ACC=25.436% , Running time=1.25 sec, And for (image size >90x90 pixel) TPR

=100%, FPR=60%, ACC=50.2% , Running time=1.87 sec. for SURF , while (image size <=90x90 pixel) TPR =95.3%, FPR=0%, ACC=97.65%, Running time=0.15 sec, and (image size >90x90 pixel) TPR =100%, FPR=0%, ACC=100%, Running time=0.18 sec.

Keywords: *Duplicated Word Detection, Copy-Move Forgery Detection, Scanned Documents, SURF, Forgery Localization, Arabic OCR.*

Introduction

The increasing of digital image editing software has made the images forgery as a popular and widely used tampering method in digital images. There are too many types of forgery; one of them is tampering images by deleting, adding objects, changing background, scaling, rotation, or illumination change. Others may aim to duplicate one or more objects within the same image or move it to another [1]. The last type is the widest, it is called copy-move forgery and aimed to cause misleading or misinterpretation [2]. Copy Move Forgery Detection (CMFD) is considered as a major challenge in digital image forensics, so many methods appeared, most of them subspecialize in images of natural scenes and few others took care of images of scanned text documents [1]. The finding of duplicated words considered as a base stone in CMFD systems for scanned documents and it faced a difficulties in detection because of the documents features have some issues such as it is consists of multiple similar looking glyphs, may have common structures or a background with little noise [2], So that effected to fail of detection forgeries in scanned text document, in other words because of the high similarity, it is a weak matching methods [1]. The researchers tried to produce and enhance a schemes and algorithms for dealing with duplicated texts [3], [4]. Some researchers used neural networks like Optical Character Recognition (OCR) [5], and others used the deep learning like Convolutional Neural Network (CNN) [6] and Artificial Neural Network (ANN) for segmenting word out of the documents [7]. The features extraction is the second important key in CMFD system that calculated the points of interest to represent each word numerically to find matched words. Many feature estimators were presented like Extract Max Frequency Intensity feature (EMFI) [8], HU moments and Zernike [5]. Speeded Up Robust Feature (SURF) the commonly and perfectly one for detection and extraction features from a natural scene images due to its ability to deal with oriented and scaled images [9], so it can be applicable with text scanned image because in case of copy move forgery the copies may exposed to distortion like scaling and orientation [10].

In this study the OCR were invested for segmenting step while SURF for feature extraction, and the Euclidian distance to find matching words. This paper build as follows: Section 2: Related works, section 3: Text-CMFD frame work, section 4: (OCR) as an supplemental tool that let us to recognize text in an image, section 5: SURF key points are employed to extract features of the possible duplicate regions in the image, section 6: Proposed matching task to finding correspondences between two images of the same scene, Section 7: Experiments and results and finally Section 8: The conclusion.

The proposed approach contributes effectively to improve the systems of detecting forgery in the official text documents of governmental or private institutions with inexpensive tools that do not require an expert compared to other usual methods that require the use of special and electronic expertise to detect forgery and may sometimes require the presence of watermark, signatures, etc. The fraud detection process avoids many financial and economic losses and administrative problems.

2: Related works

In recent years, many methods have been proposed to detect a CMFD, most of them worked on image of natural scene while some others took the forgery in scanned text documents as a matter of interest. The duplicated word is the major component in CMFD and all these systems works to develop this step as well as forgery step.

In 2008, Khanna *et. al*. [11] introduced methods to authenticate and identify tampering regions in images that have been obtained using scanners. The methods are based on applying statistical features as a fingerprint of imaging sensor pattern noise for the scanner. A SVM classifier is trained for applying statistical features of pattern noise to classify smaller blocks of an image. This feature vector is shown to identify the tampering regions with high accuracy. N. Khanna and E. J. Delp (2010) [12], introduced many extensive analysis of font shape and size on the proposed texture analysis based intrinsic signatures for scanned documents. On the other hand Prabhu, Anirudh *et. al.* (2012) [13] dealt with the documents which are clipping to damage and rigging. Many of these documents may turn out to be too risky. Their paper introduced three basic methods, the three methods as follows: 1. Border pixels 2. Block comparison 3. Pixel count. The technique used for this aim is a very easy pixel detection procedure.

Y Sun *et al.* (2013) [14], proposed a near duplicate text detection framework using signatures to save space and query time. They also proposed a new signature selection algorithm that uses the q-grams grouping frequency. They compared their algorithm with Winnowing, which is one of the modern signature selection algorithms. The result showed that their algorithm gain much better accuracy with less cost of time and space .using signatures to save space and query time. TMD Wu *et al.* (2013) [15], The duplicate text recognition system includes a primary method of dividing electronic text into multiple phrase sections; A second method for converting each segment of the phrase into a unique, fixed-length bit string; A third method of storing a set of bit strings, each group of bit strings (a set of strings) including a set of bit strings corresponds respectively to the phrase segments in a given electronic text; The fourth method is to determine whether the predetermined similarity between any two sets of strings in the third method reaches the first term, and to determine the two electronic texts corresponding to the two sets of strings are duplicate texts if the predetermined similarity between the two sets of strings reaches the first threshold. Hasan FHN. (2015) [8], was talking about a new method has two main stages to detect the scanned document forgery. The first stage consists of four operations : three operations using to preprocess the scanned image while the fourth operation was using to Extract Max Frequency Intensity feature (EMFI) from the pixels of document. Stage two was included one operation to Extract Edge Gradient (EEG) for finding the variance between the original text and the fabricated text. Abramova S. (2016) [16], showed how is the block-based detection of near-duplicates executed in a scanned documents. Steps of experiments showed modest performance in CMFD from scanned documents. Many strategies developed to further adapt block-based copy–move forgery detection to this relevant application and the results were TPR= 84.0%, FPR= 4.7 and ACC= 89.7. Carrazoni Entenza P.(2019) [5], discussed the CMFD problem in digital images that include scanned text documents (for example, payment lists, contracts, invoices, or any other information documents) by using two types of real-time methods, such as Zernike and Hu to extract features and discover if words are duplicates or not Based on the limit of similarity. He used a unique limit for each moment used in the paper for the same database and his experimental results showed that the Zernike method performed better than Hu's method in general. K Sun *et al.*(2019) [17], proposed a simple yet effective method to blindly reveal such tampering document image. The differential anomaly is often detected in the spatial domain along the boundary of the image assuming it without or with poor post-processing. In order to

obtain such legal evidence, the differences between rows and columns are calculated first order in the image and then their binaries across the threshold. These bidirectional team maps are filtered and combined to counteract background noise. A quadrangular structure consisting of four boundary lines specially latent via adaptive sliding windows is sought after within the combined map. Preliminary test results on the assembled image data set perfectly verify the effectiveness and efficiency of the tamper-evident detection method for the proposed document image with ACC= 88.3% , TPR= 87.0% and FPR= 18.5%.

Our proposed Approach was mainly aimed:

- To exploit the characteristics of SURF in dealing with distortions of Scale and orientation, which are among the most prominent side effects to the copy move in the scanned documents and to test their effectiveness in finding and matching its features as in the case of natural scene images
- speed up the match finding process and increase the accuracy rate in finding matches to reach 100% in finding duplicated words within the document in most of the documents under consideration and within an acceptable and relatively short period of time.

The drawbacks of our proposed approach were:

- Consider only printed and scanned documents. Because almost official documents are printed documents.
- The difficulty of collecting official documents to build the Arabic database due to the security of government institutions.
- Using a high resolution regions of not less than 300 x 300 pixels, to get clear details of the word to work properly with the SURF algorithm.

3: Text-CMFD frame work:

Generally, CMFD techniques work by a common framework and the researchers are adding or modifying some steps to ensure the effectiveness of their frames to get more perfect results in CMF[18].

The proposed framework in our study is illustrated with follows:

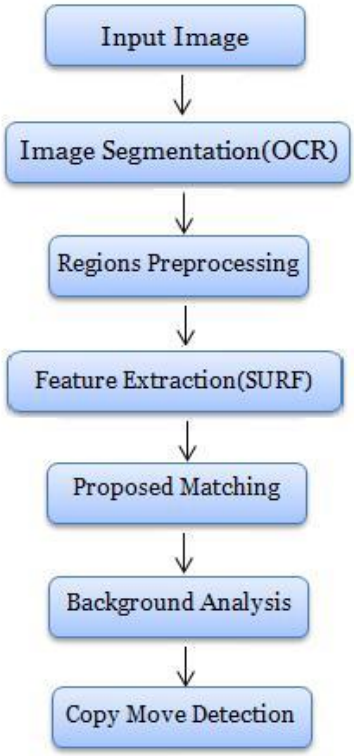


Figure1: Frame work of Text-CMFD

1. Image segmentation step was implemented with OCR to segment the scanned document image into regions (sub-images).
2. A preprocessing step was resizing the regions because OCR gives us varying size images and that may lead to unsatisfactory results in features extraction and also including the ignoring the outliers.
3. A SURF technique was used for extracting the features from each sub image and finds its strong points.
4. Finally, Applying our proposed matching method to find all regions that are similar despite of it is copy move regions or natural similarity regions with minimum false matching. The copies detection are depending on calculated ratio between features extracted from step 3, for farther details mentioned in section 6.

4: OCR

The OCR works to recognize the text and to segment the image , in fact it is a preprocessing step, it converts true-color or gray-scale specific image to a binary one and returns the recognized text, the recognition confidence and the text location in the image[18].

OCR is the operation of extracting text from an image. The main goal of an OCR is to make editable documents from existing document images[16]. Many significant algorithms is required to develop an OCR and essentially it works in two stages such as character and word detection. In other words, OCR also works on sentence detection to maintain a document's structure[19].

The methodology of OCR [18], [20]:

1) Optical Scanning : the image of original file or document is captured by digital camera or any scanner.
2) Pre-processing: It is to enhance the image for best segmentation. Preprocessing include converting gray scale, thinning, binarization, and normalization.
3) Segmentation: three steps of segmentation:
    i) line segmentation divide the character in image horizontal.
    ii) word segmentation divide words from line sentence.
    iii) character segmentation divide the characters from word.
4) Feature extraction: It is to extract important pattern from characteristics. The features differentiate one character from other
5) Training and recognition: It can be done by template matching, statistical technique, structural techniques, and artificial neural networks.
6) Post-processing: include grouping, error detection and correction.

Figure1: a and b show the execution of MATLAB built-in OCR system on the scanned document, the yellow bounding boxes appear the recognized text and the numbers on top demonstrate the recognition confidence of the text.

(a) (b)

Figure 2: (a).Text document image (b).OCR result

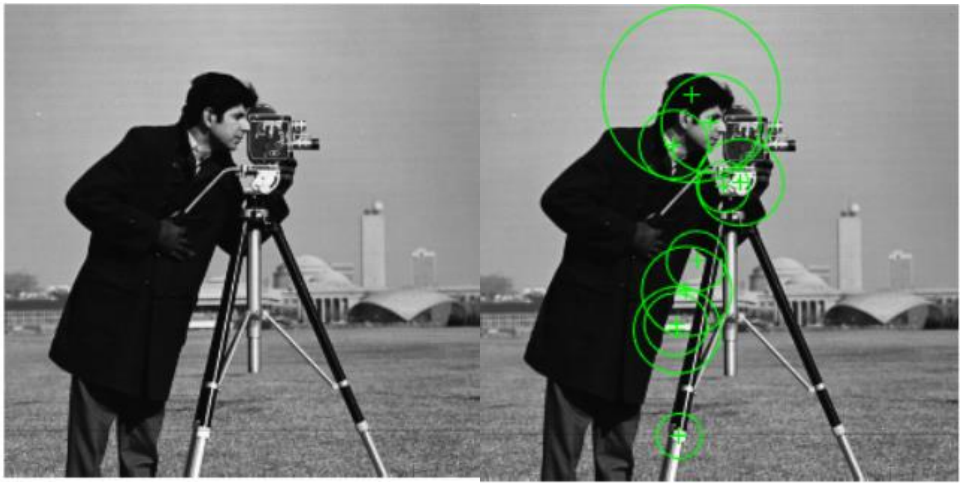
5: Feature Extraction (SURF descriptors)

H. Bay et al. in 2008 proposed an algorithm for feature extraction called SURF, which is a scale and rotation-invariant detector [21]. It is used to classify objects of images based on critical features [1] [22]. SURF is a very immunized against post-processing attacks like rotational transformation, scale, additional noise and change of brightness of an image regions [23], [24]. For that, It is considered as a suited tool for recognition system operations like object detection, object recognition or image retrieval. To detect interest point, SURF used a Hessian matrix approximation. It outperforms SIFT. That's because the SURF integrates the gradient information within a sub patch, whereas SIFT depends on the orientation of the individual gradients[25] . That's why the SURF is a less sensitive to noise as well as lower than SIFT in computational cost[25].

The Hessian matrix H(x;σ) at scale σ is shown in Equation (1)[26]:

$$H(x,\sigma) = \begin{bmatrix} Cxx(x,\sigma) & Cxy(x,\sigma) \\ Cyx(x,\sigma) & Cyy(x,\sigma) \end{bmatrix} \quad (1)$$

where Cxx(x,σ) represents the convolution of the Gaussian second-order, derivative ∂2g(σ)/∂x2 of the image in x, Cxy(x,σ) and Cyy(x,σ) are treated similarly.



(a) (b)

Figure 3: (a).natural scene image (b).SURF result

Figure 4: Texted image and SURF result in MATLAB

6: Proposed Matching

The main challenge in CMF is how to find effective features and best matching methods to find the copied regions. The feature matching is to find a high similarity between feature descriptors. Many methods can be used to process these similarities. The common ways are Euclidean distance in formula (2)[27] between the neighboring vectors, building the k-dimension

tree contain all the feature vectors, correlation, Hamming distance, hashing technique, inner product, shift frequency threshold and many others [16], [24].

The formula of Euclidean distance:

$$(p,q) = (q,p) = \sqrt{\sum(qi - pi)^2} \qquad (2)$$

Where d: The Euclidean distance, p: First vector, q: Second vector

However, every methods give incorrect matches among the regions and this can be increasing when it contains a similar or near texture such as the scanned documents with smooth ,white and little noise background and both inter and intra characters similarity may cause many false matches [8], So When using SURF for extraction feature points, if the numbers of extracted features were large, the time for matching features will increase greatly, and the numbers of false matches also will Increase, thus, the result of matching will go to worse [22]. To solve this problem of matching points in SURF algorithm, proposed some steps were added to reduce those false matches as the following algorithm:

Algorithm 1. Duplicated detection (DD)

1. Compute The Feature Descriptor And Number Of Features For Each Sub Image And Its Size Using SURF Algorithm.
2. Read A Specific Sub Image To Find Its Similarities Called I-Image.

   Create M set, Where Mi ∈ M If (Mi<I-Image+2)Or (Mi>=I-Image-2)

   For I=1,……,N Where N:Number Of Sub Images , Mi: Sub Image.
3. For Each $M_i \in M$, Compute The Euclidean Distance For Feature Descriptor Of I-Image And Feature Descriptor of mi, consequently a match Features descriptor and compute the number of matching features.
4. Compute the ratio between numbers of the feature for I-image and numbers of the matching features:

   Ri = (NumMatch / NumI-image)*100;

   where NumMatch: numbers of the matching features, NumI-image: numbers of the I-image features
5. Compare the ratio with specific threshold (**threshold=30%)

If Ri >= 30% Then

Mi match I-image

End if

go step 4

End

This proposed method reduce the numbers of the checking sub images by using size condition in step 3 (preprocessing step) thus reduced the running time and handle the false matches consequently increase the accuracy of matching results in addition to its pre increase by SURF [1], [21], [28], [29].

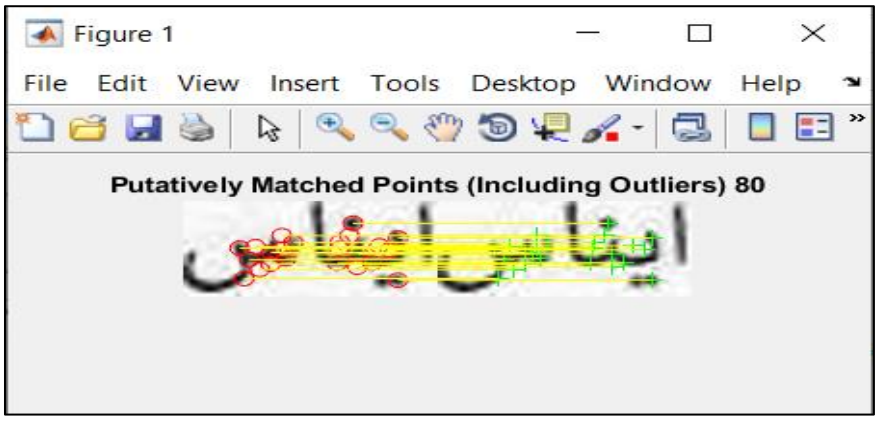


Figure 6: Matching between input image with its similarity.

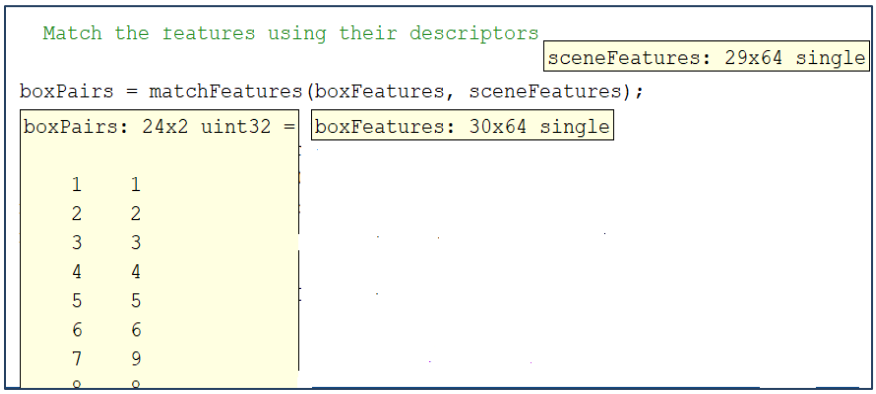


Figure 5: Match the features using their descriptors

7.1. Experimental Setup:

The proposed method has been implemented using Windows 10 pro, processor intel® core i7- 8th Gen 1.80 GHz ,installed RAM 12.0 GB , system type 64-bit and MATLAB R2018a .

Since there is no publicly available Arabic dataset, we made our dataset. It is created by 100 texted images with a font named Calibri and font size between 14-22, printed and scanned by the cannon MG1800 printer with resolution 900x1200. All scanned images were considered as OCR training data and our input image is segmented as a test image to sub-images(words). Then OCR results were sub-images with varying sizes, the SURF algorithm was applied and the sub-images were sized at least 90x90 pixel, if not, a resizing operation is done by three times to get best features extraction for each sub-image after that, the proposed matching method is executed and compute the evaluation metric then compared it with SURF algorithm.

7.2. Evaluation Metrics and results:

The matching method was measured in the true positive rate (TPR)(3), false positive rate (FPR)(4), and Accuracy(5) Where:

$$TPR = \frac{Number\ of\ true\ detected\ matches}{Number\ of\ all\ detected\ images} \qquad (3)$$

$$FPR = \frac{number\ of\ false\ detected\ matches}{Number\ of\ all\ detected\ images} \qquad (4)$$

$$ACC = \frac{TPR+(1-FPR)}{2} \qquad (5)$$

Where (TPR): points to the percentage of true matches images which are correctly detected, (FPR): points to the percentage of false matches images, that are detected as a true one and (ACC): the accuracy of duplicated detection approach.

The results were as table 1:

Table 1: TPR, FPR, ACC and running time for SURF and proposed method

	Size(pixel)	TPR	FPR	ACC	Running time(second)
SURF	Size<=90x90	50.54%	66.66%	25.436%	1.28
	Size>90x90	100%	60%	50.2%	1.87
Proposed method	Size<=90x90	95.3%	0	97.65%	0.15
	Size>90x90	100%	0	100%	0.18

8: Conclusion

CMFD methods have assured to be dependable in the detection of tampering images. many ways are tested on natural scene images and few other worked with scanned text documents because of some extra difficulties on techniques of forgery detection are being faced with the last one. For example, The features that extracted from a text is not as easy to describe as one extracted from a natural image because of the high true similarity between text characters. The second difficulty is the loss of precision when the documents may have been printed and scanned many times. In this paper, this paper aims to accurate and robust the method of finding the duplicated words in scanned document as an essential step in T-CMFD. To show the effectiveness of the proposed method, a comparison between the results in case of using SURF and Euclidian vs. the proposed approach on texted images. The OCR system was applied to segment the text in the original images then executed the two techniques each one at a time and displayed the evaluation metrics such as TPR, FPR, Accuracy rate and the running time for each algorithm. The results of proposed work showed that the proposed approach is fastest and it can handle false matches, so the accuracy is better than SURF.

References

[1] S. Dhivya, J. Sangeetha, and B. J. S. C. Sudhakar, "Copy-move forgery detection using SURF feature extraction and SVM supervised learning technique," in *Soft Computing*, 2020, pp. 1-12: Springer.

[2] B. Yang, X. Sun, H. Guo, Z. Xia, X. J. M. T. Chen, and Applications, "A copy-move forgery detection method based on CMFD-SIFT," in *Multimedia Tools and Applications*, 2018, vol. 77, no. 1, pp. 837-855: Springer.

[3] L. Ye, L. Jing-zhang, Y. Hong, and Y. J. J. o. G. U. Yun, "Study on near-replicas detection algorithm in duplicated text removal," *in Journal of Guangxi University(Natural Science Edition)*, 2010, no. 2, p. 27: en.cnki.com.cn.

[4] L. J. C. A. Shu-yi and Software, "Strategy of eliminating duplicated web pages based on text similarity," in *Computer Applications and Software*, 2011, vol. 28, no. 11, pp. 228-230: en.cnki.com.cn.

[5] P. Carrazoni Entenza, "Copy-move forgery detection in scanned text documents using Zernike and Hu moments," *Master's thesis presented to the telecommunication engineering school*, 2019, castor.det.uvigo.es.

[6] P. Forczmański, A. Smoliński, A. Nowosielski, and K. Małecki, "Segmentation of scanned documents using deep-learning approach," in *International Conference on Computer Recognition Systems*, 2019, pp. 141-152: Springer.

[7] N. K. T. El-Omari, A. H. Omari, O. F. Al-Badarneh, H. J. I. J. o. C. Abdel-jaber, and Communications, "Scanned Document Image Segmentation Using Back-Propagation Artificial Neural Network Based Technique,"in *Journal of Computers and communications* , 2012, vol. 6, no. 4, pp. 183-190: researchgate.net

[8] F. H. N. J. N. M. t. D. T. F. i. S. D. Hasan, "New Method to Detect Text Fabrication in Scanned Documents," *Master's thesis in information technology*, 2015, الجامعة الإسلامية-غزة.

[9] B. Bhosale Swapnali, S. Kayastha Vijay, K. J. I. j. o. t. r. HarpaleVarsha K., "Feature extraction using surf algorithm for object recognition," in *International Journal of Technical Research and Applications* , 2014, vol. 2, no. 4, pp. 197-199: academia.edu.

[10] F. Cruz, N. Sidère, M. Coustaty, V. P. d'Andecy, and J.-M. Ogier, "Categorization of document image tampering techniques and how to identify them," in *International Conference on Pattern Recognition*, 2018, pp. 117-124: Springer.

[11] N. Khanna, G. T. Chiu, J. P. Allebach, and E. J. Delp, "Scanner identification with extension to forgery detection," in *Security, Forensics, Steganography, and Watermarking of Multimedia Contents X*, 2008, vol. 6819, p. 68190G: International Society for Optics and Photonics.

[12] N. Khanna and E. J. Delp, "Intrinsic signatures for scanned documents forensics: effect of font shape and size," in *Proceedings of 2010 IEEE International Symposium on Circuits and Systems*, 2010, pp. 3060-3063: IEEE.

[13] A. Prabhu, Z. Shah, and M. J. I. J. o. C. S. I. Shah, "Robust detection of copy move forgeries for scanned documents using multiple methods," in *International Journal of Computer Science Issues*, 2012, vol. 9, no. 6, p. 436: search.proquest.com.

[14] Y. Sun, J. Qin, and W. Wang, "Near duplicate text detection using frequency-biased signatures," in *International Conference on Web Information Systems Engineering*, 2013, pp. 277-291: Springer.

[15] T. M. D. Wu and K. Y. Sin, "System and method for duplicate text recognition," in *US Patent 8,577,155,* 2013, ed: Google Patents.

[16] S. J. E. I. Abramova, "Detecting copy–move forgeries in scanned text documents," in *Electronic Imaging*, 2016, vol. 2016, no. 8, pp. 1-9: ingentaconnect.com.

[17] K. Sun, G. Cao, Q. Zhao, and J. Zhang, "Differential Abnormality-Based Tampering Detection in Digital Document Images," in *2019 IEEE/ACIS 18th International Conference on Computer and Information Science (ICIS)*, 2019, pp. 145-149: IEEE.

[18] R. Mithe, S. Indalkar, N. J. I. j. o. r. t. Divekar, and engineering, "Optical character recognition," in *International Journal of Recent Technology and Engineering*, vol. 2, no. 1, pp. 72-75: Citeseer.

[19] S. Tajrean and M. A. Yousuf, "Handwritten Bengali Number Detection using Region Proposal Network," in *2019 International Conference on Bangla Speech and Language Processing (ICBSLP)*, 2019, pp. 1-6: IEEE.

[20] D. Rajan Goyal, R. K. Narula, and M. K. Jindal, "A Brief Review on Optical Character Recognition Techniques." in *International Journal on Future Revolution in Computer Science & Communication Engineering*, vol. 5, no. 6, pp. 239-243: ijfrcsce.org.

[21] H. Bay, A. Ess, T. Tuytelaars, L. J. C. v. Van Gool, and i. understanding, "Speeded-up robust features (SURF)," in *Computer vision and image understanding*, 2008, vol. 110, no. 3, pp. 346-359: Elsevier.

[22] C. Wang, Z. Zhang, and X. J. S. Zhou, "An image copy-move forgery detection scheme based on A-KAZE and SURF features," in *Symmetry*, 2018, vol. 10, no. 12, p. 706: mdpi.com.

[23] W. Chen and Q. Cao, "Feature Points Extraction and Matching Based on Improved Surf Algorithm," in *2018 IEEE International Conference on Mechatronics and Automation (ICMA)*, 2018, pp. 1194-1198: IEEE.

[24] X. Bo, W. Junwen, L. Guangjie, and D. Yuewei, "Image copy-move forgery detection based on SURF," in *2010 International Conference on Multimedia Information Networking and Security*, 2010, pp. 889-892: IEEE.

[25] S. A. K. Tareen and Z. Saleem, "A comparative analysis of sift, surf, kaze, akaze, orb, and brisk," in *2018 International conference on computing, mathematics and engineering technologies (iCoMET)*, 2018, pp. 1-10: IEEE.

[26] K. Mikolajczyk and C. Schmid, "Indexing based on scale invariant interest points," in *Proceedings Eighth IEEE International Conference on Computer Vision. ICCV 2001*, 2001, vol. 1, pp. 525-531: IEEE.

[27] M. S. Mahdi and S. N. J. A.-M. J. o. S. Alsaad, "False Matches Removing in Copy-Move Forgery Detection Algorithms," in *Al-Mustansiriyah Journal of Science*, 2020, vol. 31, no. 1, pp. 47-53: iasj.net.

[28] N. B. Abd Warif et al., "Copy-move forgery detection: survey, challenges and future directions," in Journal of Network and computer applications, 2016, vol. 75, pp. 259-278: Elsevier.

[29] R. Krishnan and A. J. J. Anil, "A survey on image matching methods," *in International Journal of Latest Research in Engineering and Technology*, 2016, vol. 2, pp. 58-6: academia.edu.