SIFT BASED IMAGE TAMPERING DETECTION USING OPTIMIZATION

A Dissertation submitted in partial fulfilment of the requirement for the Award of Degree of

MASTER OF TECHNOLOGY IN INFORMATION SYSTEM

Submitted by SHIVANGI GUPTA (2K15/ISY/19) Under the guidance of Ms Ritu Agarwal Assistant Professor



Department of Information Technology Delhi Technological University Bawana Road, Delhi-110042 2015-2017

CERTIFICATE

This is to certify that Shivangi Gupta (2K15/ISY/19) has carried out the major project titled "Sift-based Image Tampering Detection using Optimization" as a partial requirement for the award of Master of Technology degree in Information System by Delhi Technological University.

The major project is a bonafide piece of work carried out and completed under my supervision and guidance during the academic session 2015-2017. The matter contained in this report has not been submitted elsewhere for the award of any other degree.

(Project Guide) Ms. Ritu Agarwal Assistant Professor Department of Information Technology Delhi Technological University Bawana Road, Delhi-110042

ACKNOWLEDGEMENT

I express my gratitude to my major project guide Ms. Ritu Agarwal, Assistant Professor in Department of Information Technology at Delhi Technological University, Delhi for the valuable support and guidance she provided in making this major project. It is my pleasure to record my sincere thanks to my respected guide for her constructive criticism, interminable encouragement and valuable insight without which the project would not have shaped as it has.

I humbly extend my word of gratitude to Dr. Kapil Sharma, Head, IT and all the other faculty members and staff of department for providing their valuable help, time and facilities at the need of hour.

Shivangi Gupta Roll No. 2K15/ISY/19 M.Tech (Information System) E-mail: shivigupta20.sg@gmail.com

ABSTRACT

Nowadays digital pictures are an effective and broadly used communication medium. They importantly affect our life. Because of the widely used editing softwares like Photoshop, it has turned out to be generally simple to make fake pictures. So there is a need to develop strategies that will detect all kinds of tampering in the images. In this thesis we propose a technique to consequently recognize the copied regions in advanced pictures. The nearness of copied regions in a picture may imply a type of fabrication called copy–move forgery.

The proposed technique coordinates both keypoint based and block based phony detection strategies. To start with, the segmentation process fragments the host picture into non-overlapping and irregular squares utilizing SLIC Algorithm. At that point, the feature points are separated from each part as block elements, and the block elements are matched to find the labelled highlight points; this technique can roughly show the presumed tampering regions. To distinguish the forged areas all the more precisely, we propose the tampered region extrication method, where the feature points are replaced with little superpixels as highlighted areas and afterward combines the neighbouring blocks that have comparable nearby shading highlights in the feature blocks to produce the merged areas. At last, it implements the morphological operation to the merged locales to obtain the identified tampered regions.

However, for some Copy Move Forgery (CMF) pictures, these methodologies can't create acceptable detection results. For example, the quantity of the coordinated key-points might be too less to end up being a CMF picture or to produce an exact outcome. Once in a while these methodologies may even deliver errors. To take care of the issue, a novel approach named as CMF Detection with Cuckoo Search Optimization (CMFD-CS) is proposed in this thesis. CMFD-CS coordinates the Cuckoo Search Optimization (CS) algorithm into the SIFT-based structure. Experimental results demonstrate that CMFD-CS has better execution.

CONTENT

Certificate	i
Acknowledgement	ii
Abstract	
List of Figures & Graphs	vi
Chapter 1 Introduction	1
1.1 Overview	1
1.2 Types of Image Forensics	2
1.2.1 Active Approach	2
1.2.2 Passive Approach	2
1.3 Types of Image Forgery	
1.3.1 Image Retouching	
1.3.2 Image Splicing	
1.3.3 Copy move attack	
1.3.4 Image Morphing	
1.3.5 Image Resampling	
1.4 Image Tampering Detection and Prevention methods	4
1.4.1 SVM Classifier	
1.4.2. Duplicate- Move Tampering Detection Using Pixel-Based Approach	
1.4.3 The partition-based copy-move tampering detection approaches	5
Chapter 2. Literature Survey	6
2.1 CMFD Techniques	
2.1.1 Types of Copy Move Tampering Detection Methods	
2.1.1.1 Block-based techniques	
2.1.1.2 Key point-based Strategies	
2.2 Literature Review	
2.2.1 DCT	
2.2.2 DWT	
2.2.3 PCA	
2.2.4 FMT	
2.2.5 Zernike Moments	10
2.2.6 SIFT	10
2.2.7 SURF	10
2.2.8 ORB	

Chapter 3. Problem Statement	12
3.1 Issues in image forgery detection	12
3.1.1 Origin of Data	12
3.1.2 Benchmarking and Standard dataset	13
3.1.3 Duplicate Regions	13
Chapter 4. Proposed Work	14
4.1 Adaptive Segmentation Algorithm	15
4.2 Block Feature Extraction Algorithm	18
4.2.1 Scale-space extrema detection	19
4.2.2 Keypoint localisation	19
4.2.3 Orientation Assignment	
4.2.4 Keypoint Descriptor	20
4.3 Cuckoo Search Algorithm	20
4.4 Block Feature Matching Algorithm	21
4.5 Forgery Region Extraction Algorithm	22
Chapter 5. Results	23
5.1 "MICC-F2000" Dataset	
5.2 Experimental Results.	
5.2.1 Original image	
5.2.2 Resized image	
5.2.3 SLIC image	
5.2.4 SIFT Features Extracted	
5.2.5 Labelled Feature points (before optimization)	
5.2.6 Merged regions (before optimization)	
5.2.7 Tampered region marked (before optimization)	
5.2.8 Tampered region detected (before optimization)	
5.2.9 Labelled Feature points (after optimization)	
5.2.10 Merged regions (after optimization)	
5.2.11 Tampered region marked (after optimization)	
5.2.12 Tampered region detected (after optimization)	
5.2.13 Performance (after optimization)	
5.2.14 Comparison of performance in existing and proposed algorithm	
Chapter 6. Conclusion	31
6.1 Conclusion.	
6.2 Future Scope	
References	

List of Figures & Graphs

Figure 1.1: Image Splicing	3
Figure 1.2: Copy Move Tampering	3
Figure 1.3: Image Morphing	3
Figure 1.4: Image Resampling	4
Figure 2.1: CMFD Techniques	7
Figure 2.2: Steps of CMFD	8
Figure 4.1: Proposed Scheme	15
Figure 4.2: Segmentation using SLIC	16
Figure 4.3: Flow of Segmentation process	18
Figure 4.4: Flow of Forgery Detection	22
Figure 5.1: Examples of Images from MICC-F220 Dataset	23
Figure 5.2: Input Image	24
Figure 5.3: Image after resizing	24
Figure 5.4: Segmentation using SLIC	25
Figure 5.5: SIFT features extricated	25
Figure 5.6: labelled feature points without optimization	26
Figure 5.7: merged areas without optimization	26
Figure 5.8: tampered region marked without optimization	27
Figure 5.9: detection of forged area without optimization	27
Figure 5.10: labelled feature points after applying optimization	28
Figure 5.11: merged areas after applying optimization	28
Figure 5.12: tampered region marked after applying optimization	29
Figure 5.13: detection of forged area after applying optimization	29
Figure 5.14: Performance Measures after applying optimization	30

Graph 5.1: Comparison of performance in existing and proposed algorithm.....30

CHAPTER 1 INTRODUCTION

1.1 OVERVIEW

Nowadays, people use imaging tools such as Magic Wand Tool, Crop Tool, and Slice Tool etc. to edit images and videos. But some are manipulating them in an unethical manner. So to prevent these types of frauds, various detection schemes have come into light. It is desirable to control these falsifications. There are various fields in which these forgeries happen like war, weather forecasting, journalism, social media which may cause misleading and blackmailing. The intruders manipulate the images or videos and send them through Internet which ruins its confidentiality.

In recent years, Image Forensics has developed a way to handle various tampering on digital media. The forgery may be of type: Copy Move Forgery, Image Splicing, Image Phylogeny and other tampering that can happen on images or videos. Human eye is unable to distinguish between the original and tampered images. Therefore detection schemes have been proposed to detect these forgeries. When a part of image is cut and pasted on the same image, it is called copy move forgery. When this happens between two images i.e. cut from one image and paste to other image then it is called image splicing. If geometric transformations such as rotation, scaling, compression etc. are applied on an image then it is image phylogeny. The detection can be done based on two methods i.e. extracting keypoint features and block matching. Both the methods have their respective pros and cons. Many detection techniques have been developed according to these two methods.

The current block-based tampering detection strategies like DCT divide the input pictures into overlapping and regular picture blocks and then, the forged area is obtained by matching blocks of picture pixels. And in the keypoint based tampering detection methods like SIFT, image keypoints are extricated and matched in the complete picture while identifying duplicated areas.

1.2 TYPES OF IMAGE FORENSICS

There are two types of methods for image forensics [3]: active protection, and passive detection. They are explained further:

1.2.1 Active Approach

In this, the image needs some sort of pre-processing, for example, watermark is produced at the time of making the picture. We can distinguish that the image is altered, if exceptional data can't be removed from that acquired picture. Different methods are proposed for giving security to the picture, which is similar to the idea of watermarking like, message validation code, picture hash, picture checksum and picture protecting.

1.2.2 Passive Approach

This approach does not require any information about the image like that in active approach. It is also called blind forensics detection. Various techniques of this approach are: Camera Based, Pixel Based, Format Based, Physics Based, Geometric Based etc.

1.3 TYPES OF IMAGE FORGERY

1.3.1 Image Retouching:

Image Retouching is regarded as less unsafe type of advanced picture tampering than others. It doesn't essentially alter the picture, but rather diminishes certain component of unique picture. This system is famous among magazine photograph editors [11]. This kind of Image fabrication is available in each magazine that will utilize this system to upgrade certain elements of a picture with the goal that it is more appealing.

1.3.2 Image splicing or photomontage:

This method for making tampered pictures is more powerful than picture modifying. In its procedure there is use of at least two pictures, which are consolidated to make a fake picture. Fig. beneath demonstrates to produce an image; by replicating a part from the source picture into an objective picture. It is like a copy paste technique.

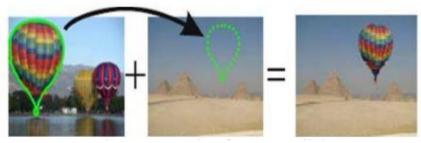


Figure 1.1: Image Splicing

1.3.3 Copy-Move Attack:

The copy move fabrication is popular as one of the troublesome and most ordinarily used method of picture altering. In this strategy, one needs to cover a piece of the picture keeping in mind the end goal to hide the data. In the Copy-Move picture [13], a piece of a same picture is duplicated and glued into another piece of that picture itself. The illustration of Copy-Move tampering is as appeared in figure below.



Figure 1.2: Copy Move Tampering

1.3.4 Image Morphing

It is an adornment in movies and developments that changes one picture or shape into another through a predictable move [11]. Every now and again it is used to depict one individual changing into another through mechanical means or as a part of a fantasy or peculiar gathering.



Figure 1.3: Image Morphing

1.3.5 Image Resampling

In advanced imaging, picture scaling alludes to the resizing of a picture. When scaling a picture, the primitives that make up the picture can be scaled with geometric changes, with no loss of picture quality. When scaling a picture, another picture with a higher or lower number of pixels must be produced.



Figure 1.4: Image Resampling

1.4 IMAGE TAMPERING DETECTION AND PREVENTION METHODS

1.4.1 SVM Classifier:

It is utilized for tampering identification, as exactness of recognizing falsification is upgraded by utilizing SVM classifier [1]. Location of manufacture Image is finished utilizing SVM classifier on methodical premise by planning a basic procedure comprising of 2 stages, i.e. preparing stage and testing stage. In the preparation stage, a database is made and prepared with various pictures. RSA key is additionally set in preparing stage and the client is made a request to enter a similar key in testing stage for distinguishing that client is an approved individual. Some Pre-processing is done on the pictures by changing over into grey scale from rgb. At that point feature extrication is done by breaking down pictures and their pixels.

1.4.2. Copy- Move Tampering Detection Using Pixel-Based Approach

This Algorithm is based on Pixel Based approach [3]. Its genuine working is done as; firstly, dyadic wavelet transform (DWT) is connected to the information picture. This change yield the first Image in a diminished measurement portrayal, i.e., LL1 sub-band. At that point this LL1 sub-band is partitioned into sub-pictures. The Copy-Move locales can be situated by pixel matching. In the last step, the Mathematical Morphological Operations (MMO) is utilized to mark the focuses to enhance the area.

1.4.3 The partition-based copy-move tampering detection approaches:

These are classified as block-based approaches and Non- block-based approaches, both are described as:

• Block-Based Approaches:

Firstly, some Pre-processing such as the colour conversion is done on the selected Images. Many existing recognition strategies needs a merger of the red, green, and blue hues as they work on dark scale pictures. The truth of the matter is that in a copy move procedure to make imitation, the source and the objective areas are both situated in a similar picture subsequently, the manufactured picture must display no less than two comparative locales [20]. It segments the picture into blocks. The blocks can be either square or round. Next, features are marked using any of the transform like Waveletbased, Discrete Cosine Transform (DCT) [17] and so forth. After the extraction of elements, copy move sets are recognized by matching those elements, which should be possible effectively via looking through the squares with comparative component vectors.

- Non-block Based Methods:
 - Sub-image Based methods:

To identify copy move imitations, different techniques that divide the picture into subpictures of same size are sub picture based methodologies. The frequency sub-band of the wavelet disintegration of the picture is first isolated into 4 non-overlapping subpictures [18]. To assess the spatial balance between the copied locale and the pasted one, the stage relationship between each match of sub-pictures is ascertained. At that point, the area of the fabrication is got by moving the information picture in-accordance with the obtained offset and removing this moved picture from the first input picture.

CHAPTER 2

LITERATURE SURVEY

The principle point of CMFD is to uncover clone areas in the picture, to the extremely slight contrast in each of such locales. Picture criminological is one of the debated issues of research as the advanced information is developing so fast over a time. As a result of this numerous specialists and the general population who are intriguing in this examination field distributed extensive number of diaries and gathering papers on CMFD, large portions of them are clarified in our writing. This specific part gives us a survey of as of now accessible writing of important research. The writing is sorted out such that it starts with a concise foundation of CMFD procedures, after that some past deals with CMFD are examined.

2.1 CMFD Techniques

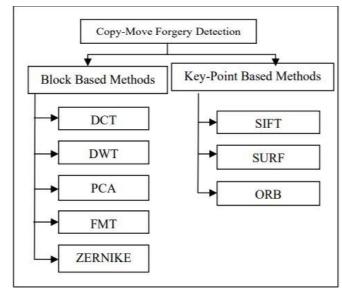
CMFD should be possible effortlessly and successfully if some regular phony operations are utilized to adequately and effectively making it the most incessant fraud detection which is used to change the pictures [3]. CMF is not simply copy and paste one of numerous areas elsewhere in that picture, it is bit keen work. In pragmatic CMF, many other sort of picture preparing operations are likewise included. These picture handling operations are arranged into predominantly two types-

Pre-processing: This operation is utilized for giving some sort of relationship between the copy region and remaining picture. Examples of some pre-processing operations are scaling, shearing, change in brightening, pivot, and change in chrominance or reflecting and so on.

Post-processing: These operations are utilized for evacuating any sort of noticeable hints of CMF operation, such as expelling sharp edges. JPEG compression, additive noise and blurring are the most well-known post processing operation. It is recommended that any of the CMFD calculation ought to must be enthusiastic to different sort of post handling operations.

In this manner, the accompanying prerequisites must be accessible in any CMFD algorithm: The time complexity of CMFD algorithm ought to be sensible, even in the presence of false positives. It ought to be powerful against any sort of Pre-processing or Post-processing operation.

2.1.1 Types of Copy Move Tampering Detection Methods



Copy Move Tampering Detection techniques are mainly categorized in 2 types:-

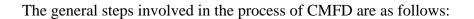
Figure 2.1: CMFD Techniques

2.1.1.1 Block-based techniques:

In block based strategies [20], input picture is firstly isolated into overlapping squares and afterward extraction of each piece is done and then matching is performed between each square to identify the tampered region. Various block based strategies include DCT(Discrete Cosine Transform), DWT(Discrete Wavelet Transform), FMT(Fourier Mellin Transform), PCA(Principal Component Analysis) etc.

2.1.1.2 Keypoint based strategies:

In Key-Point based techniques [20], information or test picture is firstly separated into corner or confined focuses to give nearby components of the picture. The Key-Point calculation for identifying of copy move fabrication begins by extricating high entropy areas i.e. Keyfocuses. Feature descriptors are separated from these elements. These element descriptors are contrasted with each other to detect the coordinated Key-Points and consequently forgery is recognized. The outstanding Key-Point descriptors are SURF (Speed-Up Robust Feature), ORB (Oriented Rotation and BRIEF) and SIFT (Scale Invariant Feature Transform)



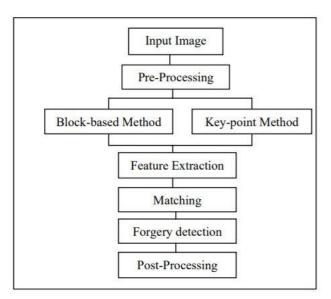


Figure 2.2: Steps of CMFD

- Pre-Processing: This procedure is application dependent. It includes picture change from shaded picture to dim scale and picture enhancement to expel the noise from input pictures.
- Feature Extraction: Feature extraction is the procedure to discover features from input picture for new portrayal of the picture in great way. Elements ought to have two essential necessities: it ought to be maintaining a distance from repetition in the first picture and decrease dimensionality of information.
- Matching: Matching is the procedure to locate a high comparability or coordinating between descriptors and if likeness between feature descriptors is found then it is deciphered as intimation for copied locales. Different procedures of matching have been created for instance Euclidean distance, KD tree, lexicographic arranging and g2NN (generalised 2 nearest neighbour).
- Post-Processing: When a picture has been named non valid; present preparing helps in discovering which change has been utilized between unique zone and its copy moved adaptation. Different algorithms have been proposed in writing such from RANSAC (Random Sample Consensus), same Affine Transformation Selection (SATS).

2.2 Literature Review

In writing, numerous CMFD strategies have been discussed in many research papers. As a matter of first importance, exhaustive search is the immediate response to this CMFD issue where a similar picture is contrasted. In any case, the primary inconvenience of this approach is that it is computationally extremely costly and would take M. N entire square strides for a picture of size 'M \times N'. In addition, this thorough inquiry won't work on account of having changes made on the copied range. The existing strategies for CMFD contrasts from each other as far as number of features extricated, sort of components utilized for matching the picture blocks. Many research papers are written based on these strategies. Maybe a couple of them are outlined beneath:

2.2.1 DCT

S. Kumar, S. Mukherjee proposed a Copy–Move fabrication recognition utilizing **Discrete Cosine Transformation (DCT)**: firstly the shading picture was changed over from RGB shading space to YCbCr shading space and after that the R,G,B and Y-part was splitted into settled size overlapping blocks and, elements separated from the DCT [17] portrayal of R,G,B and Y-segments picture square. The element vectors got lexicographically arranged to make comparable picture blocks neighbours and copied picture squares distinguished utilizing Euclidean distance as comparability measure.

2.2.2 DWT

Later on, Cao, Y., Gao, T., Fan, L. proposed a method using **Discrete Wavelet Transform** (**DWT**) [4] which is used to identify copy move imitation in advanced pictures. In DWT, rather than isolating the information pictures into overlapping blocks; input picture was partitioned into four sub-groups. The lower frequency band was additionally subdivided into overlapping squares to decrease the no. of squares which speeded up the procedure and more energy concentrated to bring down sub-band. The discrete wavelet transform was utilized to lessen the dimensionality and favourable position of DWT more than Fourier transform. It caught both frequency and location data (in time). It is joined with SVD and SIFT.

2.2.3 PCA

PCA (**Principal Component Analysis**) [1] which was proposed by Christlein, V., Riess, C. was a block based technique and it was utilized to diminish the measurement of picture. First, test picture and its measurements were diminished using PCA. PCA restores the key part coefficients of a lattice (say X). Lines and section of this network speaks to coefficients for

one main segment. The number of foremost segment sections was taken by work. PCA is consolidated with SURF (speeded up robust features) or SIFT (scale invariant feature transform) to join favourable circumstances of both block based and Key-Point based procedures to improve the speed and assessment measurements to distinguish the Copy-Move tampered area.

2.2.4 FMT

FMT is a **Fourier Millen Transform**. [1] In science, as discussed by Christlein, V., Riess, C., the Mellin transform is an essential change that might be viewed as the multiplicative form of the two sided Laplace change. In this strategy, extraction of components from the picture blocks would not exclusively be robust to lossy JPEG compression, noise expansion or blurring.

2.2.5 Zernike moment

As discussed by S. J. Ryu, M. J. Lee, and H. K. Lee, this technique was used to confine the Copy-Move falsification region in advanced pictures. The size of Zernike moment [14] was arithmetically invariant and the proposed technique was identifying produced region despite the fact that it was pivoted. This plan was likewise suitable to distinguish manufactured region by Copy-Rotate-Move imitation and resistive to noise, for example, additive white Gaussian Noise (AWGN), JPEG compression and blurring.

2.2.6 Scale Invariant Features Transform (SIFT):

SIFT was proposed by Hailing Huang, Weiqiang Guo, Yu Zhang in 2008 and it is extremely proficient strategy to distinguish tampered area. It is simply scale invariant as well as gives great location results to enlightenment and invariant perspective transforms. The Key-Point extricated by SIFT [10] are invariant to scaling since size of each Key-point is diverse. The descriptors are appointed to neighbourhood focuses as Key-features. After that every descriptors are contrasted with each other and coordinated descriptors are used to distinguish the Copy-Move forged area.

2.2.7 Speeded-up Robust Features (SURF):

SURF, proposed by Bo, X., Junwen, W., Guangjie, L., & Yuewei, is used to remove components and it is a robust nearby element indicator. SURF depends on wholes of 2D Haar Wavelet reactions. It makes a proficient utilization of Integral pictures. SURF's finder and

descriptor is said to be speedier and at same time robust to noise, discovery relocations and geometric and photometric distortions. SURF [9] is invariant to geometric change, for example, scaling and pivot. It can recognize numerous cloning and has high computational proficiency. It doesn't give great outcomes when tempered region is little. SURF components can be separated utilizing the accompanying strides:

- Integral Image
- Key-point detection
- Orientation Assignment
- Feature Descriptor Generation

2.2.8 Oriented FAST and Rotated Brief (ORB):

ORB depends on FAST identifier and BRIEF descriptor as told by Ye Zhu, Xuanjing Shen, Haipeng Chen. Because of this reason, it is called ORB (Oriented Fast and Rotated Brief) [19]. Both these systems are alluring because of good execution and ease. FAST and its variations were the strategy for decision to finding the key-focuses continuously frameworks. Filter and SURF identifiers incorporate key-point introduction however FAST finder does exclude key-point introduction. There were numerous approaches to decide the key-point introduction, histogram of inclination and estimate by square examples. Sphere is revolution invariance and resistive to noise. The effectiveness of ORB was tentatively decided on a few true applications i.e. protest recognition.

CHAPTER 3

PROBLEM STATEMENT

Constructing digital image forgeries as in knowledge of every individual has turned out to be simple either because its source is a solitary picture or different pictures in view of the accessibility and simplicity of intense tools of modelling images and different software, for example, "Photoshop". Here the fabrication is finished by replicating a specific part of a picture and afterward pasting that copied part in that same picture to shroud an essential picture data.

To the greatest extent the various procedure of CMFD experiences the ill effects of the issue of computational cost (beside different issues, for example, strength to any post or preprocessing operation to betray the indicators). This is on the grounds that the greater part of these calculations separated the picture into overlapped little squares (for instance 8 x8) moved by one pixel to one side (or to the base). At that point a few sorts of components called keypoint are acquired from each such covered square. After the extraction of components these are contrasted with each other to locate an arrangement of similar blocks.

3.1 ISSUES IN IMAGE FORGERY DETECTION

3.1.1 Origin of Data

This is important for authoritative prerequisites in numerous applications, for example, in restorative science, money related exchange and other govt. lawful interest. The information correctness is important in numerous every day circumstances, wherever the esteem and reliability of data is required [1].

The greatest risk to the computerized data is the outdated nature of innovation as opposed to the implicit physical delicacy of advanced media. The suspicion of moore's law flopped if there should arise an occurrence of advancement in innovation in future, it is substantially quicker than it was expected. In view of such fast improvement unmistakably at present accessible innovation, programming, gadgets utilized for information creation, stockpiling, recovery will be supplanted inside 3-5 years. This makes the protection of advanced information and confirmations a requesting issue.

3.1.2 Benchmarking and Standard dataset

Open informational collection is needed for sensible and some basic conditions, for example, advanced record is required. Give us a chance to take a case of advanced record like picture with various size, distinctive determination, diverse camera demonstrate and with various sorts of conceivable altering, for example, picture grafting, CM fraud [3], picture correcting, picture transforming and other sort of control like shading alteration, obscuring, differentiate modification and different other post processing operation like including noise, picture compression and so forth. In view of this, the requirement for advancing benchmarks for dataset is there.

3.1.3 Duplicate Regions

The reason of the presence of copy areas in a picture is amongst two things: in the first place, the nearness of two things or two articles with a similar size, shape, and shading; one of them might be a duplicate from the other one [1]. Second, the nearness of a moderately huge range with one shading and close in attributes, for example, backgrounds (sky, divider, and so forth) which prompts the presence of similar regions in the outcomes.

CHAPTER-4

PROPOSED WORK

The proposed technique coordinates both block based and key-point based forgery detection strategies. To start with, the segmentation [5] process fragments the picture into non-overlapping and irregular squares utilizing SLIC Algorithm [6]. At that point, the feature points are separated from each piece as block elements, and the block elements are matched with each other to find the labelled highlight points; this technique can roughly show the presumed tampering regions. To distinguish the forged areas all the more precisely, we have discussed the forgery region extrication method, which replaces the feature points with little superpixels as feature blocks and afterward combines the neighbouring blocks that have comparable nearby shading highlights into the feature blocks to produce the merged areas. At last, it applies the morphological operation [6] to the merged locales to obtain the identified tampered regions.

However, for some Copy Move Forgery (CMF) pictures, these methodologies can't create acceptable detection results. For example, the quantity of the matched key-points might be too less to end up being a CMF picture or to produce an exact outcome. Once in a while these methodologies may even deliver errors. To take care of the issue, a novel approach named as CMF Detection with Cuckoo Search Optimization (CMFD-CS). CMFD-CS coordinates the Cuckoo Search Optimization (CS) algorithm into the SIFT-based structure [16].

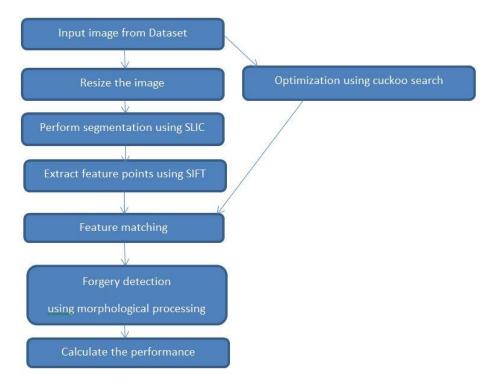


Figure 4.1: Proposed Scheme

4.1 Adaptive Segmentation Algorithm

This is like the customary block based fraud location techniques and can partition the host picture into squares. In earlier years, a lot of block based imitation discovery methods have been implemented. Of the current square based fraud location plots, the host picture was generally isolated into overlapping standard blocks. At that point, the tampered areas were distinguished by matching those squares. Along these lines, the recognized areas are constantly made out of standard blocks, which were not the exact fraud areas; therefore, the review rate of the square based strategies is constantly low. Additionally, when the measure of the host pictures builds, the coordinating algorithm of the overlapping blocks will be significantly costlier. So, the segmentation strategy [6] has been proposed, that can portion the host picture into non-overlapping locales of unpredictable shape as picture blocks, a while later, the forged areas can be identified by coordinating those non-covering and irregular areas. Since we should separate the host picture into non-overlapping areas of not a regular shape and in light of the fact that the super pixels are perceptually significant nuclear locales that can be acquired by finished division, The simple linear iterative clustering (SLIC) algorithm has been used to fragment the host picture in important unpredictable super pixels, as separate squares. The SLIC algorithm adjusts a k-means clustering way to deal with efficiently produce super pixels, and that sticks to the limits very clearly. Figure demonstrates the diverse blocking/division techniques, here (a) demonstrates covering & rectangular blocking, (b) demonstrates overlapping & round blocking, and (c) demonstrates non-covering & irregular blocking with SLIC division strategy. Utilizing SLIC division technique, the non- covering division may diminish the computational costs contrasted covering blocking; moreover, much of the time, the irregular and significant locales can indicate the fraud area superior to the customary squares.

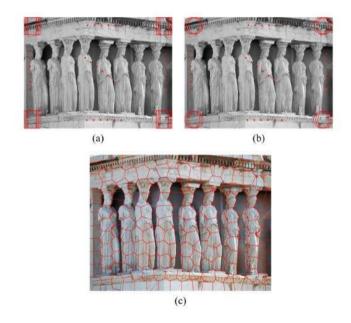


Figure 4.2: (a) overlapping & rectangular blocking, (b) overlapping & round blocking (c) non-overlapping & irregular blocking using SLIC division strategy

The size of the super pixels in SLIC is difficult to choose. It is according to the surface of the picture. If the surface of picture is smooth, size of the super pixels can be set vast and if the surface of picture is not smooth, size of the super pixels can be set little, for great tampering detection results. In this technique, the Discrete Wavelet Transform (DWT) is utilized to examine frequency dispersion of picture. If the low- frequency energy represents most of the frequency energy, picture will be smooth; but if the low- frequency energy represents just a small portion of the frequency energy, picture has all the earmarks of being a definite picture. A 4-level DWT [6] is performed, utilizing the "Haar" wavelet, on the host picture; at that point, the low-frequency energy E_{LF} and high-frequency energy E_{HF} would be ascertained utilizing equations 1 and 2, individually. Using the low- frequency energy E_{LF} and high-frequency dispersion P_{LF} can be figured utilizing equation 3. Hence the size S of the super pixels can be shown in equation 4.

$$E_{LF} = \sum_{i} |CA_{4}|$$
(1)

$$E_{HF} = \sum_{i} (\sum_{i} |CD_{i}| + \sum_{i} |CH_{i}| + \sum_{i} |CV_{i}|),$$

$$i = 1, 2, ..., 4$$
(2)

where CA_4 represents the approximation coefficients at the 4th level of DWT; and CD_i , CH_i and CV_i represents the detailed coefficients at the ith level of DWT, i =1,2,...,4.

$$P_{LF} = \frac{E_{LF}}{E_{LF} + E_{HF}} \cdot 100\%$$
(3)
$$S = \begin{cases} \sqrt{0.02 \times M \times N} & P_{LF} > 50\% \\ \sqrt{0.01 \times M \times N} & P_{LF} \le 50\% \end{cases}$$
(4)

where S implies size of super pixels; $M \times N$ shows measure of picture; and P_{LF} implies rate of low-frequency dissemination.

The steps of the existing Segmentation strategy are depicted in Figure. DWT is applied to picture for getting the coefficients of low-and high-frequency sub-groups of the host picture. Rate of the low-frequency dispersion PLF utilizing (3) is examined, as per which the underlying size S is chosen, utilizing (4). At last, the SLIC division algorithm is applied with the size S for fragmenting the picture thereby acquiring picture blocks [5].

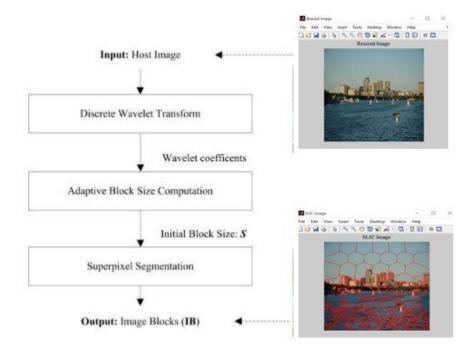


Figure 4.3: Flow of Segmentation process

4.2 Block Feature Extraction Algorithm

The square elements from the image blocks (IB) are extricated. Conventional square based tampering identification strategies extricated elements of an indistinguishable length from the block features or specifically utilized the pixels of the picture as block features. Furthermore, the components are not affected to different picture changes. Along these lines, in this strategy, we extricate feature focuses from each picture obstruct as square components, and the element focuses ought to be active to different enhancements, for example, picture scaling, revolution, and JPEG compression. We picked SIFT [8] as the component guide extraction strategy toward extricate the element focuses from every picture block, then every square is described by the SIFT feature points which were removed in the relating square. In this way, each square element consists of irregular block area data and separated SIFT include points.

The SIFT [12] algorithm removes unmistakable components of nearby picture patches which are invariant to picture scale and are vigorous to changes in noise, light and bending. It consists of four major steps:

- Scale-space extrema detection;
- Keypoint localization;
- Orientation assignment;

• Keypoint descriptor.

4.2.1 Scale-space extrema detection.

The initial step of the calculation looks for extrema over all scales and picture areas. Given an information picture I(x,y), at that point the scale space of picture I is characterized as takes after:

$$L(x,y,\sigma)=G(x,y,\sigma) *I(x,y)$$

where * is the convolution operation in x and y directions, and the Gaussian function

$$G(x, y, \sigma) = \frac{1}{2\pi\sigma^2} e^{-\frac{x^2+y^2}{\sigma^2}}$$

where σ is the factor of scale space. Keeping in mind the end goal to effectively recognize potential plot indicates that is invariant scale, which are likewise called keypoints in SIFT structure, the technique utilized the scale-space extrema in the difference of-Gaussian (DoG) work convolved with the picture, D(x,y, σ), which can be figured from the distinction of two adjacent scales isolated by a consistent multiplicative factor k :

$$D(x, y, \sigma) = [G(x, y, k\sigma) - G(x, y, \sigma)] * I(x, y)$$
$$= L(x, y, k\sigma) - L(x, y, \sigma)$$

The convolved pictures are assembled by octave, and an octave compares to multiplying the estimation of σ [12]. At that point the estimation of k is chosen so we acquire a settled number of obscured pictures per octave. This additionally guarantees similar quantities of DoG pictures are produced per octave. When DoG pictures have been acquired, keypoints are distinguished as nearby minima or maxima of the DoG pictures crosswise over scales. This is finished by looking at every pixel in the DoG pictures to its eight neighbours at a similar scale and nine comparing neighbouring pixels in each of the neighbouring scales. In the event that the pixel esteem is the most extreme or least among all thought about pixels, it is chosen as a competitor keypoint.

4.2.2 Keypoint localization.

Scale-space extrema discovery creates an excessive number of keypoint applicants, some of which are unstable. At that point keypoints are sifted in this progression with the goal that stable keypoints are held. Once a keypoint has been found by contrasting a pixel with its neighbours, next is to play out a point by point fit to the close-by information for precise area,

scale, and proportion of flows [8]. This data enables focuses to be rejected that have low difference (and are thusly touchy to noise) or are ineffectively limited along an edge.

4.2.3 Orientation assignment.

This is the key stride to accomplish invariance to picture rotation, in which each keypoint is relegated at least one of the orientations based on local image gradient directions. The keypoint orientation is calculated from an orientation histogram of neighbourhood slopes from the smoothed picture which are near to each other $L(x,y,\sigma)$. For each picture test L(x,y) at the keypoint's scale σ , the slope magnitude m(x,y) and orientation $\theta(x,y)$ is processed utilizing pixel contrasts, let

$$\begin{split} m(x,y) &= \sqrt{L_1^2 + L_2^2} \\ \theta(x,y) &= \arctan(L_2 / L_1) \\ \text{where } L_1 &= L(x+1,y,\sigma) - L(x-1,y,\sigma) \text{ , and } \\ L_2 &= L(x,y+1,\sigma) - L(x,y-1,\sigma) \text{ .} \end{split}$$

An introduction histogram with 36 bins, with each container covering 10 degrees, is framed from the gradient orientations of test focuses inside a district around the keypoint. At that point the most extreme orientation is doled out to this keypoint [10]; extra keypoints will be made with orientation which is inside 80% of the greatest orientation.

4.2.4 Keypoint descriptor.

The past operations have distributed a picture area, scale, and orientation to all keypoints, which guarantee the invariance to picture rotation, area and scale. And afterward we need to register descriptor vectors for every keypoint with the end goal that descriptors are particular and strong to different varieties, for example, enlightenments and so forth. Calculate the component descriptor as an arrangement of orientation histograms on 4×4 pixel neighbourhoods [12]. Orientation histograms are in respect to the keypoint orientation and the orientation originates from the Gaussian picture nearest in scale to the keypoint's scale.

4.3 Cuckoo Search Optimization

Copy Move Forgery (CMF) is one of the straightforward and viable operations to make computerized pictures. Strategies such as Scale Invariant Features Transform (SIFT) are generally used to distinguish CMF. Different methodologies under the SIFT-based structure are the most worthy approaches to CMF recognition because of their vigorous execution. In any case, for some CMF pictures, these methodologies can't create acceptable location results. For example, the quantity of the matched key-focuses might be too less to turn out to be a CMF picture or to create an exact outcome [7]. Once in a while these methodologies may even create mistakes. To take care of the issue, a novel approach named as CMF Detection with Cuckoo Search Optimization (CMFD-CS) [16] is proposed. CMFD-CS incorporates the Cuckoo Search Optimization (CS) algorithm into the SIFT-based structure. It uses the CS algorithm to create tweaked parameter esteems for pictures, which are utilized for CMF location under the SIFT-based system. Experimental results demonstrate that CMFD-CS has great execution.

Cuckoo-Search Algorithm [15] is portrayed utilizing three glorified principles: (i) Each cuckoo lays one egg on the double, and dump its egg in discretionarily picked settle; (ii) The best homes with high quality of eggs will proceed to the next generations; (iii) The quantity of open host homes are fixed, and the egg laid by a cuckoo is found by the host fledgling with a probability $p_a \in [0, 1]$. For this circumstance, the host winged animal can either dispose of the egg or surrender the home and fabricate an absolutely new home. For ease, this last assumption can be approximated by the division p_a of the n settle are supplanted by new homes (with new arbitrary arrangements).

Algorithm-

Objective function $f(z), z = (z_1, ..., z_d)^T$ Generate initial population of n host nests $z_i (i = 1, 2, ..., n)$ while t < MaxGeneration or stop - criterion do Get a cuckoo randomly by Levy flights Evaluate its quality/fitness F_i Choose a nest among n (say, j) randomly if $F_i > F_j$ then replace j by the new solution end if Fractions (P_a) of worse nests are abandoned and new ones are built Keep the best solutions (or nests with quality solutions) Rank the solutions and find the current best end while

4.4 Block Feature Matching Algorithm

After we have gotten the block features (BF), we should find the matched squares through the square elements. To begin with, the quantity of coordinated key points is ascertained, and the connection coefficient produced; at that point, the comparing square matching limit is computed adaptively; with the outcome, the coordinated block sets are found [5]; and finally,

the coordinated element focuses in the coordinated squares are extricated and marked to find the position of the speculated tampered area.

4.5 Forgery Region Extraction Algorithm

The labelled feature points (LFP) were removed [6] that are just the locations of the fraud areas, even the tampered locales are found. As superpixels can fragment picture very well, a technique by supplanting the LFP with little superpixels to acquire the suspected regions (SR) is discussed. Moreover, for enhancing the accuracy and review result, the nearby shading feature of the superpixels that are neighbours to the suspected regions (SR) are measured; if their shading feature matches that of the SR, then the neighbour superpixels into the comparing SR are consolidated, which creates the merged regions (MR) [6]. At last, a morphological operation is connected with blended areas for creating the identified copy move tampered areas.

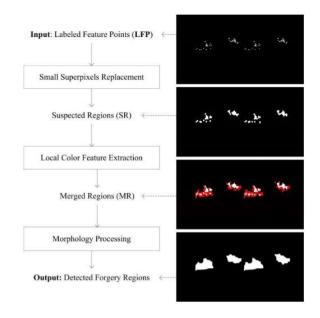


Figure 4.4: Flow of Forgery Detection

CHAPTER 5

RESULTS

The proposed strategy is assessed on numerous unique pictures with various sorts of Copy Move falsifications. We built up a database of 100 pictures including true pictures and distinctive sorts of copy move produced pictures in light of the bona fide pictures taken by various camera models and from an outstanding dataset MICC-F220. To test and look at the execution of our proposed technique with other fraud identification strategies this dataset is utilized. In this investigation the execution of copy move falsification discovery framework will be tried on (100) pictures that are chosen from this dataset.

5.1 "MICC-F220" Dataset

It contains an aggregate of 220 pictures, in which 110 pictures are real and 110 are altered. The majority of the 220 pictures have a uniform size of 737×492 pixels and having the same JPEG format. The produced pictures are acquired by arbitrarily choosing a rectangular fix and sticking it over the first picture after a few distinct assaults (revolution, scaling, interpretation, and so on.). Figure represents a few cases of legitimate and altered pictures from "MICC-F220" dataset, where every picture in top column is authentic and the pictures in the base line are altogether altered.



Figure 5.1: Examples of Images from MICC-F220 Dataset

5.2 Experimental Results:

5.2.1 Original image

The forged image is chosen from the MICC-F220 dataset for detecting the forgery.

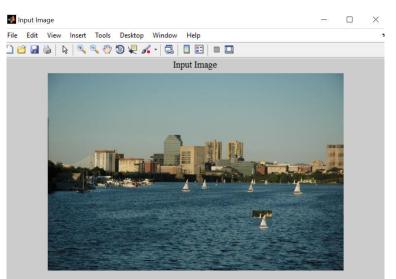


Figure 5.2: Input Image

5.2.2 Resized Image

The chosen image is resized then.

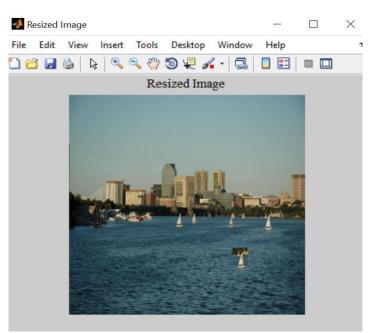


Figure 5.3: Image after resizing

5.2.3 SLIC image

Then the image is segmented to non-overlapping and irregular blocks using the SLIC(Simple Linear Iterative Clustering Algorithm).

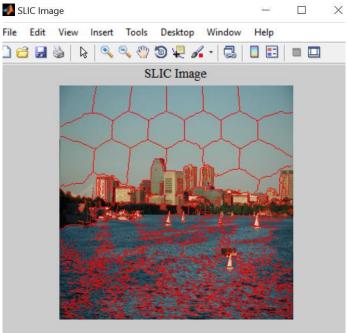


Figure 5.4: Segmentation using SLIC

5.2.4 SIFT Features extracted

Keypoints are extricated from the segmented picture. And are marked as shown in the figure.

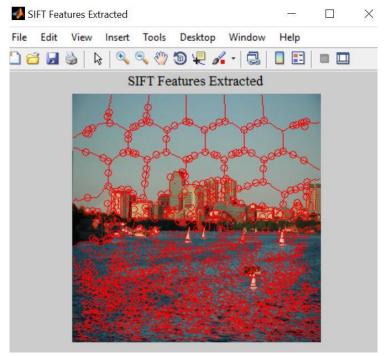


Figure 5.5: SIFT features extricated

5.2.5 Labelled Feature Points (Before optimization)

Key features are matched and the LFP are marked to obtain the suspected regions(SR). The regions which were not tampered also are marked in the picture because of the false positives.

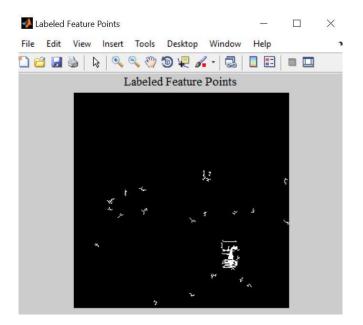


Figure 5.6: labelled feature points without optimization

5.2.6 Merged Regions (Before optimization)

The neighbourhood superpixels are compared to the suspected regions to obtain the merged regions(MR).

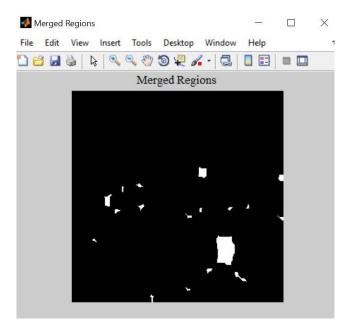


Figure 5.7: merged areas without optimization

5.2.7 Tampered region marked (Before optimization)

At last morphological operation is performed to detect to forged areas.

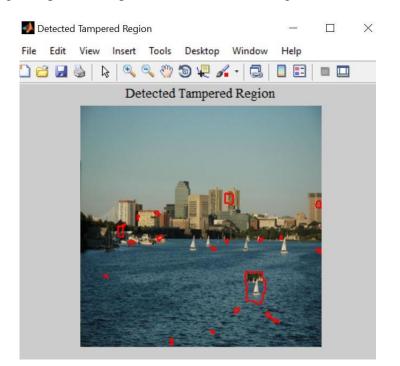


Figure 5.8: tampered region marked without optimization

5.2.8 Tampered region detected (Before optimization)

The figure shows the tampered area along with the non tampered area because of the false positives.

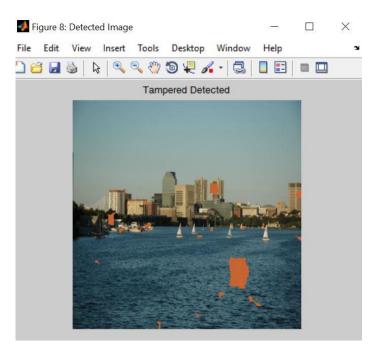


Figure 5.9: detection of forged area without optimization

5.2.9 Labelled Feature Points (After optimization)

Cuckoo Search Optimization is applied to reduce the false positives and for better performance. Here the suspected region is only the forged region.

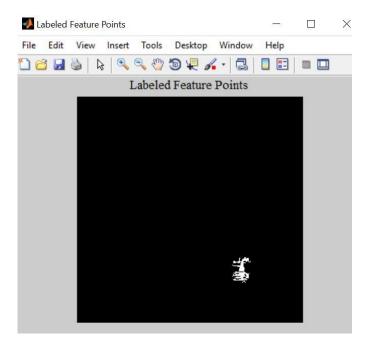


Figure 5.10: labelled feature points after applying optimization

5.2.10 Merged Regions (After optimization)

Suspected region is compared to neighbourhood superpixels to obtain the merged regions.

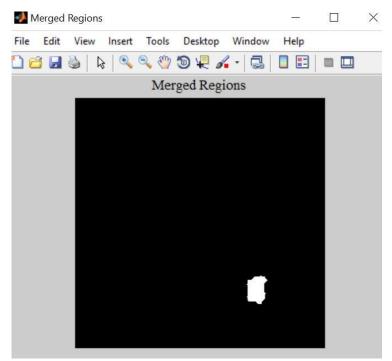


Figure 5.11: merged areas after applying optimization

5.2.11 Tampered region marked (After optimization)

The tampered region is marked in the image.

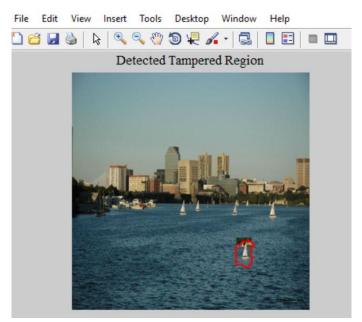


Figure 5.12: tampered region marked after applying optimization

5.2.12 Tampered region detected (After optimization)

Figure shows the forged area correctly after applying optimization.

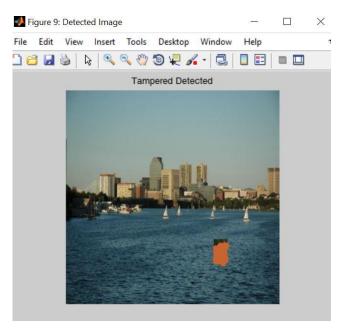


Figure 5.13: detection of forged area after applying optimization

5.2.13 Performance(After Optimization)

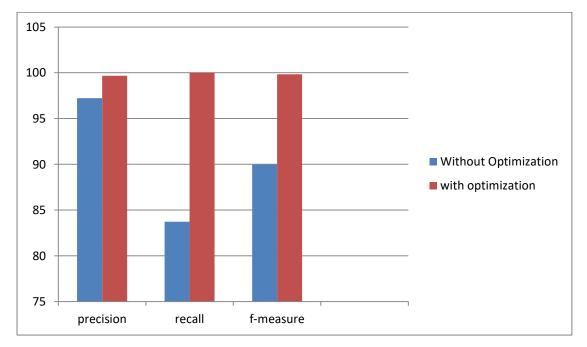
ile	Edit	View	Insert	Tools	Desktop	Window	Help
				1			
	Precisio		Recall	242 - 223	Aeasure		
h	0.9	9967		1	0.9984		

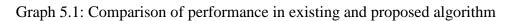
Figure 5.14: Performance Measures after applying optimization

5.2.14 Comparison of performance in existing and proposed algorithm

This graph shows the performance increase in terms of precision, recall and f-measure.

The proposed technique shows better results than existing technique.





CHAPTER 6 CONCLUSION

6.1 Conclusion

Tampered pictures made with copy move operations are identified. In this, a novel duplicate move tampering discovery plot is proposed utilizing segmentation using SLIC and feature point matching and optimization. The algorithm is implemented to fragment the picture into non-overlapping and irregular squares as indicated by the given host pictures; utilizing this method, for each picture, we can decide a fitting block beginning size to upgrade the precision of the forged location result and, in the meantime, diminish the computational costs. At that point, the component points are extricated as block features using SIFT, and then cuckoo search optimization is applied to reduce the false positives and obtain better performance. Block Feature Matching method is discussed, in which the block features are coordinated to one another for finding the labelled feature focuses; this methodology can roughly show the suspected tampered areas. In this manner, to recognize the more exact fraud regions, the Forgery Region Extrication algorithm is discussed, in which the labelled features focuses are supplanted with little superpixels as feature blocks, and neighbouring element hinders with nearby shading features that are like element squares are converged to create the merged areas. Then, the morphological operation is connected to the consolidated areas to produce the identified imitation regions. CMFD-CS coordinates the Cuckoo Search Optimization (CS) algorithm into the SIFT-based structure. Experimental results demonstrate that CMFD-CS has better execution.

6.2 Future Scope

- In future, any other optimization algorithms like Particle Swarm Optimization or Firefly Algorithm etc. can be applied instead of Cuckoo search for better performance.
- In the proposed algorithm, SIFT features have been extracted from the segmented image. Instead of SIFT many other features extraction methods can be used. For example: SURF, Harris features etc.

REFERENCES

[1]Christlein, V., Riess, C., Jordan, J., Riess, C., & Angelopoulou, E. (2012). An evaluation of popular copy-move forgery detection approaches. *IEEE Transactions on information forensics and security*, 7(6), 1841-1854.

[2] Granty Regina Elwin J, Aditya T S, and Madhu Shankar S, "Survey on Passive Methods of Image Tampering Detection, " in Proceedings of the International Conference on Communication and Computational Intelligence, 2010, pp. 431-436.

[3]Bayram, S., Sencar, H. T., & Memon, N. (2008, September). A survey of copy-move forgery detection techniques. In *IEEE Western New York Image Processing Workshop* (pp. 538-542). IEEE.

[4]Cao, Y., Gao, T., Fan, L., & Yang, Q. (2012). A robust detection algorithm for copymove forgery in digital images. *Forensic science international*, 214(1), 33-43.

[5]Li, J., Li, X., Yang, B., & Sun, X. (2015). Segmentation-based image copy-move forgery detection scheme. *IEEE Transactions on Information Forensics and Security*, *10*(3), 507-518.

[6]Pun, C. M., Yuan, X. C., & Bi, X. L. (2015). Image forgery detection using adaptive over segmentation and feature point matching. *IEEE Transactions on Information Forensics and Security*, *10*(8), 1705-1716.

[7] S. Wenchang, Z. Fei, Q. Bo and L. Bin, "Improving image copy-move forgery detection with particle swarm optimization techniques," in Proc. in China Communications, vol. 13, no. 1, pp. 139-149, Jan. 2016.

[8]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 and Security, 6(3), 1099-1110.

[9]Bo, X., Junwen, W., Guangjie, L., & Yuewei, D. (2010, November). Image copy-move forgery detection based on SURF. In Multimedia information networking and security (MINES), 2010 international conference on (pp. 889-892). IEEE.

[10]Hailing Huang, Weiqiang Guo, Yu Zhang. (2008). Detection of Copy-Move Forgery in Digital Images Using SIFT Algorithm. 2008 IEEE Pacific-Asia Workshop on Computational Intelligence and Industrial Application, Vol.2, pp 272-276, December, 2008.

[11] Tien-Ying Kuo, Yi-Chung Lo, Ssu-Neng Huang, "Image forgery detection for region duplication tampering", *Multimedia and Expo (ICME) 2013 IEEE International Conference on*, pp. 1-6, 2013, ISSN 1945-7871.

[12] Xunyu Pan, Siwei Lyu, "Detecting image region duplication using SIFT features", *Acoustics Speech and Signal Processing (ICASSP) 2010 IEEE International Conference on*, pp. 1706-1709, 2010, ISSN 1520-6149.

[13] Tu K. Huynh, Khoa V. Huynh, Thuong Le-Tien, Sy C. Nguyen, "A survey on Image Forgery Detection techniques", *Computing & Communication Technologies - Research Innovation and Vision for the Future (RIVF) 2015 IEEE RIVF International Conference on*, pp. 71-76, 2015.

 [14] S. J. Ryu, M. J. Lee, and H. K. Lee, "Detection of copy-rotate-move forgery using Zernike moments," in Information Hiding. Berlin, Germany: Springer-Verlag, 2010, pp. 51– 65

[15] Abhishek Kashyap, Megha Agarwal, Hariom Gupta, "Detection of Copy-move Image forgery using SVD and Cuckoo Search Algorithm", arxiv.org, 2017

[16] Amanpreet Kaur, Richa Sharma, "Optimization of Copy-Move Forgery Detection Technique", International Journal of Advanced Research in Computer Science and Software Engineering, Volume 3, Issue 4, April 2013

[17] Sunil Kumar, Jagannath, Shaktidev Mukherjee, "A Fast DCT Based Method for Copy Move Forgery Detection", IEEE, 2013

[18] R. Nithiya, S. Veluchamy, "Key point descriptor based copy and move image forgery detection system", *Science Technology Engineering and Management (ICONSTEM) Second International Conference on*, pp. 577-581, 2016.

[19] Ye Zhu, Xuanjing Shen, Haipeng Chen, "Copy-move forgery detection based on scaled ORB", *Multimedia Tools and Applications*, pp. , 2015, ISSN 1380-7501.

[20] Anselmo Ferreira, Siovani C. Felipussi, Carlos Alfaro, Pablo Fonseca, John E. Vargas-Muñoz, Jefersson A. dos Santos, Anderson Rocha, "Behavior Knowledge Space-Based Fusion for Copy–Move Forgery Detection", *Image Processing IEEE Transactions on*, vol. 25, pp. 4729-4742, 2016, ISSN 1057-7149.