Image Forgery Detection

Muniganti Srija[1], Marabathuni Anuja[2], Polasa Pallavi[3], Pasula Tejasri[4], Mr. P. Rajesh Kumar[5]

[1-4]B.Tech Student, BVRIT HYDERABAD College of Engineering for Women, Telangana, India
[5]Assistant Professor, Dept. of Electronics & Communication Engineering, BVRIT HYDERABAD College of Engineering for Women, Telangana, India

---

Abstract:

With the advancement in technology, there are growing challenges in the authenticity and integrity of images, image manipulation has crumbled assurance over digital image. Advanced image handling software and altering tools are manipulating an image in such a way that it cannot be distinguished to the naked eye. The detection of forgery helps in crime scene investigation and in numerous different fields. Copy-move forgery is one of the widely used techniques for tampering the images. So, our project mainly focuses on the copy-move forgery detection. In order to find the forged regions in an image, we have proposed an algorithm for the detection of key points. This algorithm uses both block based and key points based approach. This could able to detect better key points when the image is subjected to rotation, scaling, noise. Feature matching rate is high which leads to the finding of accurate forged regions in an image. The execution of the proposed algorithm is less when compared to the existing algorithm when subject to rotation. Proposed algorithm could able to detect the accurate forged regions of an image.

Key Words: **Image Manipulation, forgery, copy-move forgery, Tampering, Feature Matching, block based approach, Key point based approach**

1. INTRODUCTION

Now-a-days there is a huge response to social networking services through which there is a high volume of image data generated. Use of image processing software's like Adobe Photoshop, Pixlir editor are used to create forgery images which is the major issue. These images are subjected to forgery using these algorithms like copy-paste, slicing and copy-move and resizing and the usage of clone stamp for copy and paste.

This forgery has become a major concern for the medical department. Many are trying to tamper the doctor certified image. The malpractice's in the medical imaging is the one of the major problem. Apart from copy-move image forgery there were many tampering techniques like image splicing, watermarking, Image retouching. Image splicing is combining and cropping two or more images into a single image. Image retouching is the adding or removing something from an image for enhancing features of an image. This technique is mainly used by magazine editors. Copy-move forgery copies and pastes a specific portion of the image. Since the copied region represents the same image, the dynamic range and color remains same. In this technique, we add or remove information to cover a part of an image.

2. RELATED WORKS

SURF follows the same principles and steps of SIFT but utilize different schemes for key point detection and description in such a manner that it provides better and faster results. Both algorithms construct scale-space using Gaussian kernels. But, unlike SIFT where the image I(x,y) is progressively down sampled within an octave, SURF uses larger and larger filters as the scale increases. To detect key points, it uses a Hessian matrix with the elements being the convolution of the image and the second order derivative of the Gaussian function.

$$H(x, \sigma) = [\ Lxx(x, \sigma)\ Lxy(x, \sigma)$$
$$Lxy(x, \sigma)\ Lyy(x, \sigma)\ ]$$

Where, $Lx(x, \sigma) = I(x, y) * \partial\ 2\ \partial x2\ g(\sigma)$

$$Lx(x, \sigma) = I(x, y) * \partial\ 2\ \partial xy\ g(\sigma)$$

But these convolutions are computationally complex and therefore to speed up the computations, integral images and approximated Gaussian kernels are used in convolutions. Now, with a key point at the center, a square region is considered which is oriented in the direction of dominant orientation. This square region is divided into 4x4 sub-regions, and for each sub-region, Haar wavelet response is found. Thus, from each sub-region, a four-dimensional descriptor vector $v = (\sum d, \sum dy, \sum|dx|, \sum|dy|)$ is derived. All such vectors of the 4x4 sub-regions are concatenated to form a 64-length descriptor vector.

3. METHODOLOGY

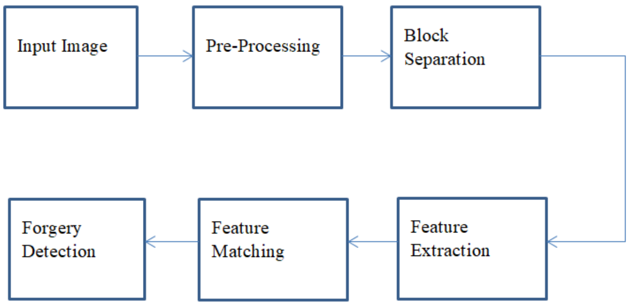

Fig1: Block Diagram

Input Image:

An image is forged using adobe photoshop and pixlir editor software tools.

a) Original Image b) Forged Image

Pre-Processing:

Image can be preprocessed that can be done by performing scaling, rotation or by adding noise and blurring of the image. The tampered image (RGB) will be taken as input and is converted into gray scale version using following formulae.

rgb2gray converts RGB values to grayscale values by forming a weighted sum of the *R*, *G*, and *B* components

$$I = 0.228R + 0.587G + 0.114B$$

Block Seperation:

The image is size MxN is divided into equal-sized(bxb) overlapped blocks based on the equation:

$$B=(M-b+1)x(N-b+1)$$

Feature Extraction:

In this we would find the key points of an image. Key points used are found using proposed algorithm as follows:

In this Feature Extraction, we have generated a new key descriptor generation of the SIFT Algorithm. Thus the proposed algorithm is the modification of the SIFT Algorithm. It follows four steps as follows:

Scale space extreme detection:

In this stage, potential key points are identified by constructing the scale space of an image. This scale space is created by the convolution of the Gaussian filter with the given image at different scales and then taking the difference of the successive Gaussian-blurred images.

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

$$G(x, y, \sigma) = 1\ 2\pi\sigma 2\ e\ -(x\ 2+y\ 2)\ 2\sigma\ /\ 2$$

$$D(x, y, \sigma) = L(x, y, k\sigma) - L(x, y, \sigma)$$

Key point Localization: A few key points needs to be discarded since not all the key points found in the first stage are used for feature description. Two criteria are used for the elimination of unused key points. The first criterion is to remove key points with low contrast value (intensities less than a specific threshold value) and the second criterion removes key points which are located on a flat or an edge region based on the gradients at each key point. Thus, key points located in corner and key points with high intensities are retained.

Key point orientation: The objective of this stage is assigning orientation to each key point by constructing a 36 bin histogram of gradient magnitudes m(x,y) and orientations θ(x,y) considering key point neighborhood of fixed size weighted by a Gaussian window.

$$(x, y) = \sqrt{(L(x+1, y) - L(x-1, y))\ 2 - (L(x, y+1) - L(x, y-1))\ 2}$$

$$(x, y) = tan{-1}\ ((L(x+1, y) - L(x-1, y))/(L(x, y+1) - L(x, y-1)))$$

Then each key point is assigned with θmax, the orientation of the bin with maximum magnitude. Usually, the range of orientations of θmax is $[-180^0, 180^0]$. The orientation is assigned to each key point by rotating the gradient image patch around the key point by an amount of θmax, to achieve image rotation invariance.

Key point Descriptor:

This project presents a calculation of the weighted sum of the features described in sub-generation algorithm using convolution. The neighborhood of the feature point is divided into multiple sub- neighborhoods. Then each statistical information for each sub-gradient neighborhood in the eight directions, To generate the feature descriptors. Given an image, first calculate the derivative images for each point on each of the eight directions.

$G_{oi} = \max(\partial I / \partial o_i, 0)$

Oi represents the direction, i =0,……7,Get the image derivative matrix on Eight directions. Statistics feature point to all the sub-neighborhood gradient information in the eight directions. calculated for each weights sub-matrix and gradient matrix convolution.

$$G_{w_{i,j}}^{k} = W_{i,j} * G_{o,k}$$

Where i = 1,..... ,4 j = 1,.... ,4 k = 1,.....8 ., $G_{w_{i,j}}^{k}$ is the gradient matrix after the original in the direction of k convolution with in the weight matrix w. $G_{w_{i,j}}^{k}$ (x, y) is a (x, y) centered 4x4 neighborhood, Each point in the direction of the gradient of the weighted sum of k. And now requires a weighted gradient, In the center of the feature point (u, v)in the16 sub-neighborhood in direction 8, Can be based only on the position of the sub-neighborhood Read the center of each sub-neighborhood of (xi, yi)values from the $G_{w_{i,j}}^{k}$.This avoids double-counting gradient weighted of feature point. Feature point descriptor vector elements consists of gradient weighted of feature points In the eight directions of its neighborhood 16 sub. After the calculation of $G_{w_{i,j}}^{k}$. The gradient weighted of the feature points in the Sub-neighborhood on Eight directions Can according to the center sub-neighborhood fast read out on the corresponding $G_{w_{i,j}}^{k}$. So you can read feature descriptor vectors obtained directly from the $G_{w_{i,j}}^{k}$.

Feature Matching:

Feature matching is calculated using Nearest Neighbor Distance ratio. The best candidate match for each key point is found by identifying its nearest neighbor in the database of key points from training images. The nearest neighbor is defined as the key point with minimum Euclidean distance for the invariant descriptor vector as was described in Section 6. However, many features from an image will not have any correct match in the training database because they arise from background clutter or were not detected in the training images. Therefore, it would be useful to have a way to discard features that do not have any good match to the database. A global threshold on distance to the closest feature does not perform well, as some descriptors are much more discriminative than others. A more effective measure is obtained by comparing the distance of the closest neighbor to that of the Second-closest neighbor. If there are multiple key points of the same object, then we define the second-closest neighbor as being the closest neighbor that is known to come from a different object than the first, such as by only using images known to contain different objects. This measure performs well because correct matches need to have the closest neighbor significantly closer than the closest incorrect match to achieve reliable matching.

Forgery Detection:

From the proposed algorithm, the matched features in a image are detected and the masking of the matched features will able to detect the forged regions. Proposed algorithm will able to detect the accurate regions when compared to the other algorithms.

4. CONCLUSIONS

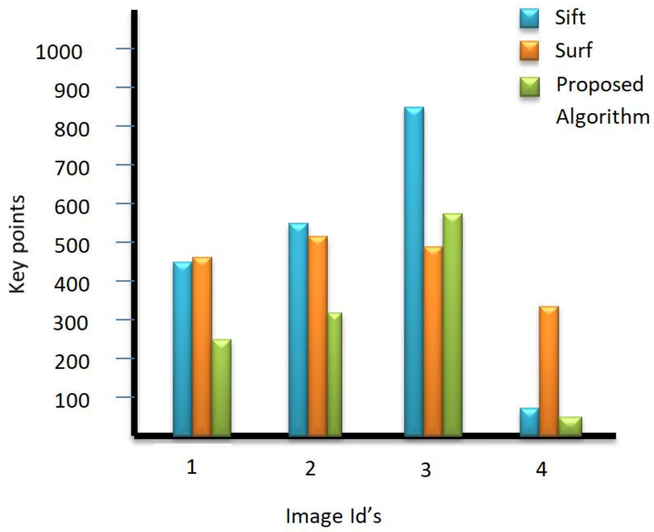

Fig2: No. of Key points when images are subjected to scaling

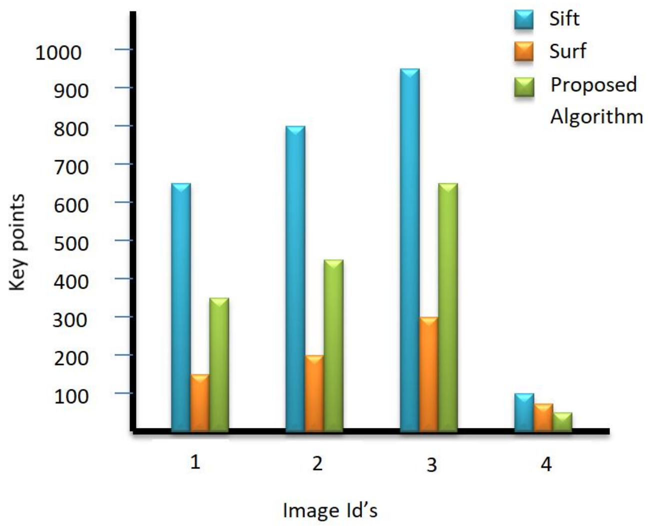

Fig 3: No. of Key points in an image

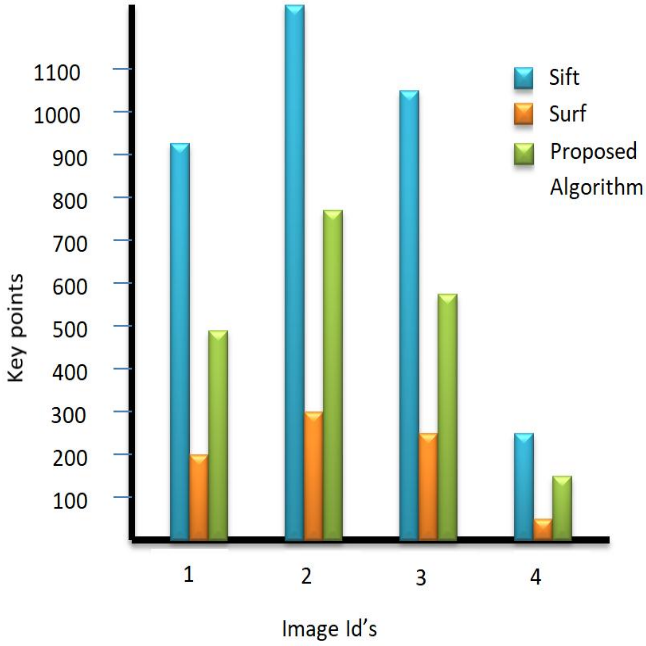

Fig4: No. of Key points when images are subjected to noise

Algorithm	Key points	Image-1	Image-2	Image-3	Image-4
SIFT	Original	8.78	12.16	7.10	7.75
	Resize	10.57	9.93	7.01	6.95
	Noise	14.62	12.29	11.23	17.79
SURF	Original	4.90	4.50	3.96	2.55
	Resize	11.57	11.82	7.88	2.33
	Noise	8.10	12.24	11.20	17.71
PROPOSED ALGORITHM	Original	7.59	9.64	6.43	7.78
	Resize	9.26	8.38	5.85	8.70
	Noise	11.83	12.29	12.21	17.79

Fig5 : Execution time of algorithms when subjected to noise,scaling.

Experimental results have shown that the proposed algorithm performs better in all cases. When the images are subjected to rotation,scaling and noise.

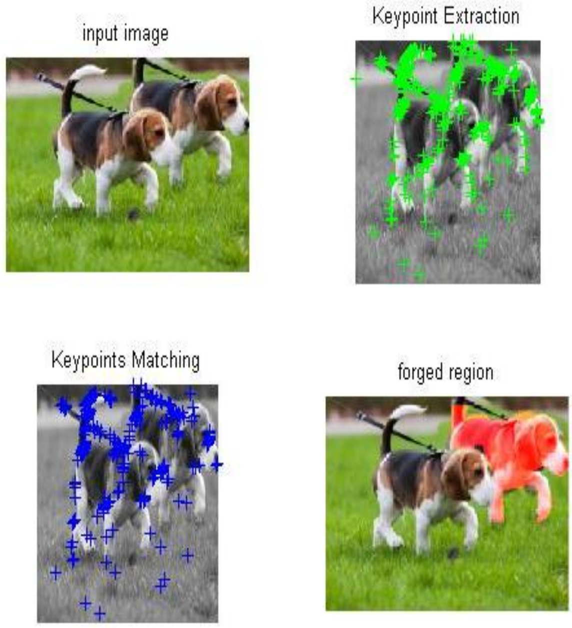

Fig6: Output of Image forgery Detection

ACKNOWLEDGEMENT:

We would like to express our deep sense of gratitude to our Principal Dr. K V N Sunitha, Head of the Department Dr. J. Naga Vishnu Vardhan and our guide Assistant Professor Mr. P. Rajesh Kumar

REFERENCES:

[1] R. N. Satare and S. R. Khot, "Image matching with SIFT feature," 2018 2nd International Conference on Inventive Systems and Control (ICISC), Coimbatore, 2018, pp. 384-387, doi: 10.1109/ICISC.2018.8399100.

[2] K. Abratkiewicz, D. Gromek, P. Samczyński, J. Markiewicz and W. Ostrowski, "The Concept of Applying a SIFT Algorithm and Orthophotomaps in SAR-based Augmented Integrity Navigation Systems," 2019 20th International Radar Symposium (IRS), Ulm, Germany, 2019, pp. 1-12, doi: 10.23919/IRS.2019.8768136.

[3] X. Zhou, K. Wang and J. Fu, "A Method of SIFT Simplifying and Matching Algorithm Improvement," 2016 International Conference on Industrial Informatics- Computing Technology, Intelligent Technology, Industrial Information Integration (ICIICII), Wuhan, 2016, pp. 73-77, doi: 10.1109/ICIICII.2016.0029

A Comparative Study of SIFT and its Variants Jian Wu1, Zhiming Cui1 , Victor S. Sheng2 , Pengpeng Zhao1, Dongliang Su1 , Shengrong Gong1 1 The Institute of Intelligent Information Processing and Application, Soochow University, Suzhou 215006, China, jianwu@suda.edu.cn 2 Department of Computer Science, University of Central Arkansas, Conway 72035, USA

International Research Journal of Engineering and Technology (IRJET) e-ISSN: 2395-0056
Volume: 07 Issue: 05 | May 2020 www.irjet.net p-ISSN: 2395-0072

BIOGRAPHIES:

Muniganti Srija is pursuing her B.Tech Final Year in Electronics & Communication Engineering Stream in BVRIT HYDERABAD College of Engineering for Women

Marabathuni Anuja is pursuing her B.Tech Final Year in Electronics & Communication Engineering Stream in BVRIT HYDERABAD College of Engineering for Women

Polasa Pallavi is pursuing her B.Tech Final Year in Electronics & Communication Engineering Stream in BVRIT HYDERABAD College of Engineering for Women

Pasula Tejasri is pursuing her B.Tech Final Year in Electronics & Communication Engineering Stream in BVRIT HYDERABAD College of Engineering for Women