A Study of Copy-Move Forgery Detection Scheme Based on Segmentation

Mohammed Ikhlayel[†], Mochamad Hariadi[††] and Ketut Eddy Pumama[†††]

[†]Department of Electrical Engineering Institut Teknologi Sepuluh Nopember, Surabaya, Indonesia
[†]Faculty of Technology and Applied Sciences, Al-Quds Open University, Palestine
[††]Department of Electrical Engineering, Institut Teknologi Sepuluh Nopember, Surabaya, Indonesia
[†††]Department of Computer Engineering Institut Teknologi Sepuluh Nopember, Surabaya, Indonesia

Summary
A copy-move forgery in digital image is a type of passive technique it will contain a part of the copied image and pasted to another parts in the same image. This may be occurring by a forger to cover part of object or validity or to enhance the visual effect in the image. Nowadays, there are many advance editing software in digital image are used to tampering, the forger can easily tamper the image, as a result, the image truth or validity is lost. In this study we will introduce three scheme for copy-move forgery detection(CMFD) based on segmentation and comparing between them, we will discuss the Segmentation-Based Image CMFD, then, adaptive Oversegmentation and Feature Point Matching, Finally, Multi-scale feature extraction and adaptive matching for CMFD. The results indicate the very good performance of each schemes.

Key words:
copy-move forgery detection, segmentation, Multi-Scale Feature Extraction. Adaptive Patch Matching.

1. Introduction

The modern technology in Multimedia and Availability of digital image is increasing nowadays. So digital images are Considered as the part of the main information source. In some cases, digital image is an important to proof the crime. In such cases, the digital content become an Evidence for the crime and the court of law to make a decision about it. However, the main problem that comes during to the wide availability of digital images are increasing digital tampering in digital image. Nowadays, many software is used in tampering, so the forgeries reduce the trustworthiness of the digital images. There are different types of forgery in digital image. Digital image forgery classified into active and passive techniques. For active technique, watermarking or signature, it's based on additional information embedded in the image used for forgery detection. While passive technique does not require additional information for forgery detection. There are many types of Passive technique, as Retouching, Splicing, Copy-Move Forgery CMF, etc.
Among the various types of digital image forgeries, copy-move is a common image tampering. A copy-move forgery in digital image contain some part of the image copied and pasted to another portion of the same image. Since source and target regions are same properties such as noise, illumination condition, color temperature etc. will be compared between source and target regions.

Fig. 1 the original image

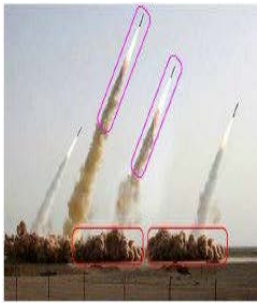

Fig. 2 the forged image.

The forgery maybe done to hide some object or authenticity or to enhance the visual effect of the image. By using image editing software such as Adobe Photoshop a forger can easily tamper the image and hide the tamper trace, thus the image authenticity is lost. Mostly the forgery occurs by making some geometric transformations such as rotation, scaling etc. The forger may hide the tampering by noise addition, lossy compression or blurring. The above operations are done to make copy-move forgery detection more difficult. So we need an effective CMFD method.
CMFD is a very important process in many areas such as medical imaging, criminal investigation, surveillance systems, transportation sector, scientific publications, intelligence services, financial document, etc. Figure 1. shows the original image. This image contains only three missiles. Figure 2. shows the forged image. This image contains four missiles. Copy-move source and target regions are shown in rectangle.
In the next section 2, we will introduce the Literature Survey of the copy-move forgery detection. In section 3, three types of copy-move forgery detection scheme based on segmentation are discuss. In section 4, introduce the

Manuscript received July 5, 2018
Manuscript revised July 20, 2018

comparative analysis. Finally, the conclusions are drawn in section 5.

2. Literature Survey

Block based approach: Block-based approach uses features extracted from the small blocks of the digital image. Fridrich et. al. [1] proposed a direct approach for CMFD by using an exhaustive search. So, this approach requires large amount of time for its processing. modifications such as scaling, rotation, translation etc. are done on the CMF. Luo et. al. [2] proposed a method based on color features of image. Bayram et. al. [3] proposed a method based on fourier transform. Both of their works can't detect forgery if noise or blur or other transformations are added to the image. Christlein et. Al. [4] method based on Same Affine Transform Selection (SATS) can't detect rotated CMF regions. Ardizzone et. al. [5] proposed a method based on texture of the image. M. Ghorbani et. al. [6] and S. Khan et. al. [7] uses discrete cosine transform and discrete wavelet transform (DWT) respecively for forgery detection. But their work can't detect forgery if any modifications such as rotation, scaling, etc. are done on the copy-move regions. Ryu et. al. [8] proposed a method based on zernike moments requires large amount of time for forgery detection.

Keypoint based approach: Keypoint-based approach uses the features extracted from the whole image. Methods of H. Huang et. Al. [9], Xunyu Pan et. al. [10], Amerini et. al. [11],[12]etc. uses SIFT features for forgery detection. But their methods can't detect noisy or blurred CMF regions and false detection was also high. The method of X. Bo et. al. [13] using SURF approach requires both original and forged image for CMF detection. Method of Kakar et al. [14] using gaussian filter can't detect multiple duplicated regions in the image.

Segmentation-based approach: The image has to be segmented into meaningful regions. Li et al. [15] tested four different image segmentation methods and used superpixels Simple Linear Iterative Clustering (SLIC) algorithm, to over-segment the images. They then extracted the SIFT features from each segment, built a k-d tree for them and used the KNN (k-nearest neighbor) to find the matching between patches. Li et al. [15] also used the SLIC method to segment the images. They used different sizes of segmentation depending on the image content itself. Bhanu and Kumar [16] used SLIC to segment the image into more than 100 patches, and extracted the SURF [17] from each patch. Sekhar and Shaji [18] presented a study taking a segmentation based approach and rotation invariant DAISY descriptors to detect copy-move forgery. In their study, they used three existing methods; they followed the same approach as [15] to segment the image. They proposed to use Adaptive Non-Maximal Suppression (ANMS) feature detection and DAISY descriptors [17] instead of SIFT. They found the matching between features using the generalization of Lowe's matching technique (g2NN) approach [11]. Since they did not implement their proposed method, there is no evidence whether or not their method would work. Bi et al. [19] segment the host image into non-overlapping irregular patches, considering that superpixel algorithms can group pixels into perceptually meaningful atomic regions, the author employ SLIC algorithm to segment the host image.

3. Copy-move forgery detection scheme based on segmentation

There are Many scheme have been developed to detect copy-move based on segmentation in digital image. In This study we will introduce three schemes, first Segmentation-Based Image copy-move forgery detection, then, Adaptive Oversegmentation and Feature Point Matching, finally, Multi-scale feature extraction and adaptive matching for copy-move forgery detection.

3.1 Segmentation-Based Image CMFD Scheme

In this scheme, as shown in Figure.3, the author in [20] proposed a scheme to detect the forgery in an image, he depend on extracting the keypoints for comparison. The main difference to the previous methods is that the proposed methods first segments the test image into semantically independent patches prior to keypoint extraction. and then, the forgery regions can be detected by matching between these patches. The matching process consists of two stages:

First Stage Matching: This stage Consists of three steps

1. Keypoint Extraction and Description

In this step the author employ the default setting of vlFeat for keypoints detection and description, namely SIFT [21]. Although the keypoint detection and description are not rather important, the number of the keypoints should be larger than 2000 for good performance.

2. Matching Between Patches

In this step we look for the suspicious pairs of patches So it contains many similar keypoints. This process is performed by comparing each patch with the other patches.

3. Affine Transform Estimation

when we find the detecting a suspicious pair of patches, we know where the copying source region and pasting target region are. Then we estimate the relationship between these two regions in terms of a transform matrix. In the second stage of matching process, where additional information of the digital image is employed to improve the accuracy of transform estimation.

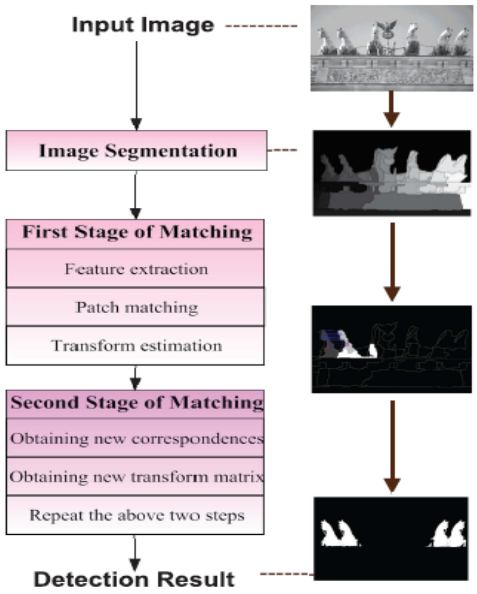

Fig. 3 Flowchart of the Segmentation-Based Image CMFD Scheme

Second Stage Matching: This stage generally speaking, they both consist of the following three steps.
1. Obtaining the matched points.
2. Calculating the transform matrix.
3. Repeating the above two steps until a convergence condition is satisfied.

3.2 Adaptive Oversegmentation and Feature Point Matching

In this study, as shown in Figure.4 the CMFD scheme using adaptive oversegmentation and feature point matching is presented by [22]. This scheme integrates both block-based and keypoint-based forgery detection methods. This scheme depends on four steps as follow:
1. segment the host image into nonoverlapping using the Adaptive Over-Segmentation algorithm and irregular blocks called Image Blocks (IB). the Adaptive Over-Segmentation method can divide the host image into blocks with adaptive initial sizes according to the given host images, with which each image can be determined to be an appropriate block initial size to enhance the forgery detection results. The proposed Adaptive Over-Segmentation method can lead to better forgery detection results compared with the forgery detection methods, which segment the host images into fixed-size blocks and, at the same time, reduce the computational expenses compared with most of the existing forgery detection methods, which segment the host images into overlapping blocks.

2. the Scale Invariant Feature Transform (SIFT) will apply in each block to extract the SIFT feature points as Block Features (BF). In this algorithm, the author chose SIFT as the feature point extraction method to extract the feature points from each image block, and each block is characterized by the SIFT feature points that were extracted in the corresponding block. Therefore, each block feature contains irregular block region information and the extracted SIFT feature points.

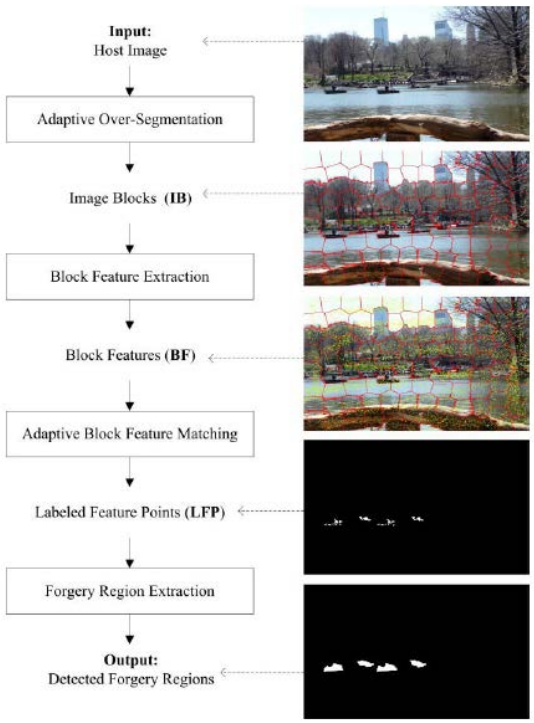

Fig. 4 Framework of the Adaptive Oversegmentation and Feature Point Matching scheme.

3. the block features are matched with one another, and the feature points that are successfully matched to one another are determined to be Labeled Feature Points (LFP), which can approximately indicate the suspected forgery regions. In this algorithm, because the block feature is composed of a set of feature points, the author proposed a different method to locate the matched blocks as follows:
STEP-1: Load the Block Features
STEP-2: Calculate the block matching threshold
STEP-3: Locate the matched blocks according to the block matching threshold.
STEP-4: Label the matched feature points in the matched blocks to indicate the suspected forgery regions.
4. we propose the Forgery Region Extraction method to detect the forgery region from the host image according to the extracted Labeled Feature Points (LFP). In this algorithm, a close morphological operation is applied to the merged regions to generate the detected copy-move

forgery regions then We can determine the Forgery Region Extraction. this algorithm, which is explained as follows:
STEP-1: Load the Labeled Feature Points
STEP-2: Measure the local color feature of the superpixels neighbor to the SR, called neighbor blocks
STEP-3: Apply the morphological close operation into MR to finally generate the detected forgery regions.

3.3 Multi-scale feature extraction and adaptive matching for CMFD

A copy-move forgery detection scheme by using multi-scale feature extraction and adaptive matching is proposed by [19]. As shown in Figure 5 The framework of the Multi-Scale Feature Extraction (MSFE) scheme discuss as follows:

First, the host image is segmented into the non overlapping patches of irregular shape in different scales. Most of the existing block based algorithms in previous studies, divided the host images only in single scale with initially predefined block size; in that situation, if block size is too small, some forgery regions will be missed. To solve this problem, the author proposed the multi-scale segmentation in this algorithm. the proposed MSFE algorithm segments the digital image into the patches with multiple scales; from each patch the feature points are then extracted, then, the author chose SIFT with default parameters as the feature extraction method to extract feature points as patch features.

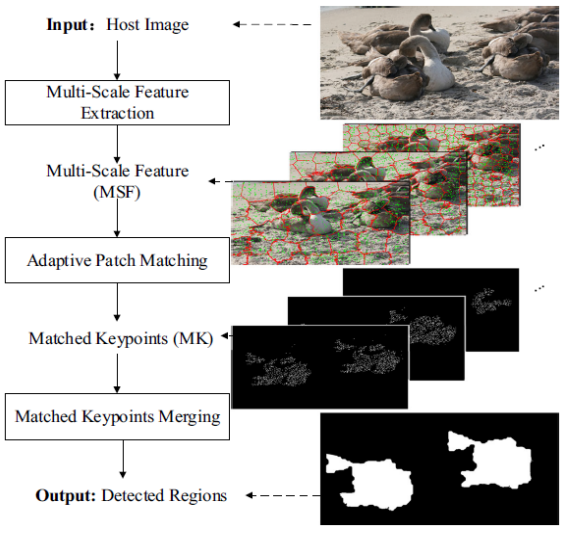

Fig. 5 framework of the Multi-Scale Feature Extraction scheme.

Figure 6 shows the flowchart of the MSFE algorithm. First, the host image is blocked into the patches with the superpixel segmentation method; then, the feature points are extracted from these patches. The whole process is repeated along with the decreasing of the size of segmentation, until feature points cannot be extracted any more in the corresponding scale. Finally, the multi-scale feature MSF is generated, which includes the patches in each scale and the corresponding feature points.

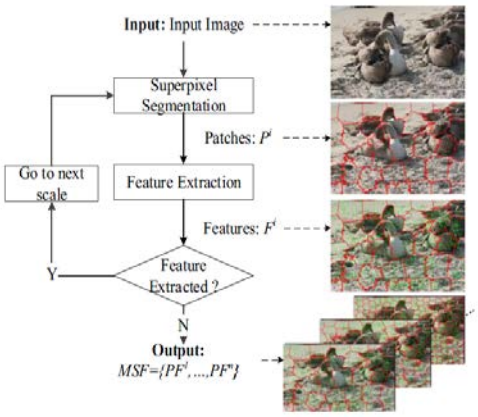

Fig. 6 Flowchart of the Multi-Scale Feature Extraction (MSFE) algorithm

In this study, the author employ the Simple Linear Iterative Clustering (SLIC) algorithm [23] to segment the host image. SLIC algorithm adapts a k-means clustering approach to efficiently generate superpixels, and superpixel adheres to boundaries very well. With SLIC, the host image is segmented into the non-overlapping superpixels that are meaningful and are of irregular shapes. The non-overlapping segmentation method can help to decrease the computational expenses, compared with the existing overlapping block method; in addition, in most of the cases, the irregular and meaningful regions can represent the forgery regions better than the regular blocks.

Second, Scale Invariant Feature Transform is applied to extract feature points from all patches, Figure.7 shows the Flowchart of the Multi-Scale Feature Extraction (MSFE) algorithm, to generate the multi-scale features. an Adaptive Patch Matching algorithm is subsequently proposed for finding the matching that indicate the suspicious forged regions in each scale.

Then, after generating MSF, we need to locate the matched patch pairs in each scale. In most of the existing block based algorithms, the block matching generates specific block pairs only if there are many other matched pairs in the same mutual position, assuming they have the same shift vector. When the number of matched block pairs, which have same shift vector, exceeds a user specified threshold, the matched block pairs that contribute to that specific shift vector will be identified as the regions that probably have been tampered. In that situation, the threshold is related to the regions that can be identified; the larger threshold may cause some not-so-closely matched blocks missing, while the smaller threshold may bring

more false matched blocks. Therefore, the threshold highly relates with the performance of the forgery detection algorithms, and how to determine the just right threshold becomes an important issue.

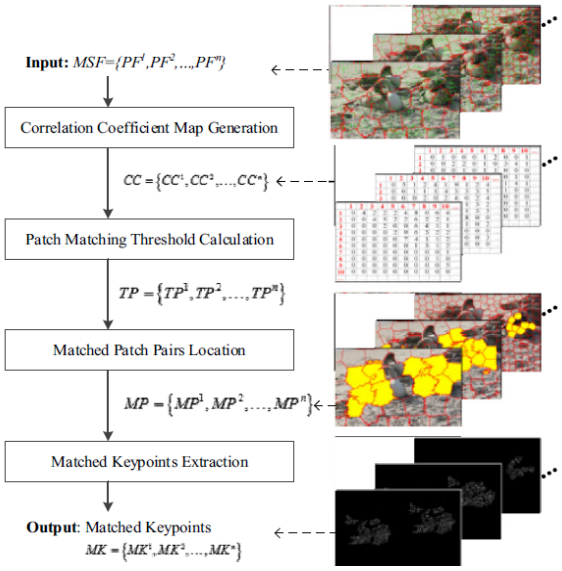

Fig. 7 Flowchart of the Multi-Scale Feature Extraction (MSFE) algorithm

Third, then finally, the suspicious regions in all scales are merged to generate the detected forgery regions. After obtaining the matched keypoints MK, we need to determine the forgery regions by turning the independent pixels/keypoints into regions. Figure 8 shows the flowchart of the MKM algorithm.

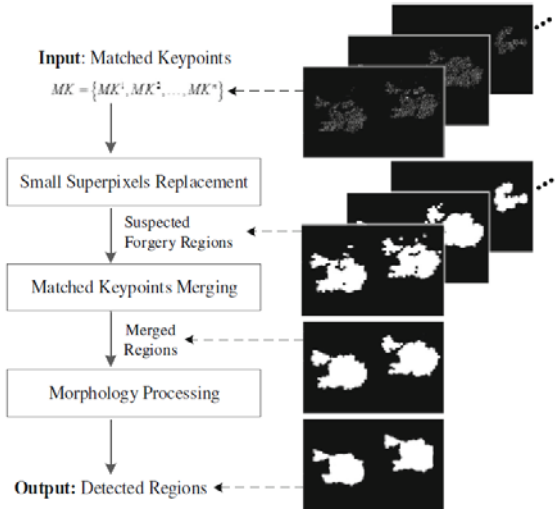

Fig. 8 shows the flowchart of the APM algorithm.

4. Comparative Analysis

In Tables 1 and 2, It can be easily seen that the MSFE scheme can achieve precision and recall, which performs better than the most of existing state-of-the-art methods at image level and pixel level.

Table 1: Detection results of the plain CMF at the image level

Image level	Precision (%)	recall (%)	F (%)
SBFD [20]	70.16	83.33	76.18
ASFPM[22]	96	100	97.96
MSFE [19]	90.57	100	95.05

Table 2: Detection results of the plain CMF at the pixel level

Pixel level	Precision (%)	recall (%)	F (%)
SBFD [20]	84.90	54.095	65.16
ASFPM[22]	89.195	83.73	86.38
MSFE [19]	95.22	90.6	92.85

5. Conclusions

Replace the overlapping blocks of regular shape in traditional forgery detection algorithms, with individual irregular patches, which can better partition the host images into non-overlapping blocks. integrates both block-based and keypoint-based forgery detection methods. Then the best scheme is the multi-scale feature extraction method because it can extract more accurate feature points. using the predefined small superpixels to replace the matched keypoints and we apply some morphology operations into the merged regions to generate more accurately detected forgery regions.

References

[1] A. J. Fridrich, B. D. Soukal, and A. J. Lukáš, "Detection of copy-move forgery in digital images," in in Proceedings of Digital Forensic Research Workshop, 2003.

[2] W. Luo, J. Huang, and G. Qiu, "Robust detection of region-duplication forgery in digital image," in Pattern Recognition, 2006. ICPR 2006. 18th International Conference on, 2006, pp. 746-749.

[3] S. Bayram, H. T. Sencar, and N. Memon, "An efficient and robust method for detecting copy-move forgery," in Acoustics, Speech and Signal Processing, 2009. ICASSP 2009. IEEE International Conference on, 2009, pp. 1053-1056.

[4] V. Christlein, C. Riess, and E. Angelopoulou, "On rotation invariance in copy-move forgery detection," in Information Forensics and Security (WIFS), 2010 IEEE International Workshop on, 2010, pp. 1-6.

[5] E. Ardizzone, A. Bruno, and G. Mazzola, "Copy-move forgery detection via texture description," in Proceedings of the 2nd ACM workshop on Multimedia in forensics, security and intelligence, 2010, pp. 59-64.

[6] M. Ghorbani, M. Firouzmand, and A. Faraahi, "DWT-DCT (QCD) based copy-move image forgery detection," in Systems, Signals and Image Processing (IWSSIP), 2011 18th International Conference on, 2011, pp. 1-4.

[7] S. Khan and A. Kulkarni, "Detection of copy-move forgery using multiresolution characteristic of discrete wavelet transform," in Proceedings of the International Conference & Workshop on Emerging Trends in Technology, 2011, pp. 127-131.
[8] S.-J. Ryu, M. Kirchner, M.-J. Lee, and H.-K. Lee, "Rotation invariant localization of duplicated image regions based on Zernike moments," IEEE Transactions on Information Forensics and Security, vol. 8, pp. 1355-1370, 2013.
[9] H. Huang, W. Guo, and Y. Zhang, "Detection of copy-move forgery in digital images using SIFT algorithm," in Computational Intelligence and Industrial Application, 2008. PACIIA'08. Pacific-Asia Workshop on, 2008, pp. 272-276.
[10] X. Pan and S. Lyu, "Region duplication detection using image feature matching," IEEE Transactions on Information Forensics and Security, vol. 5, pp. 857-867, 2010.
[11] I. Amerini, L. Ballan, R. Caldelli, A. Del Bimbo, and G. Serra, "A sift-based forensic method for copy–move attack detection and transformation recovery," IEEE Transactions on Information Forensics and Security, vol. 6, pp. 1099-1110, 2011.
[12] I. Amerini, L. Ballan, R. Caldelli, A. Del Bimbo, L. Del Tongo, and G. Serra, "Copy-move forgery detection and localization by means of robust clustering with J-Linkage," Signal Processing: Image Communication, vol. 28, pp. 659-669, 2013.
[13] X. Bo, W. Junwen, L. Guangjie, and D. Yuewei, "Image copy-move forgery detection based on SURF," in Multimedia information networking and security (MINES), 2010 international conference on, 2010, pp. 889-892.
[14] P. Kakar and N. Sudha, "Exposing postprocessed copy–paste forgeries through transform-invariant features," IEEE Transactions on Information Forensics and Security, vol. 7, pp. 1018-1028, 2012.
[15] K. Minakshi, "Digital Image Processing: In: Satellite Remote Sensing and GIS Applications in Agricultural Meteorology," World Meteorological Organization Publishing, pp. 81-102, 2003.
[16] B. B. MP and A. Kumar, "Copy-move forgery detection using segmentation," in Intelligent Systems and Control (ISCO), 2017 11th International Conference on, 2017, pp. 224-228.
[17] A. Swaminathan, M. Wu, and K. R. Liu, "Digital image forensics via intrinsic fingerprints," IEEE Transactions on Information Forensics and Security, vol. 3, pp. 101-117, 2008.
[18] R. Sekhar and R. Shaji, "A study on segmentation-based copy-move forgery detection using DAISY descriptor," in Proceedings of the International Conference on Soft Computing Systems, 2016, pp. 223-233.
[19] X. Bi, C.-M. Pun, and X.-C. Yuan, "Multi-scale feature extraction and adaptive matching for copy-move forgery detection," Multimedia Tools and Applications, vol. 77, pp. 363-385, 2018.
[20] J. Li, X. Li, B. Yang, and X. Sun, "Segmentation-based image copy-move forgery detection scheme," IEEE Transactions on Information Forensics and Security, vol. 10, pp. 507-518, 2015.
[21] D. G. Lowe, "Distinctive image features from scale-invariant keypoints," International journal of computer vision, vol. 60, pp. 91-110, 2004.
[22] C.-M. Pun, X.-C. Yuan, and X.-L. Bi, "Image forgery detection using adaptive oversegmentation and feature point matching," IEEE Transactions on Information Forensics and Security, vol. 10, pp. 1705-1716, 2015.
[23] R. Achanta, A. Shaji, K. Smith, A. Lucchi, P. Fua, and S. Süsstrunk, "SLIC superpixels compared to state-of-the-art superpixel methods," IEEE transactions on pattern analysis and machine intelligence, vol. 34, pp. 2274-2282, 2012.

Mohammed Ikhlayel is an Faculty member at Al-Quds Open University (QOU). He was Action Head of Department of Information Technology and Communications from 2009 until 2012, in QOU. Mr. Ikhlayel received his Bachelor degree in electrical engineer (Communications) in 2004, he received his master degree in Communications in 2007 from Egypt, and he Studies PhD in electrical engineering at Institut Teknologi Sepuluh Nopember in indonisia. (e-mail: missaikhlayel8080@gmail.com)

Mochamad Hariadi received the B.E. degree in Electrical Engineering Department from Institut Teknologi Sepuluh Nopember, Surabaya, Indonesia, in 1995. He received both M.Sc. and Ph.D. degrees in Graduate School of Information Science Tohoku University Japan, in 2003 and 2006 respectively. Currently, he is a staff of Electrical Engineering Department Institut Teknologi Sepuluh Nopember, Surabaya, Indonesia. He is the project leader in joint research with PREDICT JICA project Japan.(e-mail: mochar@te.its.ac.id).

Ketut Eddy Pumama, He is the Head of Computer Engineering Department Institut Teknologi Sepuluh Nopember, Surabaya, Currently, he is a staff of Computer Engineering Department Institut Teknologi Sepuluh Nopember, Surabaya, Indonesia. (e-mail: ketut@te.its.ac.id).