Detection of Image Forgery Using Information Standard Method With SVM

Alyaa Hashem Mohammed 1*, Duraid Hussein Badr 2, Fathala Ali 3
1,3 Collage of administration and Economics, Mustansiriyah University
2 Collage of administration and Economics / University of Basrah.

Email: aleaa_hashim@uomustansiriyah.edu.iq

Abstract. Depending on various features, image forgery is now used to enhance and restore images to make them more significant, whereas image forgery is performed by vulnerable people to generate fake truths by tampering images. From large (image) databases, a specified number of images forgery, new mechanism will have implemented to detect systems which is mainly based on two procedures. The first procedure relies on extract features of image through filters (LBP, Gabor, PCA and DCT) to find out its efficiency evaluated using some Statistics Measurements (Mean, Variance) and the second procedure relies on several quality metrics in optimization processes which (PSNR, MSE, AIC) have been calculated to which results in better quality compared to other image by combining selected filters with SVM, and, through experimental test, it was found that this proposed method of detect forgery technique is powerful than the classical detect system. The proposed methodology gives the better accuracy up to 98.5%, to detect and extract all type of forgery, with image enhancement capabilities.

Keywords: SVM classifier, quality metric, statistics measurements

1. Introduction

The contemporary world is witnessing easy access to information, way to process it and methods of storing and sharing it with easy to use device and available to everyone. Recently, using image editing tools to create forged images in an effortless manner, the growing usage of internet the and the propagation of digital image has become easy to do and hard to detect although image forgery has become challenging research area making us to lose the trust in legitimacy of the images. In recent years, many researchers have started to focus on the problem of forgery image forgery, and various methods have been developed to combat tampering and forgery in order to ensure the authenticity of images[10].
2. Theoretical side

2.1. Necessity of Forgery Detection of Digital Image
Image editing software tools will increase in fake images leads to a decrease in confidence in visual media, as well as the ease of simulating the origin and content of digital visual information as long as its confidence has been questioned methods of detecting forgery have emerged due to the significant impact of image processing on the medical profession, justice, and news reporting process as well as accounting professions. Detecting counterfeiting techniques are aimed at exposing inconsistent patterns that are supposed to be exist in the images due to manipulation in order to forge the image. The active and passive methods are used to detect fraud in image. [10]

2.2. Forgery Types in Image
The digital image forgery means the intentional tamper in the digital image, which aims to alter the semantic meaning of the image. There are 3 major kinds of image forgery: [1]

3. Applied side

a. Extraction of features from images
Feature extraction is depend on rebuilt statistical models applied quality metrics and statistics measurements on the (LBP, Gabor, PCA and DCT) filters, avoiding redundancy and reducing the dimensionality of the data are two requirements for features. [8]

\[\text{Fig1. Block diagram of the proposed system.}\]

i. LBP
Local binary patterns is most powerful and effective texture representation method. In this feature extraction technique, every pixel is represented in terms of a gray-scale value by considering the intensity value of adjacent pixels, operator combines features of statistical and fundamental texture analysis, using circular neighborhoods and linearly interpolation [8]

ii. Gabor Filter
Gabor filters were employed to extract features in many image forgery [1] Gabor filters have usually been used for texture segmentation and classification, are composed of a Low pass filter and a band pass filter [6]

iii. PCA
This mechanism Principal Component Analysis is a type of feature extraction process which deals in bringing out strong patterns in a dataset [2] PCA is perhaps the most widely-adopted multivariate statistical approach[3]

iv. DCT
we apply DCT so that we detect forgery on jpeg image too. After that we compare both the approaches and find out the results and compare the results[4], Blocking Matrix is to extracted form DCT blocks of the image. It exposes important structures roughly the image compression
history. To execute grey scale image is subdivided into sub blocks of 8x8 pixels. For each sub-
block and every pixel location the inter-pixel strength difference is calculated. [2]

b. Some Statistics Measurements for digital image description
One of the best vital tasks of image analysis is to determine the necessary information extracted
for the purpose of statistical data analysis, and image analysis includes the Statistics Parameters
will be used in this research to measure the image efficiency: of the image into two types.

3.1 Mean Of Image:
it is the average value image pixels intensity, which has a brightness range 0-255, so when it is
approximated to zero then image will close to be black, but when it is approximated to 255 that
will indicate the image is bright, however this could be explored using equation (1) in the
following:

$$\text{Mean} = \frac{1}{m \times n} \sum_{(m,n)}$$ (1)

3.2 Variance of image (Variance)
It is the degree of image contrast, so if the image variance value is low that will indicates better
image and therefor, few changes will appear in image intensity and small signtens whereas,
the image of high variance that will indicates high changes in image intensity and big signftens,
so this could be formulated using equation: [5]

$$\text{Variance} = \frac{\sum_{i=1}^{N} x_i^2}{N} - \text{Mean}^2$$ (2)

Table 1. Refer to filters (LBP, Gabor, PCA and DCT) to original image, Tamperd QF100
Tamperd QF80

|   | Original Image | Tamperd QF100 | Tamperd QF80 |
|---|----------------|---------------|--------------|
| A | LBP            |               |              |
| B | Gabor          |               |              |
| C | PCA            |               |              |
| D | DCT            |               |              |
Through studying the statistical Measurements of the image, it was found that the mean and variance of image (A) was (119.3945, 65.36921), to Tamperd QF100 was (142.306, 66.7944) and to Tamperd QF80 (148.4336, 60.26064) respectively, while the mean and variance of image (B) was (142.0645, 70.641), to Tamperd QF100 was (99.6679, 66.7944) and to Tamperd QF80 (100.6636, 67.79447), image (C) was (145.3445, 62.36921), to Tamperd QF100 was (144.3945, 64.29924) and to Tamperd QF80 was (142.3446, 60.36064), image (D) was (154.4545, 66.40921), to Tamperd QF100 was (142.006, 70.799) and to Tamperd QF80 was (48.6336, 69.354).

3.3 Quality metrics
We focused on 4 image quality measures are considered[10], Image quality metrics are used for evaluate the imaging systems and processing techniques. There are many varieties of these metrics, some of them[7]:

PSNR metric:
One of the most common metric is Peak Signal-to-Noise Ratio (PSNR), Equation (1) denote it. Increasing this value indicates the high quality of the image formula is: [10]

$$PSNR = 10 \log_{10} \left[ \frac{255^2}{M \sum \sum (r_{ij} - x_{ij})^2} \right]$$ ...(3)

Where: $r_{ij}$: are the element values for the reconstructed image., $x_{ij}$: are the element values for the original image.

MSE metric:
Equation (4) shows mathematically the Mean Square Error (MSE), the MSE measure is used to compare between two images by describing the degree of similarity between them it is assumed that one of the images is original and the other is reconstructed, the difference of pixel color in the two images must be known. A decrease in this value indicates a high quality [1, 3].

$$MSE = \frac{1}{MN} \sum_{ij} (r_{ij} - x_{ij})^2$$ ...(4)

Where: $r_{ij}$: are the element values for the processed image., $x_{ij}$: are the element values for the original image.

AIC metric:
Akaike's information criterion (AIC) is used for find best model, This can be expressed by the following equation:

$$AIC = N \ln \left( \frac{SSe}{N} \right) + 2q$$ ...(5)

Where: $SSe$: Sum of error squares, $q$: Number of parameters for models

N: Number of views (image size), Accordingly, the lowest AIC value represents the best model.

Where the model that corresponds to its lowest value is chosen. [9]

SSIM metric”
Structural Similarity Image Measure(SSIM) is depend on light, contrast and Structural Similarity[9]:

$$l(x, y) = \frac{2\mu_x(x, y)\mu_y(x, y) + C_1}{\mu_x^2(x, y) + \mu_y^2(x, y) + C_1}$$

$$c(x, y) = \frac{2\sigma_x(x, y)\sigma_y(x, y) + C_2}{\sigma_x^2(x, y) + \sigma_y^2(x, y) + C_2}$$
\[ s(x, y) = \frac{\sigma_{xy}(x, y)}{\sigma_x(x, y) \sigma_y(x, y)} + C_3 \]

\[ SSIM(x,y) = [l(x,y), c(x,y), s(x,y)] \]

Where: \( l(x,y) \) is lighting, \( c(x,y) \) is contrast, \( s(x,y) \) is Structural Similarity.

**Table 2. Quality metrics (PSNR, MSE, AIC, SSIM)**

|   | PSNR   | MSE    | AIC    | SSIM   |
|---|--------|--------|--------|--------|
| A | Original Image | 1946.269 | 1954.07 | 1949.07 | 1934.125 |
|   | Tampered QF100 | 1907.24  | 1912.64 | 1897.04 | 1903.238 |
|   | Tampered QF80  | 1934.157 | 1924.05 | 1934.06 | 1924.112 |
| B | Original Image | 1944.269 | 1950.07 | 1952.07 | 1944.125 |
|   | Tampered QF100 | 1901.24  | 1910.64 | 1890.04 | 1900.238 |
|   | Tampered QF80  | 1924.157 | 1944.05 | 1935.06 | 1922.112 |
| C | Original Image | 1945.269 | 1952.07 | 1955.07 | 1954.125 |
|   | Tampered QF100 | 1902.24  | 1914.64 | 1894.04 | 1902.238 |
|   | Tampered QF80  | 1934.157 | 1948.05 | 1937.06 | 1926.112 |
| D | Original Image | 1947.269 | 1955.07 | 1945.07 | 1924.125 |
|   | Tampered QF100 | 1905.24  | 1911.64 | 1888.04 | 1901.238 |
|   | Tampered QF80  | 1903.157 | 1935.05 | 1945.06 | 1936.112 |

To depict quality of images (original, fake) for extent it has the heights quality is have lowest standard of information.

PSNR\(_A2\) < PSNR\(_A3\) < PSNR\(_A1\), PSNR\(_B2\) < PSNR\(_B3\) < PSNR\(_B1\), PSNR\(_C2\) < PSNR\(_C3\) < PSNR\(_C1\),

MSE\(_A2\) < MSE\(_A3\) < MSE\(_A1\), MSE\(_B2\) < MSE\(_B3\) < MSE\(_B1\), MSE\(_C2\) < MSE\(_C3\), MSE\(_C1\), MSE\(_D2\) < MSE\(_D3\) < MSE\(_D1\),

AIC\(_A2\) < AIC\(_A3\) < AIC\(_A1\), AIC\(_B2\) < AIC\(_B3\) < AIC\(_B1\), AIC\(_C2\) < AIC\(_C3\) < AIC\(_C1\), AIC\(_D2\) < AIC\(_D3\) < AIC\(_D1\),

SSIM\(_A2\) < SSIM\(_A3\) < SSIM\(_A1\), SSIM\(_B2\) < SSIM\(_B3\) < SSIM\(_B1\), SSIM\(_C2\) < SSIM\(_C3\) < SSIM\(_C1\),

SSIM\(_D2\) < SSIM\(_D3\) < SSIM\(_D1\),

### 3.4 SVM Classifier

Support vector machine (SVM) classifier is a controlled learning method to classification [5] SVM decide decision boundaries in training step and also provide better generalization in high dimensional input spaces feature space making classification problem. In experiment, we use an adaptive support vector machine for the classification. This system is implemented using Matlab 2014a [10] it is process of feature extraction main and essential step in SVM. To approve classification process and give the final classification.

![Fig. 2. Classifier accuracy to images (“small and large spliced area”).](image-url)
4. Calculate the correct rating percentage
Detecting forgery output based on SVM classifier, the statistical Measurement, Quality metric measure, set S1, S2, and S3 which represent the Original images set, Tampered QF100 images set, and Tampered QF80 images set, respectively, the SVM classifier trained with these measures from each S1, S2, and S3 set, 12 images selected to train the SVM classifier. Then, rest of image tested correctness to SVM classifies, different experiments tested (1) trained SVM classifier and tested with the data from S1, S2 sets. (2) the SVM classifiers is trained and tested with the data from S1 and S2 sets. (3) the SVM classifier is trained and tested with the data from S1 sets and the group of S2 and S3 sets. The results are summarized in Table (S1 and S2 Set), (S1 and S3 Sets) and (S1 and (S2 & S3) Sets) respectively. Without grouping two sets, S2 and S3, used to train the SVM classifies in same time, the classification ability became low. false positive rate which represents deciding the faked image increased at 8.3%. despite of, false negative rate deciding the copy of faked image was 0%. S1, S2, and S3 represents the results for the original set image S1 set and the Tampered Qf80 S3 set, the SVM classifier identified. The overall classification accuracy was greater than 95%. [5]

| Name of Set | Training Set | Testing Set | S1 | S2 |
|-------------|--------------|-------------|----|----|
| S1 and S2 set | S1 | 100% | 0% |
| | S2 | 0% | 100% |
| S1 and S3 sets | S1 | 100% | 0% |
| | S3 | 0% | 100% |
| S1 and (S2 & S3) sets | S1 | 100% | 0% |
| | S2 & S3 | 8.3% | 91.7% |
| S1, S2, and S3 | S1 | 100% | 0% |
| | S2 | 20% | 80% |
| | S3 | 0% | 100% |

5. Conclusion
This paper presented the fraud image identification scheme using (Statistics Measurements for digital image, 4 image quality measures) and our forgery detector based on the SVM classifier. Even though we use 6 features, Tables show misclassification when we tested with several candidates. It means that selection of proper features is the key to achieve accurate results. In future works, we expect to improve the performance of the presented image forgery detection method by designing more suitable features.
6. References

[1] M. S. Mahdi and S. N. Alsaad, “Detection of Copy-Move Forgery in Digital Image Based on SIFT Features and Automatic Matching Thresholds,” in International Conference on Applied Computing to Support Industry: Innovation and Technology, 2019, pp. 17–31.

[2] A. Gupta, N. Saxena, and S. K. Vasistha, “Detecting copy move forgery using DCT,” Int. J. Sci. Res. Publ., vol. 3, no. 5, p. 1, 2013.

[3] S.-J. Ryu, H.-Y. Lee, I.-W. Cho, and H.-K. Lee, “Document forgery detection with SVM classifier and image quality measures,” in Pacific-Rim Conference on Multimedia, 2008, pp. 486–495.

[4] D. Bansal and S. Kaushal, “A novel Analysis of Image Forgery Detection Using SVM,” Int. J. Eng. Appl. Sci., vol. 3, no. 12.

[5] V. Sharma, S. Jha, and R. K. Bharti, “Image forgery and it’s detection technique: a review,” Int. Res. J. Eng. Technol., vol. 3, no. 3, pp. 756–762, 2016.

[6] V. Laxmi and P. S. Rao, “Eye detection using Gabor Filter and SVM,” in 2012 12th International Conference on Intelligent Systems Design and Applications (ISDA), 2012, pp. 880–883.

[7] H. M. Ahmed and S. R. Salim, “Method for Detect and Extract Forgery from Images,” AL-MANSOUR J., no. 23, pp. 37–60, 2015.

[8] T. M. Mohammed et al., “Boosting image forgery detection using resampling features and copy-move analysis,” Electron. Imaging, vol. 2018, no. 7, pp. 111–118, 2018.

[9] C. Deep Kaur and N. Kanwal, “An analysis of image forgery detection techniques,” Stat. Optim. Inf. Comput., vol. 7, no. 2, pp. 486–500, 2019.

[10] B. Duraid Hussein, “Comparison of some filters to remove noise for digital images using some statistical methods” Basic Education Journal, 2020.