

Image tampering detection using extreme learning machine

Cite as: AIP Conference Proceedings **2457**, 040002 (2023); <https://doi.org/10.1063/5.0123415>
Published Online: 02 February 2023

Dalia S. Sulaiman and Mohammed Sahib Mahdi Altaei



View Online



Export Citation

ARTICLES YOU MAY BE INTERESTED IN

[Using survival function and transmuted formula to produce lifetime models with application on real data set](#)

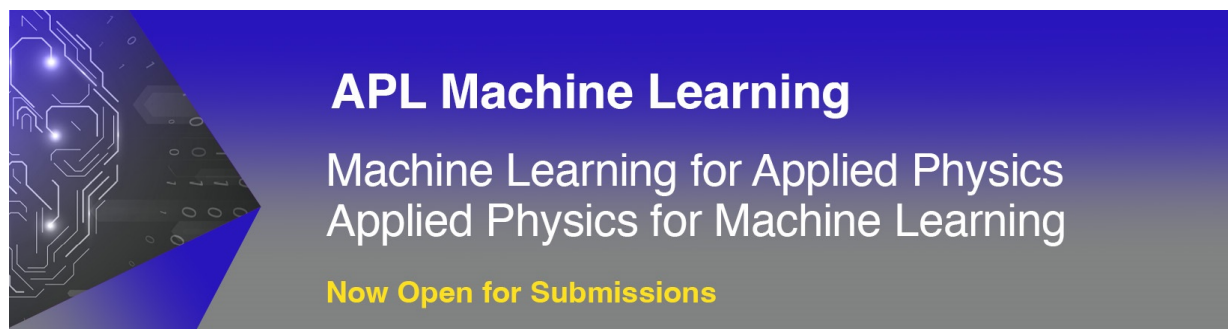
AIP Conference Proceedings **2457**, 020017 (2023); <https://doi.org/10.1063/5.0120114>

[Development of spectrophotometric method for the determination of metoclopramide hydrochloride in pharmaceutical preparations](#)

AIP Conference Proceedings **2457**, 030002 (2023); <https://doi.org/10.1063/5.0118703>

[Designed new mesogenic containing 5H-thiazolo\[3,4-b\]\[1,3,4\]thiadiazole: Synthesis and investigation of liquid crystals properties](#)

AIP Conference Proceedings **2457**, 030003 (2023); <https://doi.org/10.1063/5.0120896>



APL Machine Learning
Machine Learning for Applied Physics
Applied Physics for Machine Learning
Now Open for Submissions

Image Tampering Detection Using Extreme Learning Machine

Dalia S. Sulaiman ^{a)} and Mohammed Sahib Mahdi Altaei ^{b)}

Computer Science Department, College of Science, Al-Nahrain University, Baghdad, Iraq

^{a)} Corresponding author: st.dalia.sufiyan@ced.nahrainuniv.edu.iq

^{b)} Corresponding author: altaeimohamed@gmail.com

Abstract. The use of textural descriptors were useful in the present work, which are calculated based on the Haralick features for image tempering using the Extreme Learning Machine (ELM). This method was applied to color images after converting it into YC_bC_r color system, and then the image is divided into blocks in order to apply the Local Binary Pattern (LBP) on each block belongs to each resulted color band. The textural features are then computed and encoded for the target image to be stored in a database file for that reconstructed image. The computed features enter the ELM classifier to carry out the processes of the training and classification. The training was performed on CASSIA-II dataset while testing was performed on CASSIA-I. The classification results gave a test accuracy of tempering detection about 99.7% when using the Y -band, 99.7% when using the C_b band, and 99.4% when using the C_r band. Whereas, the evaluation of the test results was good compared to previous work, this confirms the validity of the results and ensure the correct path of the proposed method.

Keywords: Image Tempering, Tempering Detection, ELM, GLCM, LBP, Haralick

INTRODUCTION

The change occur in digital images by automatic software is nowadays an activity of simplicity with very low cost, so any person can synthesize a fake image. The incorrect evidence transmits incredibly easily to the freely available Internet. As a consequence, it is easy to misrepresent the reality and affect popular sentiment, resulting in a negative social impact. In the justice system, things could be much worse [1]. Digital images can be manipulated to such perfection today that forgery can not be visually identified. The security issue of digital content appeared a long time ago and numerous approaches have been developed to verify the integrity of digital images. One tampers pictures to enjoy visual pleasure for different reasons. It works to create fantastic pictures or to provide fake facts. The forger could use a single or a mixed sequence of image processing operations, no matter what the source of the act may be [2]. As a result, Image Tampering Detection (ITD) process very importantly. The purpose of digital image tampering detection is to check digital image accuracy without any background knowledge of the originals [3]. In an image classification scheme, there are two principal measures. The first step is to identify an effective representation of an image which, for future classification, includes appropriate image details. The second step is to classify a good classifier for the new picture. Therefore, to enhance image classification efficiency [4].

Image tampering is described as adding, altering or removing any essential features from an image without leaving any apparent sign [5]. Image tampering comes in two different types, they are: active approach and passive approach. To study the information of a tampered image, an active approach is used where any information like a watermark has previously been embedded within the image. It was necessary to do this embedding during a special camera is required to produce the image, but if the watermark was not added during the image production, it may be inserted in the post-processing process. A passive approach is the most popular technique since certain amounts of photographs have a built-in watermark, and often it could damage a watermark or because the person who conducted the test doesn't know what it is [6].

FIGURE 1 shows the basic concepts of the most common two tempering types: copy-move and splicing, while Fig. 2 shows the use of these concepts on real images. The copy-move attack is one of the easiest types of digital image forgery. There are three steps in the copy-move process: copy the selected image fragment from one location, convert it using an image processing algorithm, and paste it into another location of the same image [6]. Splicing is an image compositing technology that uses image fragments from the same or separate images without any post-processing, such as smoothing borders between different fragments. Retouching was also one way to get rid of photo errors in the days of video photography (dust, hair, etc.). Retouching was then a procedure includes a sequence of steps to enhance an image [7].

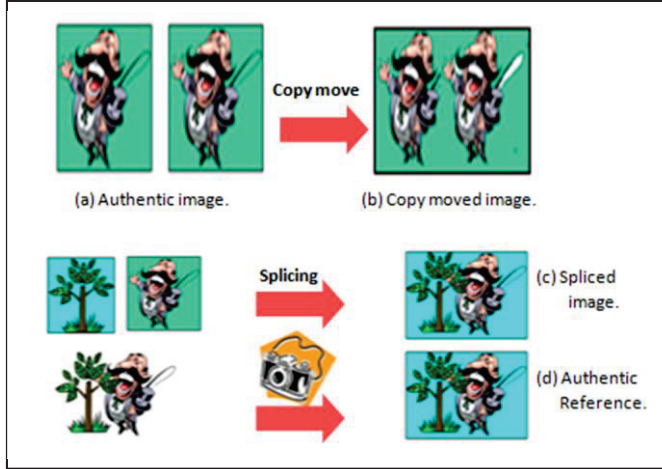


FIGURE 1. Copy move and splicing tempering processes [6].

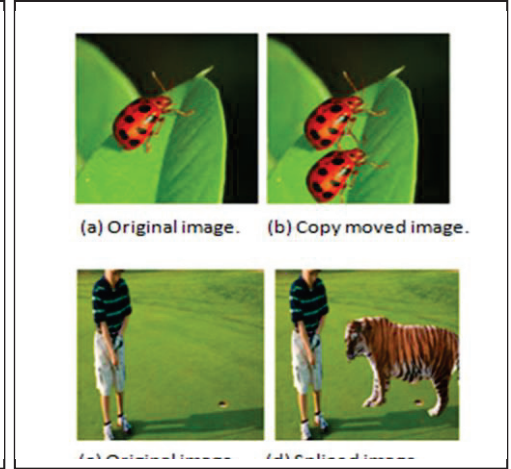


FIGURE 2. Real tempered images [6].

EXTREME MACHINE LEARNING

The forward feeding of the neural network (NN) through the learning phase is typically slow. This slow is almost coming from two reasons: the use of gradient-based learning in training the neural network when facing large databases, and also the use of learning algorithms require iteratively tuning all parameters of the networks. This problem make a huge consideration in NN experiments for years [8]. Extreme learning machine (ELM) is proposed to overcomes such problem, it does not use gradient-based techniques. This makes it run much faster than its competitors. The advantage of ELM over traditional NN algorithms is that all the parameters are tuned once due to the ELM does not need iterative training [29]. ELM is a significant new strategy for machine learning that providing a computationally efficient classifier that single-layer Feed-forward Network (SLFNs) as shown in Fig. 3, ELM theories demonstrate that in learning, hidden neurons do not need to be tuned and their parameters may be independent of the training results, but even so, ELMs have common properties for comparison and classification [9].

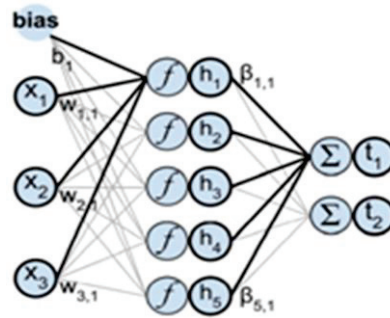


FIGURE 3. The ELM model [29].

In ELM, the number of training samples (N) is determined by the number of input features, while x_j , L presents the number of hidden neurons. The output function of the ELM is written by the following equation [8]:

$$g_L(\mathbf{x}) = \sum_{i=1}^L \alpha_i \Psi_i(w_i \cdot \mathbf{x} + b_i) = \Psi(\mathbf{x})\boldsymbol{\alpha} \quad (1)$$

where w and b are the weight and bias of the hidden neurons, $\boldsymbol{\alpha}$ is the output weight, and Ψ is the activation function.

The optimal output weights are found by using the training samples as follows:

$$\boldsymbol{\alpha}' = \mathbf{H}'\mathbf{M} \quad (2)$$

where, \mathbf{H}' is the Moore-Penrose matrix, which is formed by $\mathbf{H} = [\Psi(x_1), \Psi(x_2), \dots, \Psi(x_N)]^T$. \mathbf{M} is the training data teaching array. $\mathbf{H}' = (\mathbf{H}^T \mathbf{H})^{-1} \mathbf{H}^T$. The Rectified Linear Unit (ReLU) is an activation function that can be used when there are two used classes such as authentic and forged. ReLU has strong and solid mathematical roots are depending on setting the threshold at zero (e.g. $f(x) = \max(0, x)$), where it produces 0 when x is zero and a linear function when x is greater than zero as follows [10]:

$$\ell(\theta) = -\sum y \cdot \log(\max(0, \theta x + b)) \quad (3)$$

RELATED WORK AND CONTRIBUTION

The detection of tampered images have been given a lot of attention in the last few years. Thus, there are several works have addressed the issue of image tempering. The following subsections include a shorten survey about previous work that related to the field of interest:

Related Work

The related work includes a number of published papers that simulate the developments in the image tempering detection approaches, the following ones are representing the state of the art. In [11], test image is first automatically segmented into distinct areas to study the Camera Response Feature (CRF) that was calculated using geometric invariants from Locally Planar Irradiance Points (LPIPs). Then, CRF-based cross fitting and local image features are computed and fed to statistical classifiers to identify a boundary section between two areas as authentic or spliced. In [12], Model the edge image of the image chroma variable as a finite-state Markov chain based on modeling edge information and remove the low-dimensional function vector from its stationary distribution to detect tampering. The Support Vector Machine (SVM) is used to test the usefulness of the proposed algorithm, the use of separate picture channels (Y, Cb, and Cr) and varying thresholds gave a detection results are 66.5%, 95.6%, and 95.5% consequently. In [13], the training employed a huge number of blocks that extracted off-line and characterized by features centered on a dense local descriptor. The training features are then fitted by a multidimensional Gaussian model. Then the tampering is localized via basic thresholding. This method has a really good detection efficiency without false alarms, although often (rarely) the comparison PRNU-based method shows certain near misses and fail. Special descriptor was created In [14] to be used for each image row, merging the artificial grid function of the JPEG block with that of noise estimation. The forehead picture consistency measurement protocol reconciled these distinct characteristics by setting proper weights. Experimental findings revealed that this approach was efficient for detecting both copy-move and splicing forgery with accuracy of 100%. In [15], the camera model characteristics were extracted from source images, this algorithm uses a convolution neural network (CNN) to detect whether an image has been forged and identify the fake region. These characteristics are then analyzed using iterative cluster analysis. The evaluation of 2000 images dataset that collected from 26 camera models reveals that this algorithm is capable of detecting forged images with an accuracy of 91%. Tampering localization results showed an accuracy in between 82-90% depending on the information availability of the used camera models at the training step. In [16], the copy transfer picture forgeries was used to establish an active forgery identification approach. The picture is subdivided into smaller patches of fixed size that overlap each other and then mark areas of tampering to figure out the tampered parts in the picture. The experiment results gave an acceptable forgery detection rate and detection time. The research given in [17] proposed a technique for digital image tamper detection based on CFA artifacts resulting from variations in obtained and interpolated pixel distribution. This technique is based on estimating each pixel's probability of being interpolated and then adding the DCT to the probability map's small blocks. The value of the highest frequency coefficient for each block is used to determine whether or not the evaluated area has been tampered with. The achieved accuracy is 86%, which gives a decent outcome for a technique that doesn't even require prior training. In [18], the spatial grey level dependency system was used based on many characteristics that derived from regular and spliced images. SVM and Twin-SVM have been used for classifying the extracted features: Local Binary Pattern (LBP), Contrast, Entropy, Histogram of Oriented Gradient (HOG), Inverse Different Moment, Speeded up Robust features, Interia, Cluster shade, cluster prominence, Angular Second Moment. The experiments used two different datasets of authentic and spliced images:

Colombia which gives 88.63% accuracy, and MICC-F220 gives 93.33% accuracy. The paper in [19] proposed a new passive splicing detection approach based on textural features and Gray Level Co-occurrence Matrix (TF-GLCM) that calculates the GLCM using Differential Block Discrete Cosine Transform (DBDCT) arrays to accurately capture the textural details and relative distance between image pixels. The SVM is used to classify the TF-GLCM features to achieve response times of 98% on CASIA v1.0 and 97% on CASIA v2.0. In [20], the LBP function is applied on each block in image to produce the authentication. According to the block mapping sequence, the current image block's pixel indicate and LBP was inserted into the 2-LSBs of the related image block. The identification and recovery of image tampering is performed. The recovery is optimized by the eight binary of mean value for each eight neighboring pixels in image.

Contribution

The proposed method looks for the detection of the tampering in images that may be appeared in different situations or conditions, and places. The contribution is concentrated to the use of greater number of texture color features with the extreme learning machine (EML) classifier to solve the problem of authenticity of each image in a huge dataset. This make the number of nodes in the EML in both input and hidden layers are greater, and yielding more accurate classification decision that enable to discover fraud in the image whether it was trained or not.

PROPOSED TID METHOD

The generic structure of the proposed TID method is shown in Fig. 4. It is shown that the proposed method is designed to be consisted of two phases: training and classification (i.e. detection). Both phases are passing through one preprocessing stage, which is image conversion. The features extraction stage depends on computing the local binary pattern (LBP) that encoded in four directions to be used as distinguishing features, and input the ELM classifier. The used activation function used in the ELM classifier is the ReLU, which can label the given image block in just two classes: authentic or tempered. The training phase concerned with learning the ELM model depending on distinguishable features of image samples to be stored in a database array and used as a comparable models in the classification phase. Multi stages are found in the training phase, they are: image blocking and LBP computation, image reconstruction, differences estimation, features extraction. Whereas the classification phase is responsible on verifying the contents of the test image in comparison with the database models, this is carried out by classifying the considered image as one block with the database models to be classified if it is authentic or tempered. More details are explained in the following subsections.

Image Conversion

The input colored image is converted into the YC_bC_r color model, where Y represents the luminance, and C_b and C_r represent the two chrominance components. The conversion into YC_bC_r is a functional approach to color processing and visual standardization, where the primary colors relating approximately to red, green, and blue are transformed into perceptually relevant details and image chrominance components of the pixel array are separated into a set of non-overlapping $n \times n$ blocks, each is given in terms of Y , C_b , and C_r components using the following relation:

$$\begin{bmatrix} Y' \\ C_b \\ C_r \end{bmatrix} = \begin{bmatrix} 65.481 & 128.553 & 24.966 \\ -37.797 & -74.203 & 112 \\ 112 & -93.786 & -18.214 \end{bmatrix} \begin{bmatrix} R' \\ G' \\ B' \end{bmatrix} + \begin{bmatrix} 16 \\ 128 \\ 128 \end{bmatrix} \quad (4)$$

LBP Computations

The YC_bC_r image is divided into multi blocks of size 8×8 and then each block is treated as image and blocked into non-overlapped 3×3 neighborhood pixels. When a pixel intensity is greater than central pixel intensity, the pixel in the middle pixel is then allocated to 1, otherwise it is allocated to 0. By organizing the 1s and 0s in clockwise or anti-clockwise orders, one can gets an 8-bits binary code. Then, this generated code number is translated into a decimal and the middle pixel is allocated that number [21]. For every image pixel, this process is repeated Applying Local Binary Pattern to each block of the image. For texture analysis, image detection, and pattern recognition, the LBP operator is often used to obtain localized spatial characteristics.

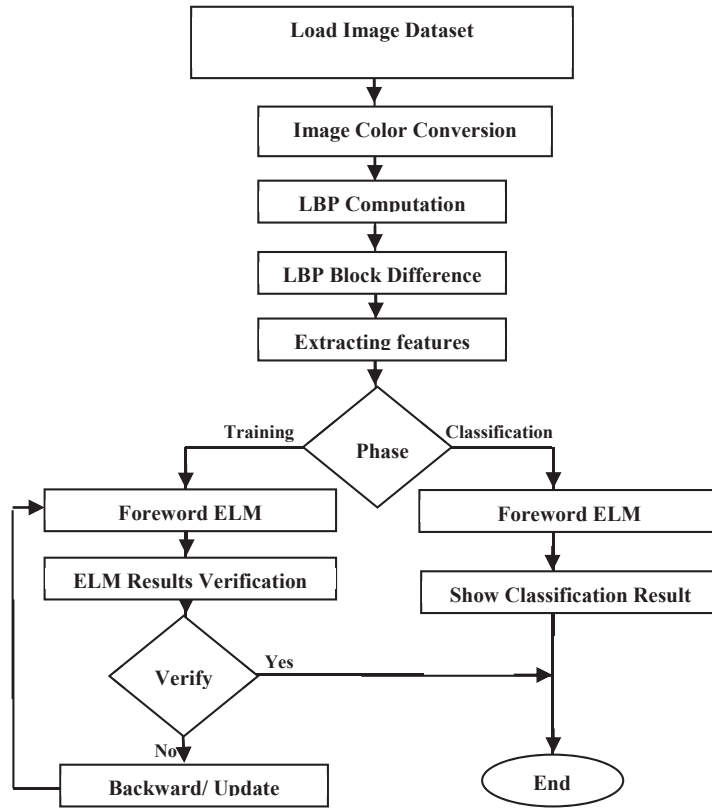


FIGURE 4. Generic structure of the proposed TID method.

LBP Block Difference

The difference between LBP block (DBLBP) arrays are calculated in the four directions, and then the Haralick features are calculate for each difference BLBP array.

Features Extraction

Features extraction used Haralick relations given in [9] to estimate the classification features. The basis for these features is the gray-level co-occurrence matrix (GLCM). These texture features, depending on the adjacency matrix, where the adjacency matrix records how many times that a pixel in location (i,j) to take i value next to j value. The average of these records over the four directions is determined to describe any rotational invariance. The first 13 features are calculated directly producing 4×13 features vector (one row per direction) for 2D images.

Training and Classification

ELM classifier is used in training and classification phases, this is because of its quick and effective learning speed, fast consistency, strong generalization performance, and ease of execution. Also, ReLU is used as an activation function to convert the summed weighted inputs entering the neuron to a specific output value that is sent to the next layers. Then, the data rearranged using One-Hot Encoding to be used in training the ELM for impalements the classification process and making the material data to be clearer and easier to understand network.

RESULTS AND DISCUSSION

The used dataset to test the proposed TID method comes from multiple datasets. In particular, two datasets are used in the training and classification phases. In the approved model, CASSIA-II was relied upon as data for the system training and instruction because it in turn contains 12616 image files divided into two groups that include 5124 tampered images and 7492 authentic images, and thus the number of images in it is greater and learning of manipulation is done by relying on more various conditions of lightening and various images sizes and they include a variety of images, including characters, landscapes, animals, plants, and others to learn machine to detect forgery. Whereas, the

classification was carried out on CASSIA-I, which in turn contains 1721 image files, including 800 authentic and 921 spliced images.

TABLE 1 shows the result of applying the proposed TID method on CASSIA-I dataset for detecting the tampering in images, in which the use of the three color bands to predict the classification decision was 798 positive prediction and 7 negative prediction when testing positive classes. Also the positive prediction was 914 and negative prediction was just 1 when testing the negative class. These results indicates that the computed classification accuracy of using C_r band is about 99.4%, while it is about 99.7% when using C_b band, and the use of Y band gave a classification accuracy of about 99.7% , in which the average runtime was about 6.7 seconds. The classification using the EML classifier which in turn is considered to be highly influential on the time factor of the chase with accuracy. For issues with imbalanced classification, precision and recall is an ineffective success matrices. Therefore, the computed evaluation measures were as follows: the precision is equal to 0.99, recall is equal to 0.991, and F-measure is at its best a perfect score which is 0.98.

TABLE 1. Resulted numbers of positive and negative classification for tempering detection.

Total=1720	Positive prediction	Negative prediction
Positive class	798	7
Negative class	914	1

CONCLUSION AND FUTURE WORK

The use of textural features in the present work gives a very good description for effective and important image properties, and enable to classify the target image into tempering or authentic one. This comes in harmonic behavior with EML that in turn showed very effective tool for achieving intended accuracy at reduced computation time. The suggestion for future work include adding more powerful features such as SIFT to overcome the problem of arise training performance and providing great ability for detection and localize the region where the tampering happened exactly.

REFERENCES

1. B. Liu, C. M. Pun, and X. C. Yuan, "Digital image forgery detection using JPEG features and local noise iscrepancies," *Sci. World J.*, 2014, doi: 10.1155/2014/230425.
2. D. Sharma and P. Abrol, "Digital Image Tampering – A Threat to Security Management," *Int. J. Adv. Res. Comput. Commun. Eng.*, 2013.
3. W. Wang, J. Dong, and T. Tan, "A Survey of Passive Image Tampering Detection," pp. 308–309, 2009.
4. F. Cao, B. Liu, and D. Sun Park, "Image classification based on effective extreme learning machine," *eurocomputing*, vol. 102, pp. 90–97, 2013, doi: 10.1016/j.neucom.2012.02.042.
5. S. Kumar, J. Desai, and S. Mukherjee, "A fast DCT based method for copy move forgery detection," *2013 IEEE 2nd Int. Conf. Image Inf. Process. IEEE ICIP 2013*, pp. 649–654, 2013, doi: 10.1109/ICIP.2013.6707675.
6. D. A. Mendoza, "Digital Forensics Method for Image Tampering Detection," 2017.
7. M. Sridevi, C. Mala, and S. Sanyam, "Comparative study of image forgery and copy-move techniques," in *Advances in Intelligent and Soft Computing*, 2012, doi: 10.1007/978-3-642-30157-5_71.
8. G. Bin Huang, Q. Y. Zhu, and C. K. Siew, "Extreme learning machine: Theory and applications," *Neurocomputing*, 2006, doi: 10.1016/j.neucom.2005.12.126.
9. A. Akusok, K. M. Bjork, Y. Miche, and A. Lendasse, "High-Performance Extreme Learning Machines: A Complete Toolbox for Big Data Applications," *IEEE Access*, vol. 3, pp. 1011–1025, 2015, doi: 10.1109/ACCESS.2015.2450498.
10. A. F. M. Agarap, "Deep Learning using Rectified Linear Units (ReLU)," *arXiv*, no. 1, pp. 2–8, 2018.
11. Y. F. Hsu and S. F. Chang, "Camera response functions for image forensics: An automatic algorithm for splicing detection," *IEEE Trans. Inf. Forensics Secur.*, 2010, doi: 10.1109/TIFS.2010.2077628.
12. W. Wang, J. Dong, and T. Tan, "Image tampering detection based on stationary distribution of Markov chain," in *Proceedings - International Conference on Image Processing, ICIP*, 2010, doi: 10.1109/ICIP.2010.5652660.
13. L. Verdoliva, D. Cozzolino, and G. Poggi, "A feature-based approach for image tampering detection and localization," in *2014 IEEE International Workshop on Information Forensics and Security, WIFS 2014*,

- 2015, doi: 10.1109/WIFS.2014.7084319.
14. W. Li, Y. Yuan, and N. Yu, "Passive detection of doctored JPEG image via block artifact grid extraction," *Signal Processing*, vol. 89, no. 9, pp. 1821–1829, 2009, doi: 10.1016/j.sigpro.2009.03.025.
15. L. Bondi, S. Lameri, D. Guera, P. Bestagini, E. J. Delp, and S. Tubaro, "Tampering Detection and Localization Through Clustering of Camera-Based CNN Features," in *IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops*, 2017, doi: 10.1109/CVPRW.2017.232.
16. P. Singh, "Correlation Based Image Tampering Detection," no. November, 2017.
17. E. González Fernández, A. L. Sandoval Orozco, L. J. García Villalba, and J. Hernandez-Castro, "Digital Image Tamper Detection Technique Based on Spectrum Analysis of CFA Artifacts," *Sensors (Basel)*, 2018, doi: 10.3390/s18092804.
18. R. Agarwal and M. Pant, "Image tampering detection using genetic algorithm," *MATEC Web Conf.*, 2019, doi: 10.1051/mateconf/201927702026.
19. X. Shen, Z. Shi, and H. Chen, "Splicing image forgery detection using textural features based on the grey level co-occurrence matrices," *IET Image Process.*, 2017, doi: 10.1049/iet-ipr.2016.0238.
20. J. D. Chang, B. H. Chen, and C. S. Tsai, "LBP-based fragile watermarking scheme for image tamper detection and recovery," in *ISNE 2013 - IEEE International Symposium on Next-Generation Electronics 2013*, 2013, doi: 10.1109/ISNE.2013.6512330.
21. D. P. Mohapatra and S. Patnaik, "Preface," *Adv. Intell. Syst. Comput.*, vol. 243, p. V, 2014, doi: 10.1007/978-81-322-1665-0.