# Recognition of Copy Move Forgeries in Digital Images using Hybrid Optimization and Convolutional Neural Network Algorithm

Anna Gustina Zainal[1], Dr. Chamandeep Kaur[2], Dr. Mohammed Saleh Al Ansari[3], Ricardo Fernando Cosio Borda[4], Dr. A. Nageswaran[5], Rasha M. Abd El-Aziz[6]

Department of Communication, University of Lampung, Indonesia[1]
Lecturer, Department of IT, Jazan University, Saudi Arabia[2]
Associate Professor, College of Engineering-Department of Chemical Engineering, University of Bahrain, Bahrain[3]
Universidad Privada del Norte, Peru[4]
Professor, Department of CSE. Sri Venkateswaraa College of Technology, Vadakkal village, Sriperumbudur- 602105[5]
Department of Computer Science-College of Science and Arts in Qurayyat, Jouf University, Saudi Arabia[6]
Faculty of Computers and Information, Assiut University, Assiut, Egypt[6]

***Abstract*—In the modern day, protecting data against tampering is a significant task. One of the most common forms of information display has been digital photographs. Images may be exploited in a variety of contexts, including the military, security applications, intelligence areas, legal evidence, social media, and journalism. Digital picture forgeries involve altering the original images with strange patterns, which result in variability in the image's characteristics. Among the most challenging forms of image forgeries to identify is Copy Move Forgery (CMF). It occurs by copying a portion or piece of the picture and then inserting it again, but in a different place. When the actual content is unavailable, techniques for detecting fake content have been utilised in image security. This study presents a novel method for Copy Move Forgery Recognition (CMFR), which is mostly based on deep learning (DL) and hybrid optimization. The hybrid Grey Wolf Optimization and African Buffalo Optimization (GWO-ABO) using Convolution Neural Network (CNN) technique i.e., GWO-ABO-CNN is the foundation of the suggested model. The developed model extracts the features of images by convolution layers, and pooling layers; hereafter, the features are matched and detect CMF. The MICC-F220, SATs-130, and MICC-F600 datasets were three publicly accessible datasets to which this methodology has been implemented. To assess the model's efficacy, the outcomes of implementing the GWO-ABO-CNN model were contrasted with those of other approaches.**

***Keywords*—Copy move forgery; convolutional neural network; image authentication; deep learning; tampered images***

## I. Introduction

Today's technological images play a crucial role in a broad range of fields. They are used in a variety of uses in the fields of broadcasting, journalism, medicine, and the army, to name a very few. The computerised image can be seen as a notable resource of information in today's advanced globe due to the advancement in the technology of sophisticated picture, such as sensors, coding, and Computers, as well as the widespread usage of the internet [1]. In addition, advanced image forgery refers to the intentional manipulation of a digitized image in order to change the conceptual interpretation of the contextual perspective contained within. Also, with availability of cutting-edge information structures editing tools like Photoshop, it becomes quite simple to create sophisticated fakes from one or more images. The reliability of photographs plays a crucial role in a variety of fields, such as measurement analysis, criminal probe, surveillance systems, organisational learning, medical imaging, and media broadcasting. Creating phony pictures is a specialty with a lengthy tradition. But in the current technology age, it showed out to be very simple to change the facts talked to by an image without any obvious consequences.

One of the most widely popular techniques for altering digital photographs is copy-move. One of two factors can explain why there are copy regions in an image: first, the proximity of two particles or objects that are identical in size, form, and colouring; one of them may be a copy of the other one. Second, the appearance of duplicate regions in the results is caused by the proximity of a reasonably massive area with one colour and similar in features, such as foundational principles sky, splitter, etc. Copy-move forgeries is created by copying and pasting a region or sector from one spot in an image to some other spot inside the same picture in order to modify or hide one or more objects and create a false vision. Moreover, copy-move forgeries identification is known to be successful when using key point-based analysis. There were some changes made to the image during copy-move forgeries. Moreover, to implement quality impersonating forgeries, techniques including turning, cropping, lightening, reduction, and force and contribute are used. Nowadays, even a non-expert may easily produce convincing forgeries in digital images because of modern digital imaging and robust image manipulation tools. Huge different forgeries have been created in recent decades as a result of digital manipulation, which involves incorporating or removing certain parts from the image. Thus, checking the materials of digital photos or discovering fraudulent areas would be immediately helpful, for example, when images are used as evidence at trial.

Evidently, there is no one single, efficient option to the challenge of detecting digital forgery [2].

Two kinds of techniques and block-based approaches— can be used to recognize the presence of a copy-move in a visual. The fundamental area of assessment consists of examining the decentness of images and identifying signs of modification without the requirement to use image files. Structurally complex transform, resilient to disturbance and geometrical modifications are only a few samples of the main aspects that feature-based techniques for copy-move verification retrieve from copied images. Block-based copy-move identification, in comparison, analyses the properties of each block in the frequency response by splitting the copied images into matching or non-matching blocks [3]. LBP uses the features retrieved from the frames of gray-level image data to pinpoint the fake portions in visuals. Moreover, with SURF, local picture features that are resistant to noisy and geometrical changes are quickly extracted. In terms of how the two methods replace a specific image location with some other picture, copy move manufacture is very similar to picture patching. For instance, copy move forgery uses a portion of the original base picture as its origin instead of using an external image. The comparable image served as both the origin and the endpoint for the modified image. Hereafter, parts of the initial image are duplicated, moved to the perfect location, and then merged in a duplicated move fraud. Typically, this is used to conceal particular nuances or to replicate specific areas of an image [4]. Moreover, to minimise the impression of irregularities between the initial and delayed region, some additional preparation, such as center separation and hiding, is often linked along the edge of the modified region.

Furthermore, the judgement regarding the validity of the data is subsequently made by evaluating the outcomes of several strategies. In single-image forgeries, the copy-move approach is used to substitute a section of the original picture with a section that was previously eliminated. Forgers employ certain comment on these thread modifications on the duplicated portions, including spin, scalability, and reflections, as well as mouldings and mixing, to make forgeries are easier to conceal. The addition of content for one image that comes from another image or photos is called forgeries using various pictures [5]. In addition, there are two distinct streams of computerized photo fraud detection. They are passive tactics and dynamic approaches. Images are extremely simple to change in the technological environment utilising well-known image manipulation software applications like Adobe Illustrator and Gimp, which results in an astounding number of phoney images entering in every day over the Web.

Identification of Digital image forgery is an important research area; however, it seems to be a significant problem. Finding proof of a fake is done by looking for unique characteristics, and traits. The actual image's analytical, architectural, or physiological qualities must be homogeneous for the detection techniques to work, but manipulated images lack this homogenity. Digital image counterfeiting comes in a variety of forms, including picture merging, image reshaping, image editing, image re-sampling, and copy-move forgeries [6]. Several modified over time in a variety of domains, including target detection, ML, and computation imagery, were successfully solved using the DL method. As a result, it can be useful in developing patching forgery investigative techniques.

Among the most successful and effective techniques applied in a variety of applications for image processing during the past ten years is the Artificial Neural Network. In addition, Convolutional Neural Network (CNN) is commonly employed for Copy-Move Forgery Detection in recent years. However, CNN method implemented on challenging datasets has attained less accuracy and high testing time in CMF recognition [7]. To address this issue, the current work proposes an effective hybrid optimization with CNN for recognizing CMFD from digital images, which can achieve great result at a very cheap computational effort. CMFD's objective is to identify areas that are comparable to certain other areas of the image. Architectural or post-processing procedures are typically performed upon altered areas during the tampering procedure to make the forgeries realistic and undetectable. The key piece of evidence in CMFD is the strong resemblance between the modified areas and the origin. Here, the GWO-ABO with CNN technique serves as the framework's main tenet.

The rest of the section is organized as follows: Section II describes the recent literatures related to CMF detection, Section III explains the materials and methods in which the dataset and proposed methodology are detailed, Section IV describes the result and discussion, and Section V concludes the paper.

## II. Related Works

Although there are many other styles of digital picture scams, it can be quite difficult to spot copy-move forgeries. In order to identify copy-move fraud, the study utilized a new robust approach which is based on the Speeded up Robust Feature descriptor, Approximate Nearest Neighbour for feature representation, and Simple Linear Iterative Clustering for categorising the actual image into large pixel blocks. The portions that are in dispute are identified by swapping out paired key frames for comparable large pixel blocks, followed by the merging of adjacent blocks based on similar LCF [8]. On a variety of samples, including MICCF220, MICC-F600, MICC-F2000, and CoMoFoD, the paper measured a time complexity of 3.84 secs with 91.95% positioning accuracy. However, because the process takes longer than the alternative way, it cannot be used in the present circumstances.

The author in [9] utilised robust feature enhancement acceleration, and the SVM is used to recognise the particular object. There were some changes made to the image during copy-move forgeries. To create effective portrayal forgery, techniques including rotating, magnification, lightening, reduction, and force and contribute are used in this paper. The dominant features from the source images are chosen in this case by the SURF extracting features. The SVM classifies these input photos using detection and recognition in order to retrieve the paired local features. Here, the outcomes demonstrate that forging images are taken from a collection of test images. When there is additional uncertainty in the data set, SVM doesn't really function very well.

The paper [10] used a reliable technique for identifying Copy-Move forgeries in digital photographs. The procedure begins by removing an image's SIFT characteristics, which are resistant to modifications in lighting, spin, scale, etc. Characteristics are then compared to one another in order to check for any potential picture forgeries due to the resemblance between the inserted region and the duplicate data region. The utilised system works well with various post-image analysis methods and is resistant to complex image analysis because of the high robustness of SIFT features extracted. However, to increase the robustness over inadequate SNR and small-size tampering regions, more research is still required.

The paper [11] utilized a technique for spotting fake parts in digital photographs. The steps in this procedure are as follows: (1) transform the colour image into gray, (2) break the image representation into overlaying blocks of pixels, (3) use DCT to identify and extract feature, (4) aggregate blocks using the K-means approach, and (5) use radix order to correlate features. Here the method was appeal four different photos and on the foundation of the experimental work. Diverse methods, such as DCT and DWT, are typically utilized to extract the characteristics from the digital photos in unrelated domains, such as feature recognition and image fraud prevention.

A reliable copy-move picture fraud prevention method used Gaussian-Hermite Moments in the paper. The method separates the image as an input into overlaying, fixed-size chunks, and then extracts the GHM values for every chunks. GHM is a powerful tool that may be used to recover picture characteristics that are rotation-, translation-, and scale-invariant. Through lexicographically ordering all the characteristics', related chunks are matched. The outcomes of the experiments demonstrate how precisely the suggested technique can find the duplicate forged portions in a forged image [12]. Here, the used technique performs better than previous relevant strategies, according to experimental findings, both at the image and pixel levels. However, the method can necessitate additional time, rendering it inappropriate for the next presentation.

The outcome of applying the lighting estimation technique reveals that the method is only marginally improved by excluding specific observations, such as geometry depending on shadowing and upgrade and improve characteristics on selected input images [13]. The technique to identify image modifications is approved by the publication in consideration of the abnormal illumination. By calculating surfaces and lighting data from the centre points, the counterfeit in images is identified. The process produces acceptable results for many source images. Here, the procedure is not just restricted to human faces and the identification of image regions with the same brightness. It can be used for photographs with any kind of item included in the scene. However, the method needed to be improved in order to estimate sections with the same brightness and to automatically choose spots.

The method developed in this work is a reliable one for spotting digital picture forgeries and is sufficiently robust to withstand attempts at image alteration. The very first process in the work is to generate the Rgb values into YCbCr space. Next, the Hilbert-Huang Transform includes are obtained from the chrominance-red element Cr. Finally, three distinct classification techniques, KNN, SVM and ANN—have been evaluated and contrasted for the task of classifying images as accurate or fake. Structural-Similarity is used to measure the precision of copy - move and to validate the findings. The suggested technique may identify pixels in the context of several images post-processing assaults, according to the findings of the robustness assessment; nevertheless, these cyber-attacks gradually degrade it [14]. However, since these techniques necessitate a large amount of storage for the training examples, the procedure might be slowed significantly.

The paper [15] employed a blind verification approach based on bundle segmentation and visual scrambling to discover repeated sections for copy-move forgeries. DCT is employed to find the DCT coefficient matrices for each level sufficient source image in suspected images. The visual hashing characteristic matrices and matrices are treated in a systematic manner. Additionally, to increase analysis is an essential part; a package grouping technique is utilized in place of conventional textual order techniques in this work. The perceptive hashing vectors in every package and its neighbouring packages, comparable chunks can be distinguished. A computerisation that builds a feature representation to describe an image block using perceptive hash characters can withstand certain common attacks like adding white Gaussian noise and GB. Additionally, the suggested approach is vulnerable to some sophisticated assaults such sector manipulation and scalability. However, it is more expensive than other ways, consequently not everyone can benefit from it.

One of the most prevalent popular forms of digital image forgeries, merging, was identified using a technology in the study [16] . Here, the VGG-16 CNNs are the foundation of the algorithm. The network architecture in used to receive visual features as information and determines whether an update is authentic or fake. During the learning phase, the work chooses portions on the edges of the inserted splice and from the original image portions. The potential for using the strategy when the JPEG technique repeatedly compresses deformed photos over a small range. As compared to other techniques, the reported findings show great generalization ability for a set of photos with intentional imperfections. However, Dynamic graphic images cannot be handled by the JPEG image standard, and stacked images are not supported.

The forensic evidence of analyser sensors was addressed in the paper using a ML-based method. Numerous diagnostic approaches, including digital image authenticating, related to Google, and security mechanisms, are crucial for digital image assessment because of the rising accessibility and effectiveness of image altering applications. In this work [17], a CNN-based method for identifying scanner framework has been used. The test results demonstrated that originating scanner authentication may be done with extreme precision. A system, which brings together the advantages of ResNet and GoogleNet while still being compact. A dependability map that shows the altered areas in a digital image is also generated by the article. However, a deeper network has a significant downside in that it typically takes weeks to train, rendering it unworkable in practical systems.

In the paper [18], two methods for recognizing image combining are supported. Both techniques are used to retrieve input image features by using overlapped chunks. The first method is used to extract LBP or LTP components based on the image's chrominance's gray-levels, while the other approach retrieves ELTP characteristics from the luminance site's fast Fourier transformation. These methodologies' outcomes have been summarized in a favourable manner. By reaching an accurateness of 88.62% on data compression from the CASIAv1.0 collection, the FFT-ELTP approach works reasonably well. However, all of the strategies that are being discussed entail intricate modifications like the DCT and FFT, which makes the technique more difficult. It may be possible to eliminate the requirement for such sophisticated activities with the effort in a similar approach. An additional subject for investigation in the investigation is the distribution of the forgeries in the photograph.

The paper [19] utilized, a key point-based image evidence collection procedure based on the Helmert conversion and super pixel recognition technique has been suggested. The method seeks to gather forensic data while detecting copy-move forgery photos. The activities or tasks make up the suggested approach's process. First, use a SIFT approach to identify the key details and related attributes. The resemblance between key points will then be determined based on the descriptors to produce matched pairings. In order to determine the precise counterfeit zones, the Helmert conversion to group the matched pairs according to geographical distances and geometry restrictions. In respect to province approaches, the experimental outcomes from evaluating multiple samples show that the suggested approach achieves excellent highly precise rates. However, the used method is not resistant to symmetrical, recurrent, and smoother processes for regional manipulation.

## III. Materials and Methods

### A. Dataset Description

The most difficult or well-known databases in the evaluation of CMF recognition methods are MICC-F220 [20], SATs-130 [21], and MICC-F600 [22]. Table I provides detailed information of various assessment datasets. The MICC-F600 database challenged attacks have been organised into four tiers with various forgery attacks. The MICC-F600 and MICC-F220 databases had been used to effectively test the presented technique. While the SATs-130 database has only 96 images, which is too little and collide with DL method's nature. Large databases are required for the training phase of DL methods in order to accurately extract feature maps and construct system behaviour.

TABLE I. Dataset Specification

| Sl. no | Database | Image Composition | | | Tampered region size | Images size (Pixels) |
|---|---|---|---|---|---|---|
| | | *Total* | *Forged* | *Original* | | |
| 1 | MICC-F600 | 600 | 152 | 448 | The size of the forged region changes from image-to-image. | 800×532 to 3888×2592 |
| 2 | SATs-130 | 96 | 48 | 48 | The size of the forged area changes from image-to-image. | 1024×683 to 3264×2448 |
| 3 | MICC-F220 | 220 | 110 | 110 | 1.2% of the entire image seems to be the forged area. | 722×480 to 800×600 |

### B. Proposed Methodology

The suggested model has four steps, and Fig. 1 depicts the proposed method's workflow. Pre-processing seems to be the initial step. Two things are being accomplished by this technique. The first objective is to uniformly scale all of the input images, and the second objective is to transform the images into tensors. The extraction of features seems to be the second stage. To retrieve the features from the raw images, this step is obtained. Feature matching seems to be the third stage, which is obtained to demonstrate the presence of forgeries. The final stage is post processing. First, potential erroneous matches are removed at this stage. The CMF in digital images is afterwards identified.

*1) Pre-processing by GWO-ABO:* GWO appears to be a bionic optimization method. It mimics the cooperative, well delineated working relationships found in grey wolf hunting behaviour. Grey wolves often live in packs of five to twelve people and have a rigid, dominant hierarchy based on wolf leadership skills. The three phases of the GW pack's predatory process: hunting, encircling, and attacking. The most prominent wolf in the pack, also known as wolf, typically served as its leader. The GWO terminology for the second and third tiers of leadership wolves is wolf and wolf, respectively. These second- and third-ranking auxiliary wolves aid the main wolf in making hunting decisions. All other following wolves are classified as wolves, and they pursue and engage in combat with the prey alongside these powerful wolves.

For hunting, the enclosing of prey approach is used. For iteration $i$, the mathematical framework for this method is shown in below Eqns. (1) and (2).

$$\vec{F} = \left|\vec{Y} \times \overrightarrow{Q_p}(i) - \vec{Q}(i)\right| \tag{1}$$

$$\vec{Q}(i+1) = \overrightarrow{Q_p}(i) - \vec{D}.\vec{F} \tag{2}$$

Here, $\vec{D}$ and $\vec{Y}$ are coefficient vectors, which is described as $\vec{D} = 2\vec{d}.\overrightarrow{v_1} - \vec{d}$ and$\vec{Y} = 2.\overrightarrow{v_2}$. Where, the random vectors $\overrightarrow{v_1}, \overrightarrow{v_2}\ \varepsilon\ (0,1)$ and$\vec{d} = d_1\left(1 - {}^{i}/_{maxi}\right)$, linearly decreases from $d_1$ to zero; $d_1$ value was set as 2 in actual GWO. Moreover, $maxi$ represents maximum number of iterations. The GWO's hunting process has been headed by three finest solutions i.e., $\propto$, $\beta$ and $\gamma$ wolves. Thus, these 3 leading solution's positions have been saved in the pack and the remaining $\omega$ wolves update their positions predicated on them.

This position updating technique's mathematical model is represented in Eq. (3).

$$\vec{Q}(i+1) = {(\overrightarrow{Q_1} + \overrightarrow{Q_2} + \overrightarrow{Q_3})}\big/_{3} \qquad (3)$$

Where, $\overrightarrow{Q_1}, \overrightarrow{Q_2}, and\ \overrightarrow{Q_3}$ is computed by Eqn. (4)

$$\begin{aligned}\overrightarrow{Q_1} &= \overrightarrow{Q_\propto}(i) - \overrightarrow{D_1}.\overrightarrow{F_\propto}\\ \overrightarrow{Q_2} &= \overrightarrow{Q_\beta}(i) - \overrightarrow{D_1}.\overrightarrow{F_\beta}\\ \overrightarrow{Q_3} &= \overrightarrow{Q_\gamma}(i) - \overrightarrow{D_1}.\overrightarrow{F_\gamma}\end{aligned} \qquad (4)$$

Here, $\overrightarrow{F_\propto}$, $\overrightarrow{F_\beta}$, and $\overrightarrow{F_\gamma}$ are computed by Eqn. (5)

$$\begin{aligned}\overrightarrow{F_\propto} &= \left|\overrightarrow{Y_1} \times \overrightarrow{Q_\propto}(i) - \vec{Q}\right|\\ \overrightarrow{F_\beta} &= \left|\overrightarrow{Y_2} \times \overrightarrow{Q_\beta}(i) - \vec{Q}\right|\\ \overrightarrow{F_\gamma} &= \left|\overrightarrow{Y_3} \times \overrightarrow{Q_\gamma}(i) - \vec{Q}\right|\end{aligned} \qquad (5)$$

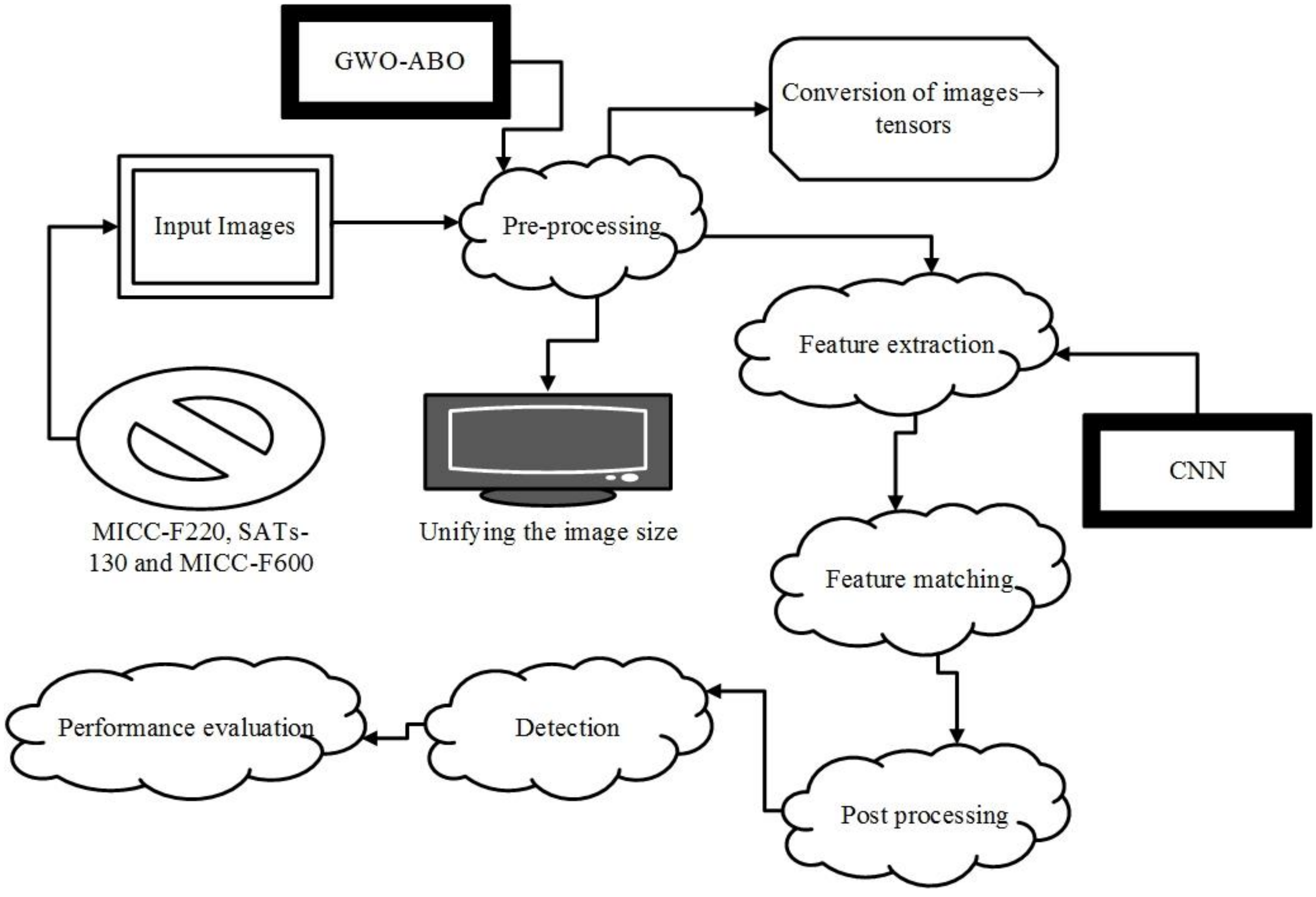

Fig. 1. Presented methods workflow

African buffalos' three principal characteristics are extensive memory capacity, cooperative cum communicative ability in both good and bad times, and extreme intellect gives origin to a democratic character. There are two sounds: '*maaa*' and '*waaa*', which are indicated by $P_k$ and $S_k$, respectively. The buffalo's movement is determined by Eq. (6).

$$P_k + 1 = P_k + V_1\big(\vec{Q}(i+1) - S_k\big) + V_2(d_u.k - S_k) \quad (6)$$

Here, the input image is denoted as $S_k$, $P_k$ is the image size, $k$ represents the iteration, the learning factors are represented as $V_1$ and $V_2$, the wolf's best fitness is represented as $\vec{Q}(i+1)$ and $d_u.k$ is buffalo's best in each iteration. The image → tensors process is done by Eq. (7).

$$S_k + 1 = {(S_k + P_k)}\big/_{\pm 0.5} \qquad (7)$$

*2) Feature extraction:* Three convolutional layers, typically preceded by a pooling layer, make up the suggested feature extraction method. A max pooling layer comes after a convolutional layer that combines 32 filters in total. In complement to the ultimate combo of a convolutional layer with 128 filters and the final max pooling layer, a convolutional layer with 64 filters as well as a max pooling layer are also used. A feature map, which reflects the input image, has been produced by this process. The layers in extracting features from input images are shown in Fig. 2.

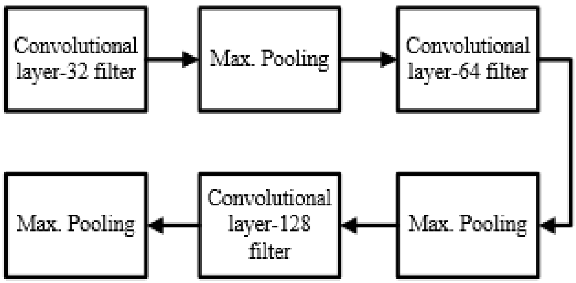
Convolutional layer-32 filter
Max. Pooling
Convolutional layer-64 filter
Max. Pooling
Convolutional layer-128 filter
Max. Pooling

Fig. 2. Layers in feature extraction

The convolutional and pooling layer sequencing creates a feature vector out of the feature map. The CNN structure is shown in Fig. 3. The input for the following step is this feature vector.

*a) Convolutional layer:* A series of 2D digital filters make up the extracting features layer known as the convolutional layer. A convolution layer was used to minimise the data variables and extract the important features. The convolution layer includes scale invariance, interpretation invariance, and rotation invariance. Both the over-fitting problem is lessened and the generalisation idea is added to the fundamental framework. Eq. (8) illustrates the input of the convolutional layer like a collection of GWO-ABO pre-processed images.

$$h_k^l = f\left(\sum_{j\epsilon Y_k} h_k^{m-1} * R_{ki}^l + Y_k^m\right) \qquad (8)$$

Here, $h_k$ is the given map cluster; $R_{kj}^m$ is the convolution kernel, which has been employed for joining the $i^{th}$ input feature map with $k^{th}$ output feature map; $Y_k^m$ defines the consistent bias to the $i^{th}$ input feature map, the activation function is denoted as $f$.

$$\delta_k^m = \delta_k^{m+1} C_k^{m+1} \times (s^m) = \beta_k^{m+1} up(\delta_k^{m+1}) \times (s^m) \qquad (9)$$

Here, $m+1$ describes the pooling layer; the convolution kernel is denoted as $C$. The error function’s partial derivatives convolution kernel and cost is described in Eq. (10) and (11)

$$^{\partial G}/_{\partial R_{ki}^m} = \sum_{v,n} (\delta_k^m) v, n (A_k^{m-1}) v, n \qquad (10)$$

$$^{\partial G}/_{\partial Y_k^m} = \sum_{v,n} (\delta_k^m) \mathrm{v, n} \qquad (11)$$

Where, $(A_k^{m-1}) v, n$ is $h_k^{m-1}$ patch for every convolution and $B_{kj}^m\ (v, n)$ is the patch centre. Lower random weights are used to initialize these filter's values. Then, during the training phase, these weights' values have been modified. The characteristics from the input images should be extracted by the filters that have been used. A feature map with an input image's filtered copies depth is produced by this procedure. A pixel's updated value $(P_n)$ is equal to the sum of its former surrounding pixels' $(p)$ values multiplied by the filtration elements $(g)$ that have been used. The calculation is shown in Eq (12).

$$P_n = \sum_{j \in w} P_j \times g_j \qquad (12)$$

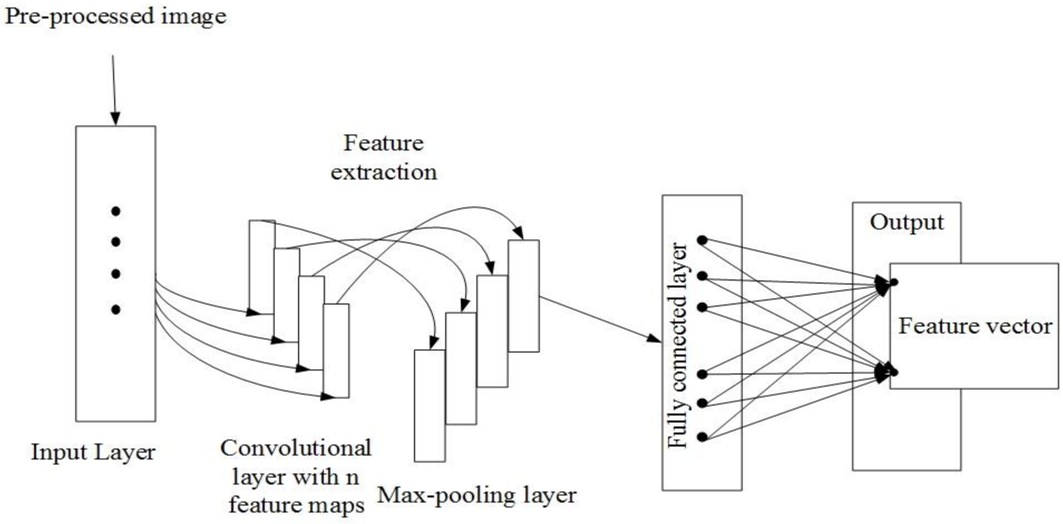

Fig. 3. CNN structure

*b) Pooling layer:* The convolutional layer's feature map's dimensions are reduced using the pooling layer, a form of feature reduction technique. Windows are created using the split input feature map. Every window is divided into its highest value from its contained values by the max pooling method. In contrast, the mean pooling divides the windows into the average value of the window's contained values. The suggested approach has implemented the max pooling layers.

*3) Feature matching:* To detect forgeries after obtaining the feature matrix, relevance searching has been carried out. In order to expedite the matching process, the feature matrix has been first lexicographically processed for this reason. Following that, the Euclidean distance used to compare vector similarities is provided. By using the provided Eq. (13) to contrast the vectors distances with a predefined threshold $\delta$, the matching vectors have been identified.

$$f_k^i = (f_1, f_2, \dots, f_{10}), \sqrt{\sum_{k=1}^{10}\left(f_i - f_j\right)^2} \le \delta \quad (13)$$

Where, $\delta$ = 1.5. The Euclidean distance between the matched blocks is used to check similarities in the following stage of matching; it must be larger than threshold $\beta$ to prevent false matches.

*4) Post processing:* First, potential erroneous matches have been removed in this stage. The shift vector between identical blocks has been computed for this reason. Shift vectors have been produced from the suspect pairs' top left coordinates. Additionally, it is assessed whether the quantity of blocks with the similar shift vector surpasses a predefined threshold value ($\gamma = 32$). If this criterion is met, it is demonstrated that CMF has finished handling the relevant blocks. The CMF on digital photos is finally discovered.

## IV. Results and Discussion

The assessment of the suggested algorithm's findings was reported in the results section. A comparative study with earlier methods working with CMF recognition was conducted after findings listing and discussion. A computer system running Windows 10 with an Intel Core i7 8th generation CPU, 4 GB of GPU hardware supporting CUDA, processor-64-bit, and RAM 8 GB was utilised to construct the suggested method. In additional to Keras and TensorFlow, Python 3.5 software tools were used to implement the suggested method's backend toolkits.

### A. Evaluation Metrics

The major metrics for evaluating the DL model's detection performance are: F-measure, TPR, accuracy, FNR, FPR, TNR, and TT. These metrics were employed for evaluating the presented GWO-ABO-CNN models performance in CMF recognition.

True negative ($T_n$) represents the number of real images that were really identified as real images. False negative ($F_n$) refers to the number of altered images that were mistakenly identified as real images. True positive ($T_p$) represents the number of manipulated images that were accurately identified and false positive ($F_p$) represents the number of real images that were mistakenly identified as altered images.

*1) Accuracy:* Accuracy has been referred to as the percentage of correctly detected image pixels. Absolute pixel precision is another phrase for this. Despite being the most fundamental performance indicator, anytime there is a class disagreement, it may lead to erroneous image detection findings. When one recognised category outperforms another, there is a category disparity. In this scenario, the dominant class's superior accuracy will outweigh the other group's poorer accuracy, leading to biased findings. If there has been no group discrepancy, the accuracy measure was advised for evaluating detection results using images.

Accuracy is described in Eq. (14),

$$Accuracy = {}^{T_p + T_n}\big/_{T_p + F_p + T_n + F_n} \quad (14)$$

*2) True positive rate (TPR):* TPR, or the proportion of images which have been accurately identified as CMF, relates to the capacity to positively detect an instance of forgery. Moreover, sensitivity is described as the proportions of predicted forgery among those images are real.

TPR is described in Eqn. (15),

$$TPR = {}^{T_p}\big/_{T_p + F_n} \quad (15)$$

*3) True negative rate (TNR):* The chance that an image which does not have any forgery will have a negative test outcome is known as TNR, and it is the proportion of forgery which is appropriately detected as CMF. Moreover, specificity is described as the proportions of predicted negative outcomes among those images are tested as original.

TNR is described in Eq. (16),

$$TNR = {}^{T_n}\big/_{T_n + F_p} \quad (16)$$

*4) False positive rate (FPR):* The FPR is determined by dividing the total quantity of forgery images by the number of forgery images that were mistakenly classified as original images (false positives).

TNR is described in Eqn. (17),

$$FPR = {}^{F_p}\big/_{T_n + F_p} \quad (17)$$

*5) False negative rate (FNR):* The likelihood that a true positive would be overlooked by the sample is measured by the FNR, also known as the miss rate. The FNR seems to be the percentage of positive test results that result in failed tests. The FNR has been computed as the difference between the number of altered images that were incorrectly identified as true images and the total number of altered images that were incorrectly identified as altered images as well as the number of altered images that were correctly identified as altered images.

FNR is described in Eq. (18),

$$FNR = {}^{F_n}\big/_{T_p} + F_n \qquad (18)$$

*6) F-measure:* The effectiveness of the procedures is demonstrated through a pixel-based assessment that involves determining whether or not each individual pixel is false. The proposed as well as considered approaches are evaluated in terms of performance using the F-measure. Its improved performance in identifying the CMF in digital images is shown by a higher F-measure outcome.

F-measure is described in Eq. (19),

$$F - measure = 2 \times {}^{T_p}\big/_{2} \times (F_p + F_n + T_p) \qquad (19)$$

*7) Testing time (TT):* Testing Time (TT) has been further employed to assess the presented model and compare it to other models. TT seems to be the average amount of time required to evaluate the images for the specified number of runs (k). Additionally, as this step has only been completed offline once, Learning Time has been not taken into account.

## B. Results

The research introduced a novel DL (CNN) method for the procedure of detecting image forgery. The suggested model combines a CNN model with a hybrid optimization of GWO and ABO. The suggested DL model has been tested using the MICC-F600, MICC-F220, and SATs-130 databases, among other datasets. The goal of the computational experiments has always been to develop the best model possible in terms of complexity and TT. Moreover, the suggested method had been tested using the k-fold cross-validation method. This framework starts with a learning phase that will be repeated (k) times to get a diversity of the images being looked at and provide accurate estimation by thoroughly looking through the datasets. The dataset is further divided into (k) groups (folds) with almost the similar dimension in this approach.

The suggested model employed (k-1) the residual categories for testing and the classes for training. For both testing and training there are (k) iterations. The suggested approach made advantage of a method for 5-fold cross-validation. Accordingly, for each of the five rounds, 30% of the dataset's images were utilised for testing while the remaining 70% were randomly chosen for training. Each iteration will employ a different 30% of the images instead of the previous 30% of images for assessment.

TABLE II. PRESENTED MODEL PERFORMANCE @ 25 EPOCHS

| **Metrics** | **Datasets** | | |
|---|---|---|---|
| | ***MICC-F600*** | ***SATs-130*** | ***MICC-F220*** |
| Accuracy (%) | 95 | 92 | 97 |
| TPR (%) | 90.4 | 84 | 100 |
| F-measure (%) | 93.98 | 88.79 | 91.55 |
| FNR (%) | 28.5 | 6.7 | 0 |
| TT (sec) | 1.23 | 5.02 | 1.24 |
| TNR (%) | 85.19 | 77.5 | 98.2 |
| FPR (%) | 5.7 | 10.4 | 10.8 |

Five trials at training epochs of 25, 75, 50, and 100 make up the test. At each epoch, the TPR, FNR, F-measure, TNR, FPR, and accuracy are recorded. In the computational experiments, a recording of the TT has been also taken into account. The simulation outcomes for every epoch for all databases are presented in Tables II to V.

TABLE III. PRESENTED MODEL PERFORMANCE @ 50 EPOCHS

| **Metrics** | **Datasets** | | |
|---|---|---|---|
| | ***MICC-F600*** | ***SATs-130*** | ***MICC-F220*** |
| Accuracy (%) | 97.33 | 93.04 | 91.45 |
| TPR (%) | 97.5 | 93.97 | 100 |
| F-measure (%) | 93.4 | 88 | 82 |
| FNR (%) | 34.4 | 1.6 | 0 |
| TT (sec) | 1.31 | 4.25 | 1.25 |
| TNR (%) | 87 | 93 | 97 |
| FPR (%) | 10.1 | 5.6 | 4.3 |

It is evident that the implementation of the suggested model grows as the number of epoch’s increases. Additionally, how the suggested technique is implemented varies from one database to another. Therefore, the suggested model cannot be implemented due to two limits. The first constraint seems to be the number of training epochs, and the other is the databases construction, based on which the suggested model has been used.

TABLE IV. PRESENTED MODEL PERFORMANCE @ 75 EPOCHS

| **Metrics** | **Datasets** | | |
|---|---|---|---|
| | ***MICC-F600*** | ***SATs-130*** | ***MICC-F220*** |
| Accuracy (%) | 94.4 | 97.8 | 98.7 |
| TPR (%) | 100 | 97.5 | 100 |
| F-measure (%) | 98.2 | 98.2 | 99.05 |
| FNR (%) | 9.7 | 0 | 0 |
| TT (sec) | 1.23 | 4.14 | 1.27 |
| TNR (%) | 97.6 | 96.25 | 98.54 |
| FPR (%) | 1.7 | 5.4 | 3.2 |

The effectiveness metric variable findings for the various four databases at epochs 25, 50, 75, and 100 are shown in Tables II, III, IV, and V, correspondingly. The findings shown in the tables show that after 100 epochs, F-measure, TPR, accuracy, and TNR increase by over 100%. In contrast, FNR and FPR decreased almost to zero at the same time that the number of epochs reached 100. Additionally, as the number of epochs increased, the TT decreased to a minimum.

TABLE V. PRESENTED MODEL PERFORMANCE @ 100 EPOCHS

| Metrics | Datasets | | |
|---|---|---|---|
| | *MICC-F600* | *SATs-130* | *MICC-F220* |
| Accuracy (%) | 98.2 | 100 | 100 |
| TPR (%) | 98.6 | 99 | 100 |
| F-measure (%) | 99.9 | 100 | 100 |
| FNR (%) | 2.2 | 0 | 0 |
| TT (sec) | 1.24 | 1.32 | 1.132 |
| TNR (%) | 100 | 98.05 | 100 |
| FPR (%) | 1.04 | 2.84 | 0 |

It is clear from the previous findings that the number of training components has a significant impact on the performance of the entire model. Additionally, it is noted that 100 epochs yield the greatest results. The suggested model included a number of identifiers that might affect the results in terms of accuracy, F-measure, or TT. These variables include the volume of data utilized for training (which is connected to the utilized databases dimension), the volume of data employed as inputs to the network, the volume of hidden prototype layers, and the volume of the chosen epochs. The Identifiers for how much data are employed for training, how much data is employed as input, and how much data is employed for the network's hidden layers all remain constant. The number of the chosen epochs seems to be the only identifier left that might affect testing time.

TABLE VI. PRESENTED MODEL'S OPTIMAL PERFORMANCE

| Datasets | Metrics | | |
|---|---|---|---|
| | *Accuracy (%)* | *F-measure (%)* | *TT (sec)* |
| MICC-F600 | 98.2 | 99.9 | 1.24 |
| SATs-130 | 100 | 100 | 1.32 |
| MICC-F220 | 100 | 100 | 1.132 |

It was discovered that the number of epochs chosen has a significant effect on how long the method takes to reach the optimal state feature map. The method can obtain the best performance by taking the time necessary to produce the appropriate feature map. The optimum feature map would be retrieved by choosing a precise number of epochs. Along with the number of epochs, interpolation will occur. In addition, the model will perform worse if fewer epochs than the chosen number are used, also failed to obtain the optimal feature map. As a result, if the other identifiers have been kept constant, the number of chosen epochs affects the TT. The optimal performance of proposed model in CMF detection on different databases is shown in Table VI.

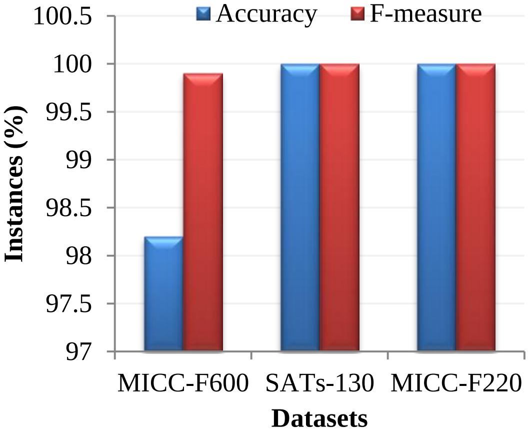

Fig. 4. Different dataset's optimal accuracy and F-measure of detection

The best accuracy and F-measure representation for each database is shown in Fig. 4. According to the simulation findings, the suggested algorithm is a valid and practical method for CMF detection. Moreover, the presented model's MICC-F220 dataset's accuracy was compared with other existing models like Speeded Up-Robust Feature (SURF) [8], Spatial features-based Image CMF detection (IC-MFDs) [23], CNN [24], and Convolutional Long-Short Term Memory (ConvLSTM) [25].

TABLE VII. MICC-F220 DATABASE ACCURACY COMPARISON

| Method | Accuracy (%) |
|---|---|
| SURF | 91.95 |
| IC-MFDs | 98.44 |
| CNN | 100 |
| ConvLSTM | 100 |
| Presented (GWO-ABO-CNN) | 100 |

The MICC-F220 dataset accuracy comparison of different models is shown in Table VII and Fig. 5. The comparison outcome indicated that the CMF detection performed by CNN model has attained 100% accuracy.

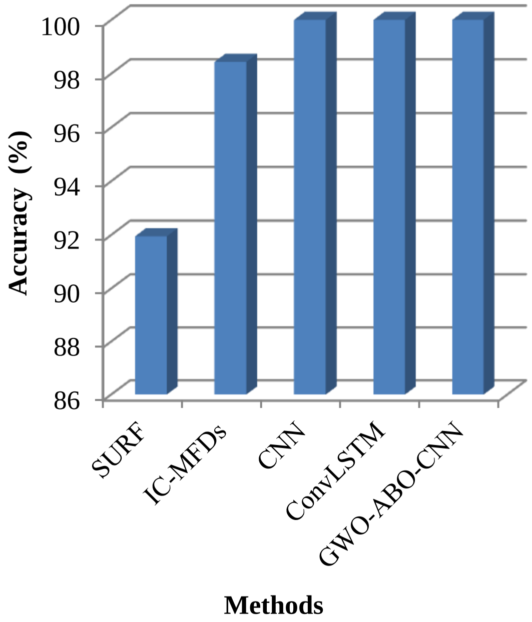

Fig. 5. Comparison of accuracy

The accuracy attained by the model SURF was 91.95%, IC-MFDs was 98.44%, CNN was 100%, ConvLSTM was 100%, and the presented GWO-ABO-CNN was 100%. The models with higher accuracy are reliable for CMF detection. The main problem in the CMF recognition is less accuracy on challenging database. From the current research results, it is evident that the presented GWO-ABO-CNN method detects CMF in challenging datasets like SATs130, MICC-F220, and MICC-F600 with 100% accuracy. Thus, the presented method is reliable for CMF recognition in digital images.

## V. Conclusion

This research presents a deep learning system for CMF detection predicated on hybrid optimization GWO-ABO with CNN (GWO-ABO-CNN). The main goal of this study is to construct and enhance the DL classification framework for identifying real and forged classes in alleged digital image forgeries. The suggested approach anticipates creating a new paradigm that will offer improved performance, demand less testing time, and incur little computational expense. Four layers—pre-processing, feature extraction, feature matching, and post-processing (detection)—are addressed in the creation of the proposed forgery detection method. GWO-ABO has been used to unify the picture size and turn the images into tensors in the pre-processing layer. The suggested method is predicated on a novel invention that builds a serial series of convolutional layers and a pooling layer to speed up the detection process in the feature extraction layer. Several challenging datasets, including the SATs130, MICC-F220, and MICC-F600 datasets, have been employed in the evaluation process. The test findings revealed that 100 training epochs had produced the best accuracy.

## References

[1] P. Jain, “Shaina shainawadhwa79@gmail.com Adesh College of Engineering and Technology, Faridkot, Punjab,” p. 6, 2019.

[2] A. H. Saber, M. Khan, and B. Mejbel, “A Survey on Image Forgery Detection Using Different Forensic Approaches,” Adv. Sci. Technol. Eng. Syst. J., vol. 5, pp. 361–370, Jan. 2020, doi: 10.25046/aj050347.

[3] W. D. Ferreira, C. B. R. Ferreira, G. da Cruz Júnior, and F. Soares, “A review of digital image forensics,” Comput. Electr. Eng., vol. 85, p. 106685, Jul. 2020, doi: 10.1016/j.compeleceng.2020.106685.

[4] Vivek Singh, Neelesh Kumar Jain, and Jaypee University, guna(M.P.), “Digital Image Forensics in Multimedia Security: A Review,” Int. J. Eng. Res., vol. V4, no. 05, p. IJERTV4IS051057, May 2015, doi: 10.17577/IJERTV4IS051057.

[5] A. Ferreira, T. Carvalho, F. Andaló, and A. Rocha, “Counteracting the contemporaneous proliferation of digital forgeries and fake news,” An. Acad. Bras. Ciênc., vol. 91, no. suppl 1, p. e20180149, 2019, doi: 10.1590/0001-3765201820180149.

[6] M. A. Elaskily et al., “A novel deep learning framework for copy-moveforgery detection in images,” Multimed. Tools Appl., vol. 79, no. 27–28, pp. 19167–19192, Jul. 2020, doi: 10.1007/s11042-020-08751-7.

[7] H. Chen, X. Yang, and Y. Lyu, “Copy-Move Forgery Detection Based on Keypoint Clustering and Similar Neighborhood Search Algorithm,” IEEE Access, vol. 8, pp. 36863–36875, 2020, doi: 10.1109/ACCESS.2020.2974804.

[8] A. Badr, A. Youssif, and M. Wafi, “A Robust Copy-Move Forgery Detection In Digital Image Forensics Using SURF,” in 2020 8th International Symposium on Digital Forensics and Security (ISDFS), Beirut, Lebanon, Jun. 2020, pp. 1–6. doi: 10.1109/ISDFS49300.2020.9116433.

[9] S. Dhivya, J. Sangeetha, and B. Sudhakar, “Copy-move forgery detection using SURF feature extraction and SVM supervised learning technique,” Soft Comput., vol. 24, no. 19, pp. 14429–14440, Oct. 2020, doi: 10.1007/s00500-020-04795-x.

[10] H. Huang, W. Guo, and Y. Zhang, “Detection of Copy-Move Forgery in Digital Images Using SIFT Algorithm,” in 2008 IEEE Pacific-Asia Workshop on Computational Intelligence and Industrial Application, Wuhan, China, Dec. 2008, pp. 272–276. doi: 10.1109/PACIIA.2008.240.

[11] A. Parveen, Z. H. Khan, and S. N. Ahmad, “Block-based copy–move image forgery detection using DCT,” Iran J. Comput. Sci., vol. 2, no. 2, pp. 89–99, Jun. 2019, doi: 10.1007/s42044-019-00029-y.

[12] K. B. Meena and V. Tyagi, “A copy-move image forgery detection technique based on Gaussian-Hermite moments,” Multimed. Tools Appl., vol. 78, no. 23, pp. 33505–33526, Dec. 2019, doi: 10.1007/s11042-019-08082-2.

[13] M. Kumar, A. Rani, and S. Srivastava, “Image Forensics Based on Lighting Estimation,” Int. J. Image Graph., vol. 19, no. 03, p. 1950014, Jul. 2019, doi: 10.1142/S0219467819500141.

[14] H. Kasban and S. Nassar, “An efficient approach for forgery detection in digital images using Hilbert–Huang transform,” Appl. Soft Comput., vol. 97, p. 106728, Dec. 2020, doi: 10.1016/j.asoc.2020.106728.

[15] H. Wang and H. Wang, “Perceptual Hashing-Based Image Copy-Move Forgery Detection,” Secur. Commun. Netw., vol. 2018, pp. 1–11, 2018, doi: 10.1155/2018/6853696.

[16] A. Kuznetsov, “Digital image forgery detection using deep learning approach,” J. Phys. Conf. Ser., vol. 1368, no. 3, p. 032028, Nov. 2019, doi: 10.1088/1742-6596/1368/3/032028.

[17] R. Shao and E. J. Delp, “Forensic Scanner Identification Using Machine Learning,” in 2020 IEEE Southwest Symposium on Image Analysis and Interpretation (SSIAI), Albuquerque, NM, USA, Mar. 2020, pp. 1–4. doi: 10.1109/SSIAI49293.2020.9094618.

[18] N. Kanwal, A. Girdhar, L. Kaur, and J. S. Bhullar, “Detection of Digital Image Forgery using Fast Fourier Transform and Local Features,” in 2019 International Conference on Automation, Computational and Technology Management (ICACTM), London, United Kingdom, Apr. 2019, pp. 262–267. doi: 10.1109/ICACTM.2019.8776709.

[19] H.-Y. Huang and A.-J. Ciou, “Copy-move forgery detection for image forensics using the superpixel segmentation and the Helmert transformation,” EURASIP J. Image Video Process., vol. 2019, no. 1, p. 68, Dec. 2019, doi: 10.1186/s13640-019-0469-9.

[20] 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 Trans. Inf. Forensics Secur., vol. 6, no. 3, pp. 1099–1110, Sep. 2011, doi: 10.1109/TIFS.2011.2129512.

[21] V. Christlein, C. Riess, J. Jordan, C. Riess, and E. Angelopoulou, “An Evaluation of Popular Copy-Move Forgery Detection Approaches,” IEEE Trans. Inf. Forensics Secur., vol. 7, no. 6, pp. 1841–1854, Dec. 2012, doi: 10.1109/TIFS.2012.2218597.

[22] 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 Process. Image Commun., vol. 28, no. 6, pp. 659–669, Jul. 2013, doi: 10.1016/j.image.2013.03.006.

[23] I. T. Ahmed, B. T. Hammad, and N. Jamil, “Image Copy-Move Forgery Detection Algorithms Based on Spatial Feature Domain,” in 2021 IEEE 17th International Colloquium on Signal Processing & Its Applications (CSPA), Langkawi, Malaysia, Mar. 2021, pp. 92–96. doi: 10.1109/CSPA52141.2021.9377272.

[24] M. A. Elaskily et al., “A novel deep learning framework for copy-moveforgery detection in images,” Multimed. Tools Appl., vol. 79, no. 27–28, pp. 19167–19192, Jul. 2020, doi: 10.1007/s11042-020-08751-7.

[25] M. A. Elaskily, M. H. Alkinani, A. Sedik, and M. M. Dessouky, “Deep learning based algorithm (ConvLSTM) for Copy Move Forgery Detection,” J. Intell. Fuzzy Syst., vol. 40, no. 3, pp. 4385–4405, Mar. 2021, doi: 10.3233/JIFS-201192.