Enhanced Image Forgery Detection In Digital Images Through Advanced Yolo And Rcnn-Based Deep Learning Techniques

Navid Ahmad Dar[1*], Rani Astya[2], Anuj Kumar[3]

[1*]Department of Computer Science and Engineering, Sharda University

[2]Associate Professor, Department of Computer Science and Engineering, Sharda University, Greater Noida U. P, India
[3]Professor, Department of Computer Science and Engineering, I. T. S Engineering College, Greater Noida U. P, India

Abstract
Fake image detection is crucial these days, because any image can be altered easily. In this regard, this research proposes a simpler yet powerful approach that employs state-of-the-art Deep Learning (DL) techniques along with traditional image processing. Combining the powers of You Only Look Once (YOLO) and Recurrent Neural Network (RCNN) for object detection, Discrete Cosine Transform (DCT) for fast frequency domain analysis, and the Watershed algorithm for accurate image segmentation. This study presents a fine blend of traditional techniques with DL, providing better results than the single use of either of the two mentioned above. The proposed technique has been empirically tested on three public standard datasets: CASIA 2.0, CoMoFoD, and MICC-F2000. The proposed technique achieves an accuracy of 98%, 99%, and 97%, respectively. The superior results prove that this technique is not only strong but also reliable; thus, it fulfills the demand for its potential use in the real world for digital image forensics and security. Accordingly, this work presents a full framework that merges modern DL approaches with traditional image processing to accurately detect forged areas in digital photos and meets an important demand of today's digital world.

Keywords: Image Forgery Detection, DL Techniques, YOLO, RCNN, Digital Image Forensics

1. Introduction

With the dawn of the multimedia era, digital images became easy to manipulate due to the widespread accessibility of image-editing tools. As a result, fabricated images have become pervasive, with a particularly serious concern being the creation of manipulated photographs intended to influence public opinion [1]. Digital image manipulation means changing an image's content in order to achieve harmful or fraudulent goals. Digital forensics refers to these kinds of alterations as forgeries. Faked photos have existed for as long as there have been photographs. Images that have been digitally altered date back to the 1860s. In the past, image forgeries have damaged many people and businesses by adding unnecessary complexity to their lives [2,3]. Image editing programs have become so accessible that photo altering and forgeries are commonplace. For example, in copy-move forgeries, the content of one image is hidden by being copied and pasted onto a virtually identical part of another image [4]. This could be done to modify the image by adding or removing elements that were originally visible [5]. Figure 1 describes the category of Image Forgery Detection Methods.

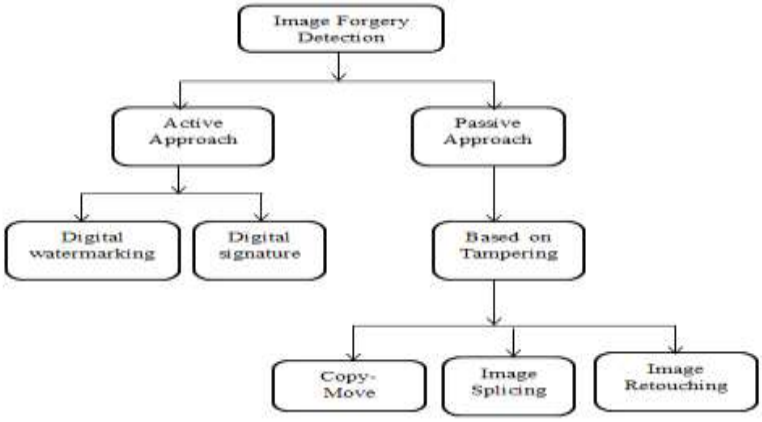


Figure 1. Classification of Image Forgery Detection Techniques [6]

- ***Active Approach:*** The active technique records digital images with pre-inserted information like a watermark or signature. However, it would set a limit to the active approach during exercise. However, many cameras lack these features, leading to active approach failure [7].
- ***Passive Approach:*** A blind approach to image forgery detection has gained popularity among researchers due to its ability to detect forgeries without keeping hidden data in the digital image. These methods verify digital images by analyzing changes in image attributes during manipulation. Several methods can locate faked regions in digital images by detecting forgery and tampering traces [8].

Techniques for Identifying Forgery in Digital Images:

- ***Image Retouching:*** This method of forgery is relatively harmless in comparison to others since the final fabricated image is barely distinguishable from the original, with some modest improvements [9].
- ***Image Splicing:*** This forgery approach is more harmful than retouching and much easier to perform. The image is created by combining components or objects from many photographs into one [10].
- ***Copy-Move Image Forgery:*** This forging technique is challenging and widely utilized. The image is forged by extracting a certain region of the unusual image and reusing it in a similar area to hide or add features. Thus, it radically alters the information communicated by the image [11].

Figure 2 displays an example of a copy-move image forgery in which 2 (a) shows the real image, and 2 (b) shows the copy-move image.



Figure 2. A simple case of copy-move forgery: (a) Original image (b) Copy-move image [12]

The Copy-Move Forgery Detection (CMFD) [13] has remained one of the central focuses in image forensics research. Examples of copy-move forgery involve at least two regions in an image that have the same content [14]. Basically, it involves obscuring or augmenting existing content by duplicating a segment and inserting it into another part of an image [15].

That is, DL-based methods now dominate a host of computer vision applications. In particular, one of the most successful classes of DL models, Convolutional Neural Network (CNNs), has already shown impressive performance in the case of complex visual input processing [16]. This inherent capability of deep neural networks to learn discriminative features automatically provides adequate discrimination between authentic photographs and those that have undergone any alteration. Considering their proficiency in learning minute patterns and changes, DL models are aptly suited for handling every nuance of image forgery detection, including copy-move manipulations [17].

This work enhances digital image forensics by detecting copy-move forgery using a DL approach. It leverages the DL models designed for image analysis with enhanced capabilities for forgery detection. That makes this technique more robust compared with most of the traditional forgery techniques, since it utilizes deep neural networks in order to find duplicated content. This

enhances the domain of digital image forensics and aids in constructing more robust tools to guard against image tampering in criminal investigations, authentication, and content verification.
Section 1 provides an overview of the concept, Section 2 describes the most relevant contributions from many researchers, Section 3 describes the research challenges, and Section 4 specifies the datasets used in this work. Section 5 describes how the research approach is performed, specifying the employed technique, proposed method, and suggested algorithm. Section 6 presents the results, whereas Section 7 offers the conclusion and possible future implications.

2. Related Work

In this section, some associated work based on Image Forgery Detection in Digital Images is discussed below:

2.1 Copy-Move Forgery in Digital Images

Copy-move forgeries in digital photos entail duplicating a segment of an image and repositioning it. Somewhere else in the same image, to either duplicate or hide something. Detection thus seeks duplication of areas, which can be challenging in cases where the copied part has been rotated or scaled.
Tahaoglu et al. (2022) [18] suggested a way to detect copy-move forgery in digital images in real-time. The method focused on the features of the image's texture. This method demonstrated improved performance and resistance to several attacks on the geometry, such as rotation and scaling. Dhanaraj et al. (2021) [19] conducted a considerable examination on the methods utilized to identify copy-move forgeries. Two major types were key point-focused and block-focused. The key point-focused approaches were found to be fairly successful. Muzaffer et al. (2019) [20] proposed detecting and localizing copy-move fraud through DL as an alternative to traditional techniques based on feature extraction methods. Gurbuz et al. (2019) [21] obtained successful results in their study regarding the detection of counterfeited images, even though the images were subjected to Gaussian filtering and JPEG compression parameters, which made attempts to detect counterfeited images even more difficult. Meena et al. (2020) [22] advanced copy or move forgery detection by addressing the weaknesses of block-based and keypoint-based techniques in isolation. A hybrid Scale Invariant Feature Transform (SIFT) for textured areas and Fourier–Mellin Transform (FMT) for smooth areas were employed, and, under different geometric and post-processing attack scenarios, the detection accuracy was superior to that of the current state of the art. Yang et al. (2020) [23] presented a feature-based approach to CMFD. Using a modified SIFT detector, key points were detected. Key Points were systematically allocated to the image. Using the modified SIFT descriptor, locations of copy-move forgery detection were found. Comprehensive experimental findings verified the efficiency.

2.2 Recent Advances in Forgery Detection

Traditional forgery detection techniques check an image for mistakes or oddities. They consider a number of things, including pixel values, strange noise patterns, and metadata. Many techniques working on blocks or on key points seek duplicates or other mismatches inside the image.
Zhang et al. 2024 [24] identified small-region forgeries in medical Computed Tomography (CT) images where Generative Adversarial Network (GANs) were used, employing a two-stage cascading framework that integrated local sub-image forgeries. It involved the local Gray-Level Co-occurrence Matrix (GLCM) classification. This approach recorded 93.5 percent accuracy with an F1 score equal to 0.933, leading the competition among peer models. Rao et al. 2023 [25] developed a technique for image fraud detection using a combination of ResNet and Transformer decoding. This method demonstrated exceptional performance in identifying splicing manipulations. It achieved a high level of accuracy on three major datasets: CASIA, National Institute of Standards and Technology (NIST), and the Image Manipulation Dataset (IMD2020). Singh et al. (2022) [26] advanced the methodology even further by creating a model that combined the efficiency of EfficientNetB0 and the sentence transformer (for text) to detect counterfeit images in microblogging. It achieved to outperform all other state-of-the-art models with 85.3% accuracy in Twitter and 81.2% accuracy in Weibo.
Krishnaraj et al. (2022) [27] presented a DL-based fusion model for Detecting and Localizing Copy-Move Forgeries (DLFM-CMDFC). This method could improve both detection and localization by fusing GAN with DenseNet. Experimental results proved that the proposed model performs better

than the latest methods. Rathore et al. (2021) [28] represented the Biorthogonal Wavelet Transform with Singular Value Decomposition (BWT-SVD) features. A detection accuracy of 92.22% was achieved for the CoMoFoD dataset. Kasban et al. (2020) [29] provided a method capable of detecting even minute changes in digital images, resistant to attacks relying on image manipulation. A novel approach to identifying forgeries was presented. When compared to Artificial Neural Network (ANN) and K-Nearest Neighbour (KNN) classifiers, the results demonstrated that the Support Vector Machine (SVM) classifier had the maximum accuracy.
Wang et al. (2020) [30] applied the magnitude and the three phases of Quaternion Discrete Cosine Transform (QDCT) to obtain Markov features. The obtained results showed 99.16% detection accuracy on the CASIA-V1 dataset and 97.52% detection accuracy on CASIA-V2 dataset. Ouyang et al. (2019) [31] introduced a method for detecting CMFD cases via a pyramid structure and Zernike moments. The authors reported the method had the most optimal results for any given angle of rotation from 50% to 200%.

2.3 Emergence of DL in Image Forensics

The arrival of DL forensics Image detection, and forgery detection technologies has advanced the field considerably and allowed for the detection of a forgery at a higher level of speed and precision than ever before. To understand this. Speed and accuracy improvements trickle down from a number of advancements and various methodologies.
Detecting and locating image forgery with a tiny ViT was employed by Aryan et al. (2025) [32] developed a model with multi-head self-attention and patch embedding together with a tiny classifier. The model achieved very high performance. It AUC/F1 scores were 0.9712/0.9587 and 0.9993/0.9914 on CASIA v1 and DEFACTO, respectively, outperforming all other models in the literature. Elaskily et al. (2023) [33] developed a hybrid CNN-Conventional Long-Short Term Memory (ConvLSTM) model to find instances of copy-move forgery in digital images with an outstanding accuracy (100%) and a mere testing time of 1 second. The authors' method improved detection compared to previous approaches. A lightweight CNN has been proposed by Hosny et al. (2022) [34] CNN achieved 100 % accuracy on copy-move image forgery detection tasks and has a quick processing time of 0.83 seconds per test. Castillo Camacho et al. (2021) [35] developed image forensics approaches with a specific focus on deep-learning-based methods. According to the findings, the proposed approach decreased the detection accuracy from 98% to 85%. Ding et al. (2020) [36] provided an innovative framework for use in digital forensics as a means of user safety. Its foundation lies in DL, allowing it to classify threats and thereby detect them. Consequently, the DL model's detection performance was enhanced, and 5G's speedy data transmission was realized. With HetNets, authors could rest assured that they always received excellent service.

3. Research Problem

Traditional techniques for the detection of copy-move forgery suffered disappointments when applied to images that had been rotated, resized, or captured in different lighting environments. Otherwise, the majority of DL-based forgery detectors spent an implausibly long time for the purpose of real-time implementations in standard computer settings. In the face of the novel forgeries, these systems malfunctioned, having been over-specialized in particular kinds of copy-move forgery. In addition, the DL models, like many others, had glaring vulnerabilities to adversarial attacks; a few slight modifications to the input images would circumvent the model and cause wrongful classifications. All the above demonstrate the critical need to develop advanced techniques that cover a broader spectrum of forgery detection.

4. Novelty

The novelty of the study is to combine frequency-domain analysis, region-based segmentation, and dual-DL in a seamless, ordered, and systematic manner in a unified framework. By this combination, this research achieves performance on par with the best in the field. With the forgery location and detection in real-time, it integrates DCT-based frequency features and Watershed-based region segmentation with an ensemble of YOLO and RCNN. By using RCNN to slow down on the fine-grained analysis of certain regions, and using YOLO to zoom back out and quickly analyze the whole image, a nice sweet spot is found in the trade-offs of detection speed and accuracy

5. Research Methodology

An analysis of the notion of designed architecture is conducted within the framework of research methodology.

5.1 Dataset Description

In this study, secondary benchmark datasets CoMoFoD, MICC-F2000, and CASIA v2.0 are used to evaluate the effectiveness of the proposed image forgery detection methodology.

(a) CoMoFoD Dataset [37]

CoMoFoD has 5,000 fake photos in 25 different categories. The base category contains 200 known forgeries, while the additional 24 categories are generated by applying various transformations to the base category photos to mask forgery signs. The CoMoFoD dataset contains additional kinds of attacks, such as distortion, post-processing attack, and colour reduction, among others; these kinds of attacks were not present in the earlier datasets. It is made up of two distinct kinds of datasets, namely, a dataset with a tiny image size and a dataset with a large image size. Each dataset includes the original photos together with the post-processing treatment, and there are a total of 10,000 photographs in the collection. Despite this, academics have not made use of this dataset because it is of such a massive size.

(b) MICC-F2000 Dataset [38]

The MICC-F2000 dataset has a total of 2000 photos, 700 of which have been edited and 1300 of which are original. The digital photographs included in the MICC-F2000 dataset could have been altered using copy-move forgery, in which a small portion of a image is copied and pasted elsewhere to make it look original. This dataset might be used as a standard against which to measure how effective DL techniques are in spotting forged images.

(c) CASIA-v2 Dataset [39]

In this sample of 12,614 images, 7,491 are high-quality originals, whilst 5,123 are presumably forgeries. All these counterfeit photographs have undergone post-processing procedures to enhance their believability. It was originally produced for research on issues related to photo splicing or copy-move fraud. The collection comprises photos with dimensions varying from 240 pixels on the shorter side to 900 pixels on the longer side. Both JPEG and Tagged Image File Format (TIFF) formats are available, with JPEG utilized for compressed images.

5.2 Data Pre-processing

The main goal of this step is to provide a consistent, high-quality input to the intended classification model. For example, normalizing and resizing the images to the same portion, contrast stretching, and reducing noise would improve the visibility of the finer textures and edges. This would further enhance the quality of the model because of the better segmentation precision.

5.3 Data Augmentation

Data augmentation is applied only to training data, after pre-processing, for better generalization by the model. Rotating images by a small angle makes the model robust against the forged regions that may appear at a different angle, whereas changing brightness and contrast makes it robust under changing lighting conditions. Mild JPEG compression enables the model to detect forgeries in compressed images. Together, these augmentation methods enhance generalization while preserving essential forgery-related features.

5.4 Technique Used

DCT, Watershed, RCNN, and YOLO are the two methodologies that are utilized in this process.

- ***DCT***

DCT is the most common method, where the image is described in terms of cosine frequencies. In this study, apply DCT to extract frequency-based features from images since the forged areas usually introduce minor frequency problems that cannot be easily viewed in the normal, pixel-based representation. Such a transformation to the frequency domain by DCT helps in locating

manipulation artifacts more easily and also enhances the subsequent steps of segmentation and detection. The 2-D DCT is defined as:

$$F(u,v) = \sum_{x=0}^{N-1} \sum_{y=0}^{N-1} f(x,y) \cos\left(\frac{(2x+1)u\pi}{2N}\right) \cos\left(\frac{(2y+1)v\pi}{2N}\right) \quad (1)$$

where $F(u,v)$ represents the DCT coefficients capturing frequency information, and (u, v) denote the frequency indices.

- ***Watershed***

The Watershed algorithm is a region-based image segmentation method inspired by topography. It considers different heights as pixel brightness. Feature maps generated in this study utilize the Watershed to split the DCT-based feature map into meaningful regions, allowing it to sharply delineate possible forgery areas from the background. By computing the gradient magnitude of the image, it finds edges between the regions and spots suspicious areas for deeper study. For instance, the gradient image is defined as:

$$G(x,y) = \sqrt{\left(\frac{\partial I}{\partial x}\right)^2 + \left(\frac{\partial I}{\partial y}\right)^2} \quad (2)$$

where $I(x,y)$ represents the input image and $G\sout{(x,y)}(x,y)$ highlights intensity transitions used by the Watershed transform to perform precise region segmentation. Figure 3 demonstrates the image segmentation using the watershed transformation.

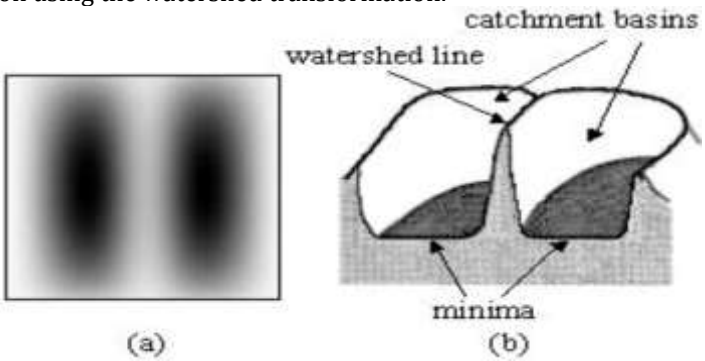


Figure 3. Image Segmentation using Watershed transformation [40].

- ***RCNN***

RCNN is a DL model used in object detection, essentially checking through different regions of the image for objects [41]. In this study, RCNN looks through the Watershed parts and spots suspicious areas that might have been forged. Now, RCNN is appropriate because it carefully extracts features from each candidate region to accurately spot the copy-move or splicing forgeries. The RCNN classification probability for a region r_i is defined as

$$P(y \mid r_i) = \text{softmax}(W \cdot \phi(r_i) + b) \quad (3)$$

where $\phi(r_i)$ represents CNN-extracted features, W and b are learnable parameters, and y denotes the class label (real or forged). Figure 4 depicts the RCNN Architecture used for object detection.

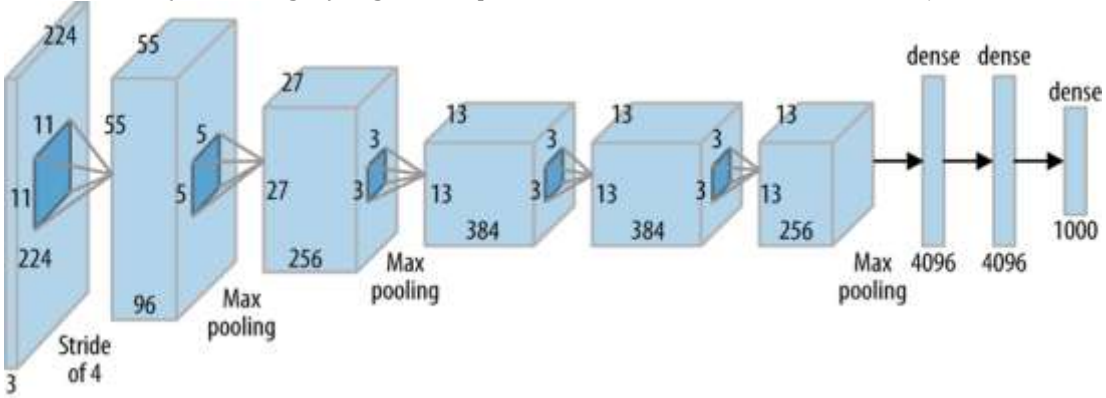


Figure 4. RCNN Architecture (Object Detection) [42]

- ***YOLO***

YOLO is an object detection model that classifies and detects objects in an image in real-time and in an unprecedented single pass [43]. YOLO is used in this study for the purpose of quickly and accurately locating all forged regions within an image. Unlike region-based algorithms, YOLO is capable of computing an image in its entirety at once, which is why it is a suitable computational and time-efficient model for use in real-world scenarios. The predicted bounding box by YOLO is given by:

$$\hat{B} = (x, y, w, h, C \quad (4)$$

where (x, y) represent the center coordinates, w and h denote the width and height of the bounding box, and C is the confidence score indicating the likelihood of forgery presence. The YOLO system architecture is shown in Figure 5.

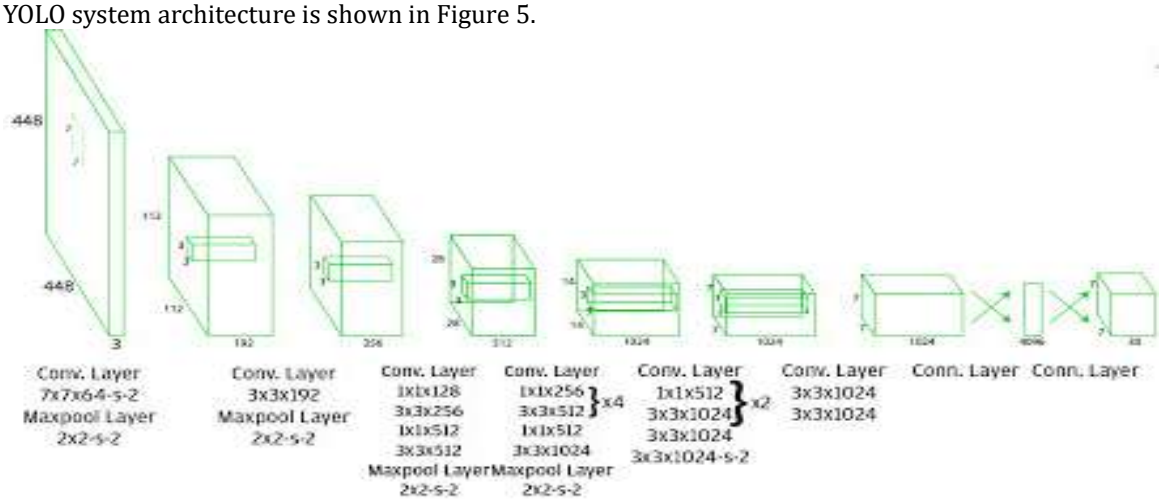


Figure 5. Architecture of YOLO [44]

5.5 Proposed Methodology

The proposed methodology presents a hybrid image forgery detection framework evaluated on three benchmark datasets: CASIA 2.0, MICC-F2000, and CoMoFoD. Initially, images from these datasets are pre-processed through normalization, noise reduction, and contrast enhancement, followed by training-stage data augmentation to improve robustness. DCT draws out the frequency-based features showing forgery clues. The Watershed algorithm segregates these features into useful regions for accurate pinpointing. Then, RCNN does a detailed region-by-region analysis for finding forged areas, and YOLO provides fast and accurate global localization with the help of bounding box predictions. Finally, the combination of the outputs of RCNN and YOLO decides on an image being real or forged and marks the manipulated regions, which ultimately gives a fast and trustworthy image forgery detection framework. The architecture proposed is diagrammatically presented in Figure 6.

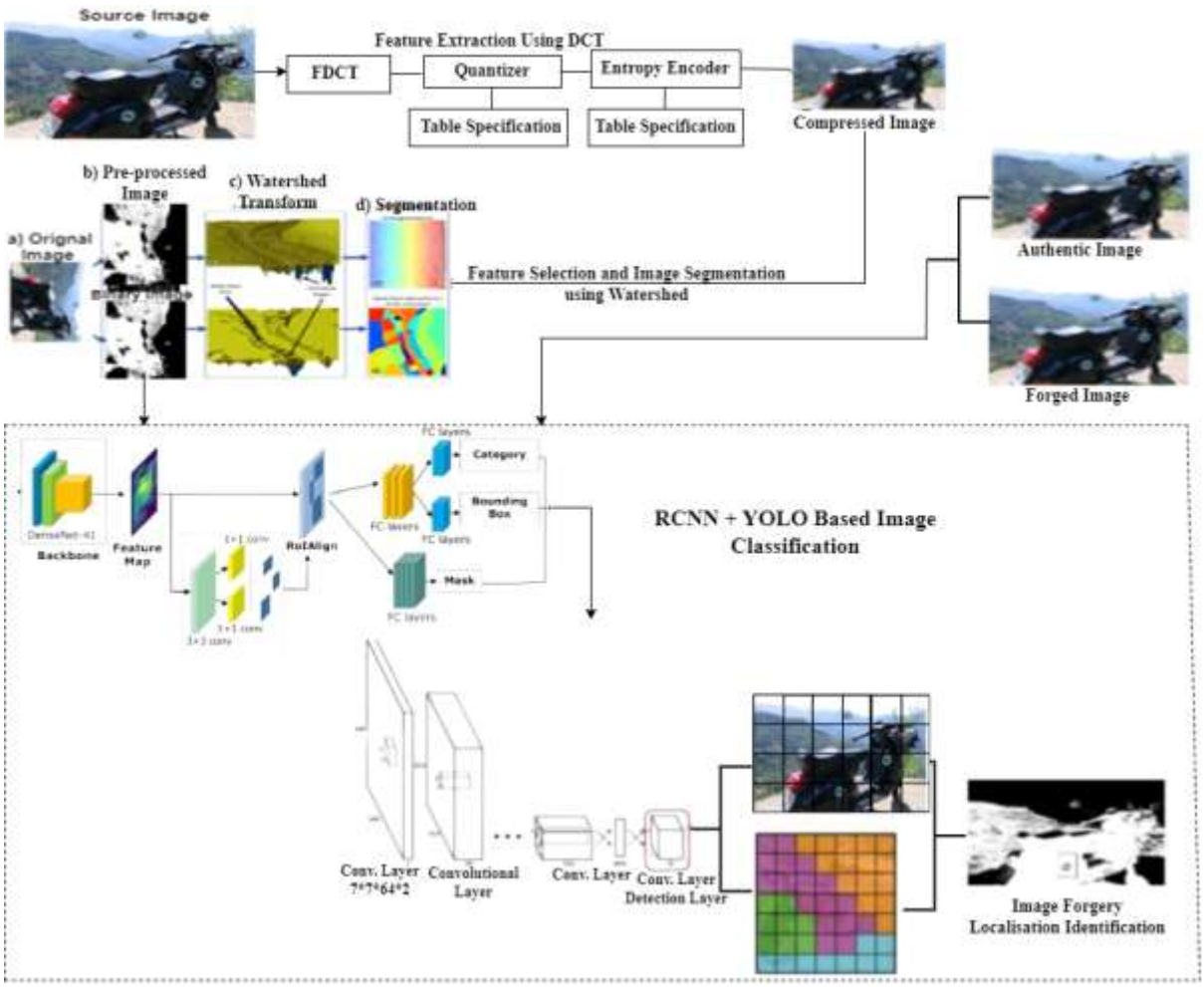


Figure 6. Proposed Methodology

5.6 Proposed Algorithm

Start

Step 1: Capture images from secondary benchmark datasets (CASIA 2.0, MICC-F2000, and CoMoFoD).

dataset_images = load_datasets()

Step 2:Preprocess the images using normalization, noise reduction, and contrast enhancement.

preprocessed_images = preprocess(dataset_images)

Step 3: Apply data augmentation (rotation, brightness–contrast adjustment, mild JPEG compression)

augmented_images = augment(preprocessed_images)

Step 4: Extract frequency-domain features from the images using the DCT.

dct_features = DCT(augmented_images)

Step 5: Perform feature selection and image segmentation using the Watershed transform.

segmented_regions = watershed(dct_features)

selected_features = select_features(segmented_regions)

Step 6: Split the dataset into training and testing sets.

train_set, test_set = split_data(selected_features)

Step 7: Train the hybrid DL model using RCNN for region-level analysis and YOLO for fast localization.

rcnn_model = train_RCNN(train_set)

yolo_model = train_YOLO(train_set)

Step 8: Combine RCNN and YOLO outputs using an ensemble strategy for final classification.

classification = ensemble(rcnn_model, yolo_model, test_set)
Step 9: Identify and localize forged regions using the predicted masks and bounding boxes.
forgery_locations = localize_forgeries(classification)
Step 10: Return the final forgery detection results.
return forgery_locations
End

6. Results and discussion

This section provides the system configuration and evaluation metrics used and presents a comparative analysis that evaluates the performance of the system in relation to existing methods. Here, Table 1 provides the system configuration of the proposed model. Table 2 describes key GAN-based forgery detection parameters for the detection model:

Table 1. System configuration

System	Configuration
Tool	Google collab
computer	Windows 10 pro
Processor	Intel core i5 2.70 GHZ
Ram	8 + 8GB
Type	X64-based processor

Table 2: Hyperparameters of the proposed model.

Hyperparameter	Value
Learning rate	0.001
Batch size	32
Epochs	50
Input size	416*416*3
IoU threshold	0.5
Rol pooling size	7*7
Number of Rols	128
optimizer	Stochastic Gradient Descent

6.1Results analysis

The CASIA 2.0 dataset example shown in Figure 7 begins with the original image, followed by the three image improving techniques (normalization, noise reduction, contrast and detail increase). Figure 8 uses the same image pre-processing techniques with samples taken from the MICC-F2000 dataset, and the same noise enhancement and detail recovery are evident from the illustrations. Also, the original image of the CoMoFoD dataset presented in Figure 9 has also been pre-processed with the same techniques, resulting in holdout values and clearer textures. In all of these illustrations, the same techniques applied consistently in the various datasets were successful in creating a polished image, ready for further processing.





Figure 7: Pre-processing Stages on CASIA Dataset



Figure 8: Visualization of Pre-processing Stages on MICC Dataset

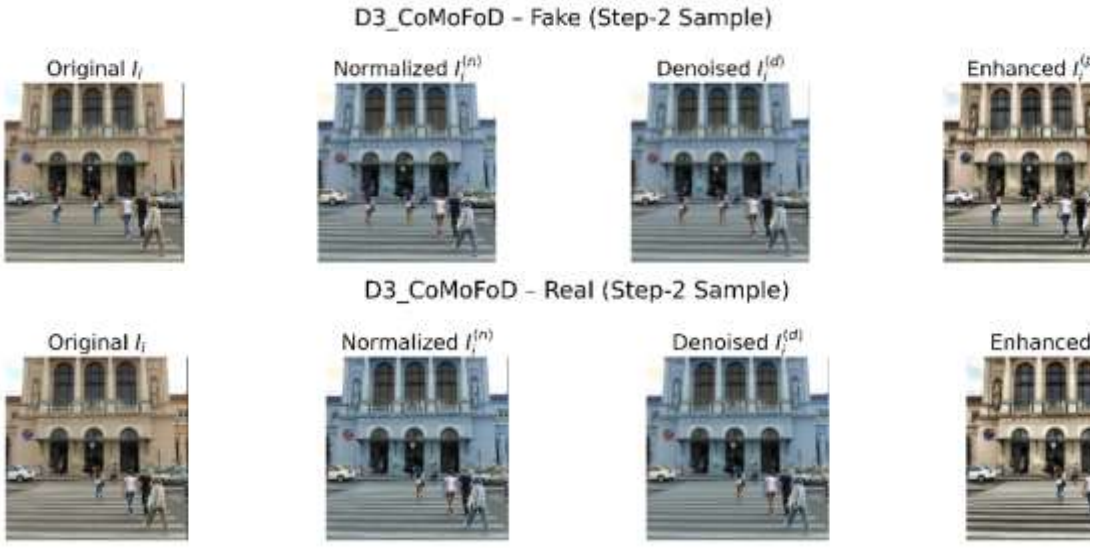


Figure 9: Visualization of Pre-processing Stages on COMOFOD Dataset

The augmentation procedure done on the images in the CASIA 2.0 dataset as demonstrated in Figure 10, starts with the original image, then continues with the augmented image in which apply rotational, brightness and contrast changes, with small JPEG compression and other changes to capture real-life differences. Figure 11 employs the same augmentation pipeline on the images in the MICC-F2000 Dataset. These transformations enhance the diversity of the dataset while preserving the important characteristics related to the forgery. Moreover, Figure 12 contains examples of augmentation on the images from the CoMoFoD Dataset, which indicates the same

systematic augmentation across the datasets. These transformations improve the robustness and generalization of the forgery detection system, as illustrated in these figures.



Figure 10: Augmentation on CASIA Dataset

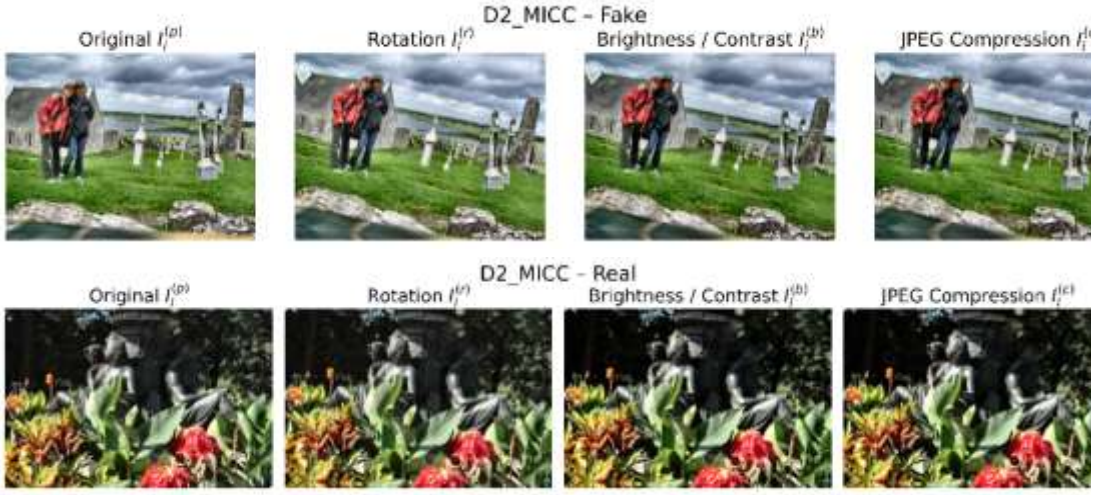


Figure 11: Augmentation on MICC Dataset

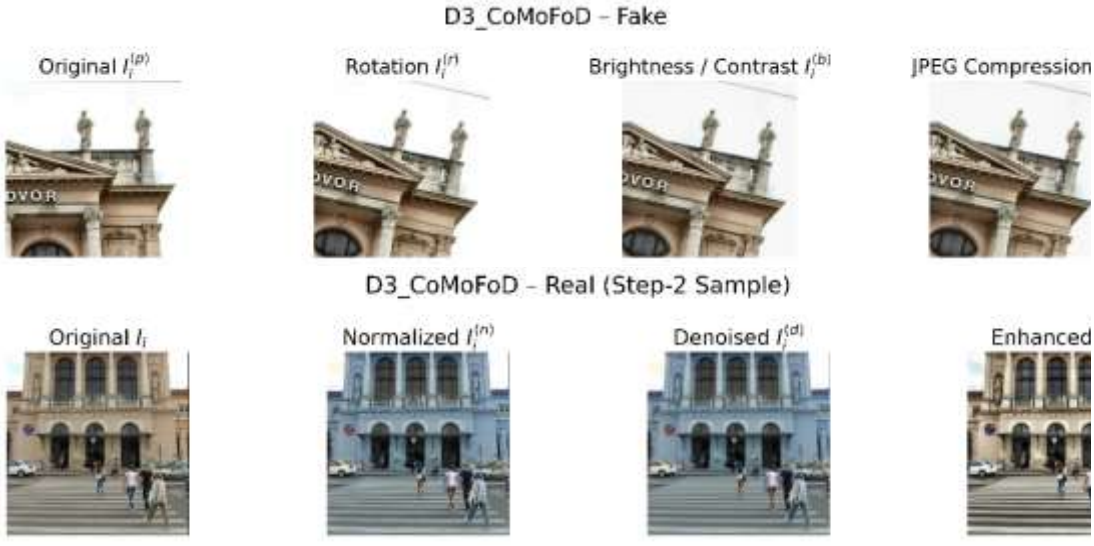
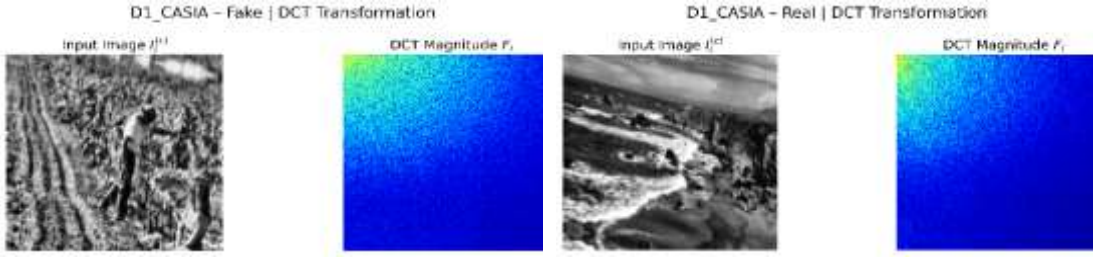


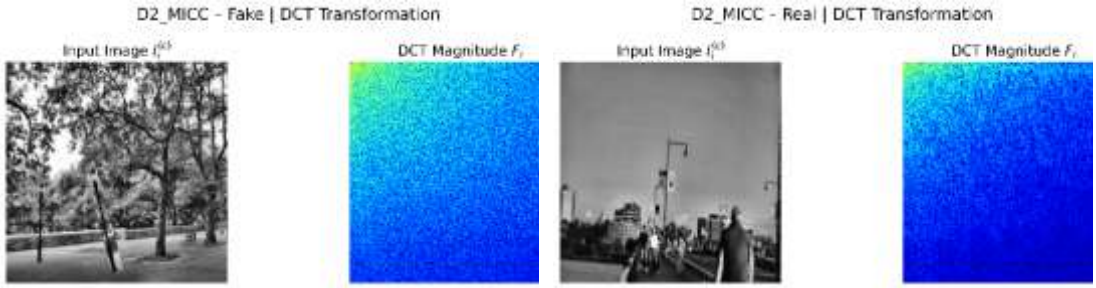
Figure 12: Augmentation on the COMOFOD Dataset

The benefit of DCT for feature extraction for images using the CASIA 2.0, MICC-F2000, and CoMoFoD datasets is outlined in Figures 13, 14, and 15, respectively. For each figure, the first subfigure (a) shows one forged and one genuine image and their corresponding DCT magnitude maps, while the second subfigure (b) shows the other genuine image along with the features that were extracted from it in the frequency domain. The extracted features that are visualized for the DCT magnitude images signify that the frequency components of the images were altered and thus present the DCT features of interest. The features were then used in the following steps of the segmentation and DL based forgery detection tasks.



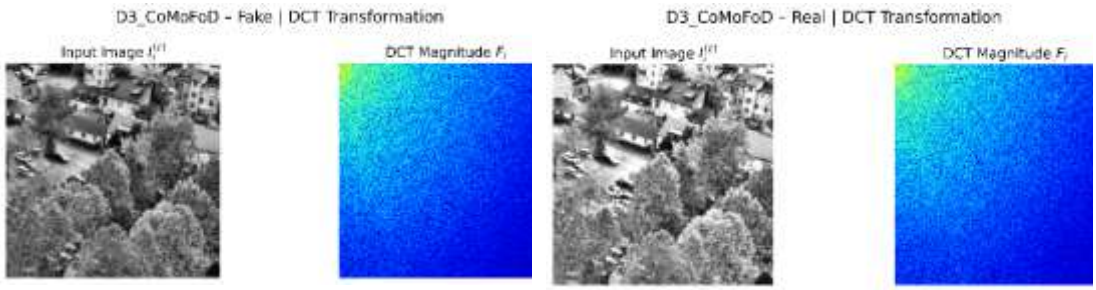
(a) (b)

Figure 13: DCT Feature Extraction on CASIA



(a) **(b)**

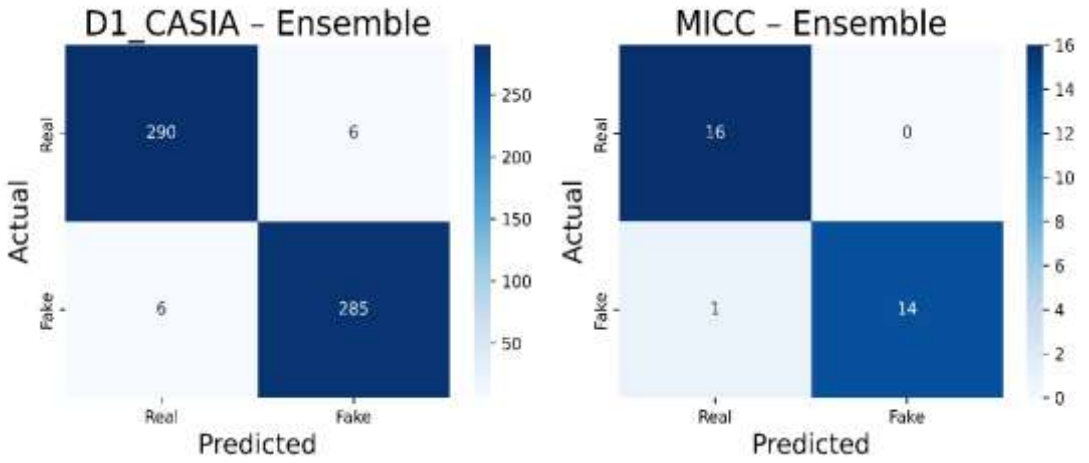
Figure 14: DCT Feature Extraction on MICC



(a) **(b)**

Figure 15: DCT Feature Extraction on COMOFOD

The proposed ensemble-based forgery detection model's ensemble-based confusion matrix is shown in three datasets in Figure 16. In Figure 16(a), it is shown that in the CASIA 2.0 dataset, the model correctly classifies 290285 real and forged images with six false classifications in each direction. So, the model has balanced classification performance. In Figure 16(b), with the model's results in the dataset, MICC-F2000, it correctly identifies 16 real and 14 forged images, having no false positives and 1 false negative, and demonstrates overall accuracy within a high range of accuracy. In the same way, Figure 16 (c) presents the confusion matrix results of the CoMoFoD dataset, describing that the model fully classification of 191 real images against 199 forged images with just 1 false classification in each direction, which demonstrates the considerable robustness of the proposed ensemble approach against the available datasets.



16(a)

16(b)



16(c)

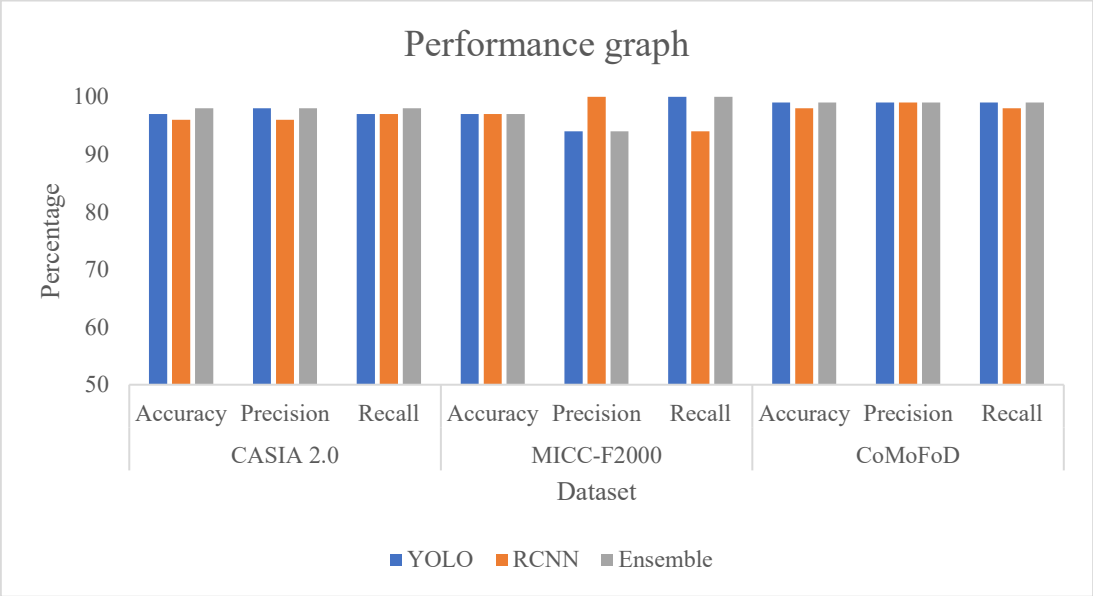
Figure 16 (a) confusion matrix on CASIA dataset, (b) MICC-F2000 dataset, (c) CoMoFoD dataset

Table 3 outlines a comparison of the performance of YOLO, RCNN, and Ensemble approaches on three datasets. As for CASIA 2.0, YOLO records an accuracy rate of 98%, with precision and recall standing at 97% and 98%, respectively. Meanwhile, RCNN surpasses YOLO with an accuracy rate of 96% and a recall rate of 97%, while precision is equal 96%. The Ensemble method, however, surpasses the two with 98% accuracy, 98% precision, and 98% recall. The MICC-F2000 dataset illustrates YOLO with 97% accuracy and 94% precision, but the recall rate stands at 100%. In contrast, RCNN records 97% accuracy, with 100% precision, and 94% recall, marking improved results. Teaming the two surpasses both, leading with 97% accuracy, 94% precision, and 100% recall. CoMoFoD records YOLO with 99% accuracy, 99% precision, and 99% recall. The accuracy goes up to 98% with RCNN, while precision increases to 99% and recall to 98%. Other than Ensemble, the rest are surpassed with 99% accuracy, 99% precision and 99% recall. The Ensemble method has proven to stand out among other methods for all datasets and parameters, proving to be reliable and efficient.

Table 3: Comparison of YOLO, RCNN, and Ensemble Techniques on Multiple Datasets

Techniques	**CASIA 2.0**			**MICC-F2000**			**CoMoFoD**		
	Accuracy	Precision	Recall	Accuracy	Precision	Recall	Accuracy	Precision	Recall

YOLO	97	98	97	97	94	100	99	99	99
RCNN	96	96	97	97	100	94	98	99	98
Ensemble	**98**	**98**	**98**	**97**	**94**	**100**	**99**	**99**	**99**



6.2 Ablation Study

This section shows an ablation study on all three benchmark datasets CASIA 2.0, MICC-F2000, and CoMoFoD, by systematically removing or isolating individual components. The study evaluates the contribution of DCT-based feature extraction, Watershed segmentation, and the YOLO–RCNN ensemble toward overall performance.

- ***Ablation Study on CASIA 2.0 Dataset***

Table 4 depicts the accuracy of different image forgery detection techniques on the CASIA 2.0 dataset with respect to variation in the train-test split (90:10, 80:20, and 70:30). Again, it is evident that the reduction in the size of the training data reduces the model performance; the larger size of the training set enables better learning of the models. Overall, feature transformation combined methods with DL outperform their single-model counterparts.

Table 4: Ablation study on CASIA 2.0 Dataset

Techniques	Accuracy (90:10)	Accuracy (80:20)	Accuracy (70:30)
YOLO only	92.4	90.2	88.1
RCNN only	95.3	94.1	92.0
YOLO + RCNN (No DCT)	94.0	92.6	90.4
DCT + RCNN (No YOLO)	96.2	95.3	93.6
Proposed ensemble model	98.0	96.8	95.1

- ***Ablation Study on MICC-F2000 Dataset***

Table 5 below depicts the accuracy achieved for different image forgery detection methods for the MICC-F2000 dataset at varying splits, such as 90:10, 80:20, and 70:30 between training and testing data. It can be observed that model performance goes down with less training data; a larger training set enables better learning. Overall, combining feature transformation with DL tends to give better results than a single-model approach.

Table 5: Ablation study on MICC-F2000 Dataset

Techniques	Accuracy (90:10)	Accuracy (80:20)	Accuracy (70:30)
YOLO only	94.5	92.8	91.0
RCNN only	96.2	94.9	93.3

YOLO + RCNN (No DCT)	95.4	94.1	92.2
DCT + RCNN (No YOLO)	96.8	95.6	94.1
Proposed ensemble model	97.0	96.1	94.8

- ***Ablation Study on CoMoFoD Dataset***

This ablation study investigates the various parts of the model with respect to forgery detection performance on the CoMoFoD dataset using 90:10, 80:20, and 70:30 train-test splits. It is observed that the accuracy drops when the number of training data is less, which signifies that sufficient learning samples are required by the model. The two comparisons reveal a significant improvement in the detection using frequency-domain features and region-based learning compared to using only one or partial models. The proposed ensemble model outperforms all splits in CoMoFoD, confirming that combining YOLO, DCT, and RCNN helps in more reliable detection of image forgeries.

Table 6: Ablation study on CoMoFoD Dataset

Techniques	Accuracy (90:10)	Accuracy (80:20)	Accuracy (70:30)
YOLO only	95.8	93.4	91.2
RCNN only	97.2	96.1	94.0
YOLO + RCNN (No DCT)	96.3	95.2	93.1
DCT + RCNN (No YOLO)	98.0	97.4	95.6
Proposed ensemble model	99.0	98.2	96.9

6.3 Comparative analysis

Table 7 shows Performance comparison of the proposed method with each dataset separately. On CASIA 2.0 dataset, the accuracy of ResTran is 84.7%, Multi-Modal is 87.13%, and for the ViT-based approach, it is 98%. The proposed approach has an accuracy of 98% as well, outperforming traditional and multi-modal approaches while still remaining competitive. MICC-F2000: The ConvLSTM model achieves 93.9% in accuracy, while the standard CNN is 76% in this regard. For the proposed model, it is 98% accurate, thus showing a great detection capability. CoMoFoD: BWT-SVD achieves 92.2%, ResNet50+ELA reaches 98%, while the affine transformation-based approach is 77.9%. However, the proposed approach surpasses all these with an accuracy of 99%, establishing its robustness on such a large, challenging forgery dataset.

Table 7: Comparison of accuracy across CASIA 2.0, MICC-F2000, and CoMoFoD Datasets

Dataset	Technique	Accuracy
CASIA 2.0	Restran [25]	84.7
	Multi-Modal [26]	87.13
	VIT [45]	97
	Proposed	**98**
MICC-F2000	ConvLSTM [33]	93.9
	CNN [46]	76
	Proposed	**97**
CoMoFoD	BWT-SVD [28]	92.2
	Affine Transformation [47]	77.9
	ResNet50+ELA [48]	98
	Proposed	**99**

7. Conclusion and future scope

This study demonstrates that advanced variants of YOLO and RCNN-based DL methods enhance image forgery detection in digital images. When applied to different datasets, these methods raise the accuracy to a great extent. They consistently outperform existing methods, such as 98% on the CASIA 2.0 dataset, 99% on the CoMoFoD dataset, and 97% on the MICC-F2000 dataset, proving that the proposed DL models are strong and efficient at finding forgeries in many kinds of image datasets and thus have the potential to be used in real-world applications of image authentication and security. The framework has shown quite promising performance on several forgery types due

to the usage of both frequency and region information. However, extending it to cope with more advanced attacks, like deepfakes, is an interesting avenue of future research. This model also suffers from some limitations. First, its performance may degrade under challenging conditions such as very low light, heavy compression of images, or highly sophisticated manipulations including deepfakes.

In future work, limitations identified would be addressed to increase robustness for challenging conditions: low-light imagery, heavy compression, and sophisticated forgery techniques such as deepfakes. Generalization and overall detection performance would be enhanced by expanding the training datasets and incorporating adaptive architectures, including transformer-based models with enriched frequency-domain features.

References

1. Xiao, Bin, Yang Wei, Xiuli Bi, Weisheng Li, and Jianfeng Ma. "Image splicing forgery detection combining coarse to refined convolutional neural network and adaptive clustering." Information Sciences 511 (2020): 172-191.
2. Farid, Hany. "Image forgery detection." IEEE Signal Processing Magazine 26, no. 2 (2009): 16-25.
3. Walia, Savita, and Krishan Kumar. "Digital image forgery detection: a systematic scrutiny." Australian Journal of Forensic Sciences 51, no. 5 (2019): 488-526.
4. Bansal, Ankita, and Aditya Atri. "Study of copy-move forgery detection techniques in images." In 2020 8th International Conference on Reliability, Infocom Technologies and Optimization (Trends and Future Directions) (ICRITO), pp. 630-635. IEEE, 2020.
5. Soni, Badal, Pradip K. Das, and Dalton Meitei Thounaojam. "CMFD: a detailed review of block-based and key feature-based techniques in image copy-move forgery detection." IET Image Processing 12, no. 2 (2018): 167-178.
6. Thakur, Tulsi, Kavita Singh, and Arun Yadav. "Blind approach for digital image forgery detection." International Journal of Computer Applications 975 (2018): 8887.
7. Raja, A. "Active and Passive Detection of Image Forgery: A Review Analysis." IJERT-Proc 9, no. 5 (2021): 418-424.
8. Al-Qershi, Osamah M., and Bee Ee Khoo. "Passive detection of copy-move forgery in digital images: State-of-the-art." Forensic Science International 231, no. 1-3 (2013): 284-295.
9. Abd Warif, Nor Bakiah, Ainuddin Wahid Abdul Wahab, Mohd Yamani Idna Idris, Roziana Ramli, Rosli Salleh, Shahaboddin Shamshirband, and Kim-Kwang Raymond Choo. "Copy-move forgery detection: survey, challenges and future directions." Journal of Network and Computer Applications 75 (2016): 259-278.
10. Kotsiantis, Sotiris B., Ioannis Zaharakis, and P. Pintelas. "Supervised machine learning: A review of classification techniques." Emerging artificial intelligence applications in computer engineering 160, no. 1 (2007): 3-24.
11. Li, Weihai, and Nenghai Yu. "Rotation robust detection of copy-move forgery." In 2010 IEEE International Conference on Image Processing, pp. 2113-2116. IEEE, 2010.
12. Liu, Ke, Wei Lu, Cong Lin, Xinchao Huang, Xianjin Liu, Yuileong Yeung, and Yingjie Xue. "Copy move forgery detection based on keypoint and patch match." Multimedia tools and applications 78 (2019): 31387-31413.
13. Ardizzone, Edoardo, Alessandro Bruno, and Giuseppe Mazzola. "Copy–move forgery detection by matching triangles of key points." IEEE Transactions on Information Forensics and Security 10, no. 10 (2015): 2084-2094.
14. Yang, Fan, Jingwei Li, Wei Lu, and Jian Weng. "Copy-move forgery detection based on hybrid features." Engineering Applications of Artificial Intelligence 59 (2017): 73-83.
15. Lin, Cong, Wei Lu, Wei Sun, Jinhua Zeng, Tianhua Xu, and Jian-Huang Lai. "Region duplication detection based on image segmentation and keypoint contexts." Multimedia Tools and Applications 77 (2018): 14241-14258.
16. Krizhevsky, Alex, Ilya Sutskever, and Geoffrey E. Hinton. "Imagenet classification with deep convolutional neural networks." Advances in neural information processing systems 25 (2012).

17. Bayar, Belhassen, and Matthew C. Stamm. "A DL approach to universal image manipulation detection using a new convolutional layer." In Proceedings of the 4th ACM workshop on information hiding and multimedia security, pp. 5-10. 2016.
18. Tahaoglu, Gul, Guzin Ulutas, Beste Ustubioglu, Mustafa Ulutas, and Vasif V. Nabiyev. "Ciratefi based copy move forgery detection on digital images." Multimedia Tools and Applications 81, no. 16 (2022): 22867-22902.
19. Dhanaraj, Rachel Selva, and M. Sridevi. "A study on detection of copy-move forgery in digital images." In 2021 Third International Conference on Intelligent Communication Technologies and Virtual Mobile Networks (ICICV), pp. 900-905. IEEE, 2021.
20. Muzaffer, Gul, and Guzin Ulutas. "A new DL -based method to detection of copy-move forgery in digital images." In 2019 scientific meeting on Electrical-Electronics & Biomedical Engineering and computer science (EBBT), pp. 1-4. IEEE, 2019.
21. Gürbüz, Emre, Guzin Ulutas, and Mustafa Ulutas. "Detection of free-form copy-move forgery on digital images." Security and Communication Networks 2019 (2019): 1-14.
22. Meena, Kunj Bihari, and Vipin Tyagi. "A hybrid copy-move image forgery detection technique based on Fourier-Mellin and scale invariant feature transforms." Multimedia Tools and Applications 79, no. 11-12 (2020): 8197-8212.
23. Yang, Bin, Xingming Sun, Honglei Guo, Zhihua Xia, and Xianyi Chen. "A copy-move forgery detection method based on CMFD-SIFT." Multimedia Tools and Applications 77 (2018): 837-855.
24. Zhang, Jianyi, Xuanxi Huang, Yaqi Liu, Yuyang Han, and Zixiao Xiang. "GAN-based medical image small region forgery detection via a two-stage cascade framework." Plos one 19, no. 1 (2024): e0290303.
25. Rao, Jingjing, Songpon Teerakanok, and Tetsutaro Uehara. "ResTran: Long distance relationship on image forgery detection." IEEE Access (2023).
26. Gupta, R., Singh, P., Alam, T., & Agarwal, S. (2023). A deep neural network with hybrid spotted hyena optimizer and grasshopper optimization algorithm for copy move forgery detection. Multimedia Tools and Applications, 82(16), 24547–24572.
27. Krishnaraj, N., B. Sivakumar, Ramya Kuppusamy, Yuvaraja Teekaraman, and Amruth Ramesh Thelkar. "Design of automated DL -based fusion model for copy-move image forgery detection." Computational Intelligence and Neuroscience 2022 (2022).
28. Rathore, Neeraj Kumar, Neelesh Kumar Jain, Prashant Kumar Shukla, Uma Shankar Rawat, and Rachana Dubey. "Image forgery detection using singular value decomposition with some attacks." National Academy Science Letters 44 (2021): 331-338.
29. Kasban, Hany, and Sabry Nassar. "An efficient approach for forgery detection in digital images using Hilbert–Huang transform." Applied Soft Computing 97 (2020): 106728.
30. Wang, Jinwei, Renfeng Liu, Hao Wang, Bin Wu, and Yun-Qing Shi. "Quaternion Markov Splicing Detection for Color Images Based on Quaternion Discrete Cosine Transform." KSII Transactions on Internet & Information Systems 14, no. 7 (2020).
31. Lai, Yuecong, Tianqiang Huang, Jing Lin, and Henan Lu. "An improved block-based matching algorithm of copy-move forgery detection." Multimedia Tools and Applications 77 (2018): 15093-15110.
32. Aryan, Police, Rama Muni Reddy Yanamala, Archana Pallakonda, Rayappa David Amar Raj, and K. Krishna Prakasha. "Lightweight End-to-End Patch-Based Self-Attention Network for Robust Image Forgery Detection." IEEE Access (2025).
33. Elaskily, Mohamed A., Monagi H. Alkinani, Ahmed Sedik, and Mohamed M. Dessouky. " DL based algorithm (ConvLSTM) for copy move forgery detection." Journal of Intelligent & Fuzzy Systems 40, no. 3 (2021): 4385-4405.
34. Hosny, Khalid M., Akram M. Mortda, Mostafa M. Fouda, and Nabil A. Lashin. "An efficient cnn model to detect copy-move image forgery." IEEE Access 10 (2022): 48622-48632.
35. Castillo Camacho, Ivan, and Kai Wang. "A comprehensive review of deep-learning-based methods for image forensics." Journal of Imaging 7, no. 4 (2021): 69.
36. Ahmed, Farah, Fouad Khelifi, Ashraf Lawgaly, and Ahmed Bouridane. "Temporal image forensic analysis for image dating with DL." In 2020 International Conference on Computing, Electronics & Communications Engineering (iCCECE), pp. 109-114. IEEE, 2020.
37. Rathi, Kavita, and Monika Urvashi. "DATASET for Image Forgery Detection."

38. Gupta, Ruchi, Pushpa Singh, Tanweer Alam, and Shivani Agarwal. "A deep neural network with hybrid spotted hyena optimizer and grasshopper optimization algorithm for copy move forgery detection." Multimedia Tools and Applications 82, no. 16 (2023): 24547-24572.
39. Dong, Jing, Wei Wang, and Tieniu Tan. "Casia image tampering detection evaluation database." In 2013 IEEE China Summit and International Conference on Signal and Information Processing, pp. 422-426. IEEE, 2013.
40. Kaur, Amandeep. "Aayushi. Image segmentation using watershed transform." International Journal of Soft Computing and Engineering (IJDCE) 4 (2014): 5-8.
41. Girshick, Ross, Jeff Donahue, Trevor Darrell, and Jitendra Malik. "Rich feature hierarchies for accurate object detection and semantic segmentation." In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 580-587. 2014.
42. https://medium.com/@selfouly/r-cnn-3a9beddfd55a
43. Redmon, Joseph, Santosh Divvala, Ross Girshick, and Ali Farhadi. "You only look once: Unified, real-time object detection." In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 779-788. 2016.
44. https://www.geeksforgeeks.org/yolo-you-only-look-once-real-time-object-detection/
45. Manu, V. T., and Babu M. Mehtre. "Copy-move tampering detection using affine transformation property preservation on clustered keypoints." Signal, Image and Video Processing 12 (2018): 549-556.
46. Pawar, Digambar, Raghavendra Gowda, and Krishna Chandra. "Image forgery classification and localization through vision transformers." International Journal of Multimedia Information Retrieval 14, no. 1 (2025): 1-11.
47. Muniappan, Thiiban, Nor Bakiah Abd Warif, Ahsiah Ismail, and Noor Atikah Mat Abir. "An Evaluation of Convolutional Neural Network (CNN) Model for Copy-Move and Splicing Forgery Detection." International Journal of Intelligent Systems and Applications in Engineering 11, no. 2 (2023): 730-740.
48. Dabhole, Smruti Dilip, and G. G. Rajput. "Copy Move Image Tampering: Genuine and Tampered Region Identification and Classification using Various Residual Network Approaches."