International Journal of
INTELLIGENT SYSTEMS AND APPLICATIONS IN ENGINEERING

ISSN:2147-6799 www.ijisae.org Original Research Paper

Hybrid Dense Net Based Segmentation Framework for Automated Forgery Detection: Analyzing Copy-Move and Image Splicing Techniques

Anshul Kumar Singh[1], Vivek Kumar[2], Brajesh Kumar Singh[3]

Submitted:10/03/2024 **Revised**: 25/04/2024 **Accepted**: 02/05/2024

Abstract: Protecting information from modification is the biggest challenge in current digital world. To cope up with this a unique hybrid dense net-based segmentation framework for automated forgery recognition is designed. Proposed methodology place a particular emphasis on the investigation of copy-move and image-splicing techniques. When analyzing digital images, system looks for signs of tampering by employing a hybrid approach that combines deep learning with a dense network (DenseNet121) structure. In this paper, authors used a two-stage approach, beginning with coarse-grained segmentation using a modified version of the U-Net structure, and then moving on to fine-grained segmentation using a hybrid dense net. Suggested system is successful for detecting copy-move and image splicing forgery, exceeding DCT & DWT Based forgery detection and CNN-CovLSTM approaches achieves less accuracy then our proposed model, according to extensive testing carried out on a diverse dataset. Notably, when it comes to recognizing challenging copy-move along with frauds, proposed method can obtain far greater rates of accuracy, precision, and recall than prior methods have been capable to achieve. The proposed model achieves an accuracy of 98% for the CASIA-2 dataset a precision of 92% and an F1 score of 90%. For the MICC-F2000 dataset, the proposed model achieves an accuracy of 99%, a precision of 98%, and an impressive F1 score of 99%.

Keywords: *Hybrid Dense Net Based Segmentation, Automated Forgery Recognition, Copy-Move, Image Splicing, Deep Learning, Digital Image Forensics.*

1. Introduction

Electronic devices are commonplace and affordable because of technical progress and international trade. Therefore, digital cameras have become more common. All around there are camera sensors, and the utilization of them to amass a vast visual database. Many forms of compulsory online filing demand soft copies of photographs, and many people post images from their daily lives on social media.

Those who have trouble reading may still be able to understand what is being said if it is accompanied by an image.

Therefore, images play a key role in the digital world since they are used for storing and sharing information. Numerous quick-editing image-processing tools are ready [1]. These resources were developed to help users to get better results while working with photos. However, some individuals use their ability to spread lies and misinformation rather than improve their image. The harm done by these phony pictures is often irreparable and poses a serious hazard [2]. Image splicing & copy-move, two common kinds of photo manipulation, are explained:

[1]*Computer Science & Engineering, Quantum University, Roorkee, India, akrajawat20@gmail.com* (Corresponding Author).*
[2]*Vice Chancellor, Quantum University, Roorkee, India, vicechancellor@quantumuniversity.edu.in*
[3]*Computer Science & Engineering, R. B. S. Engineering Technical Campus, Bichpuri, Agra, India, brajesh1678@gmail.com.*

Image Splicing: A piece of an image from a donor is grafted onto an existing image. The final forged picture may also be constructed from a series of donor photographs. Figure 1 displays the visual depiction of Image Splicing, with the original image shown on the left and a tampered version created using copy-move forgery shown on the right.

Fig. 1. Original (left), tampered with using image splicing forgery (right).

Copy-Move: There is just one picture in this situation. A cut-and-paste operation is performed on the picture itself. This is often used as a cover for something else. No parts of any other photographs were used in the creation of the final forgery. In Fig. 2, there is an instance of picture copy-move in which missile and smoke of the same image is copied to other location in that image it.

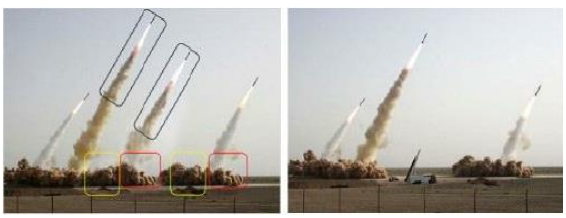


Fig. 2. Copy-Move forgery in which same image is used.

In both situations, the original content of a picture is replaced with something else for the express aim of spreading false information. Image fraud has transformed what was once a very reliable medium for sharing information into a tool for spreading falsehoods. The public's faith in pictures is being shaken since it's not always easy to tell whether an image has been tampered with. Therefore, it is crucial to identify picture forgeries to curb the propagation of false information and restore faith in visual media. This may be accomplished by employing different image processing methods to detect and analyze artifacts left behind during the counterfeit process [4].

The purpose of picture forgery detection is to establish whether a given digital image is fake. Active and passive (or blind) methods may be broadly classified as approaches to this issue. In active method, a watermark is embedded into an image using active techniques, and the genuineness of the image is determined by comparing the extracted watermark to the original watermark to determine whether the image is genuine. However, this method is restricted due to the lack of watermark information in most situations [5]. Whereas passive method works rather than relying on the existence of a signature or watermark in the source picture, a forgery These are the primary components of the overarching framework for the detection of copy-move and splicing forgeries in photos using passive methods.

Image Pre-processing: Before feature extraction, an image is divided into frequency-scale components to better represent the geometric information of a texture. This is achieved by converting a red, green, and blue (RGB) picture to the desired color space.

Feature Extraction: The purpose of the extraction of features is to find the most informative representation of the data. A classifier's performance may suffer because the feature space built using feature extraction methods has an unmanageably high dimension, contains redundant information, or both. After that, it employs the feature selection procedure to simplify the system.

Classifier Selection and Modelling: The photos in the training set are used to train a classifier that has been carefully chosen or created. A classifier's settings are fine-tuned throughout the training process.

Classification: The goal of the classifier is to sort the input photos into two groups: authentic and fake. Forgery detection has been approached with several different

2. Literature Review

Numerous writers have tried this approach and reported their findings following a survey of the relevant literature.

Krishnaraj et al., (2022) [8] stated that the generative adversarial network and dense network models are included in the detecting and localizing copy-move forgeries approach suggested. Detecting and localizing copy-move detection approach examines the image's contents and structure to determine its validity or integrity. Virtually no telltale signs of manipulation remain in digital forgeries, but the introduction of artifacts and other types of inconsistency may be traced back to a disruption in the image's structure. These discrepancies may be utilized to identify a fake image. This method of detection is widely used since it doesn't require any background knowledge of the picture to function [6].

There are methods now in use that can single out certain signs of tampering and locate the area where they occurred independently. Figure 3 is a taxonomy of methods for detecting picture fraud.

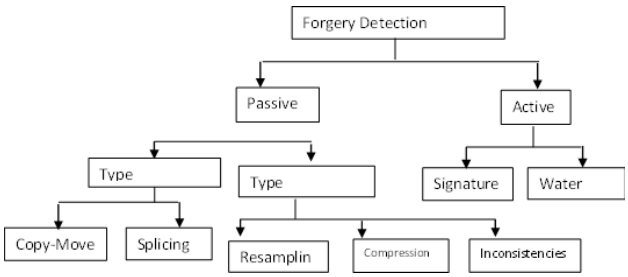


Fig. 3. Methods for detecting fake photos: a taxonomy. [7].

classifiers, including support vector machine (SVM), Naive Bayes, and Artificial Neural Networks (ANN). Figure 4 illustrates the overall structure of a system specifically developed to detect and localize instances of digital image manipulation, such as copy-move and splicing, as described earlier.

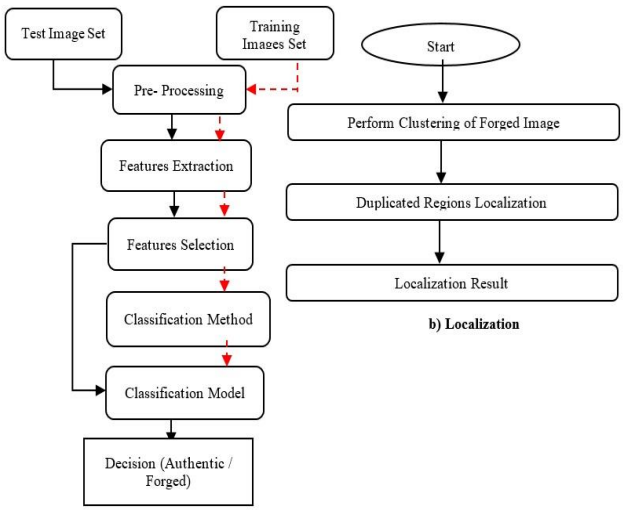


Fig. 4. (a) Forgery detection and (b) localization general framework.

forgeries (DLFM-CMDFC) combines the two outputs to produce a new layer, which is then used to integrate the input vectors using the first layer which is an extreme learning machine classifier. Both the weight and bias parameters of the ELM model used in extreme learning are improved with the aid of the synthetic fish swarm method. Inputs from networks are transmitted into the unit doing the merging.

Koul et al., (2022) [9] proposed a unique approach to automated copy-move forgery detection proposes utilizing a convolutional neural network (CNN). Specifically, MICC-F2000, a benchmark dataset consisting of 2000 pictures, is considered for the experiments. Compared to state-of-the-art methods for identifying copy-move forgeries, the experimental results show that the proposed model achieves better results. Copy-move forgeries performed very well, with an accuracy of 97.52%, 2.52% better than the state-of-the-art approaches.

Pawar et al., (2022) [10] offer our approach, which feeds ELA-pre-processed photos into a CNN for classification to identify image forging, and introduces our custom dataset, FIDAC (Forged photos Detection and Classification), that includes both the genuine camera-clicked images and their tampered counterparts. Sensitivity, specificity, precision, accuracy, and F1 Score for the suggested design all come in at 0.8269, 0.8723, 0.8485, and 0.8515, respectively.

Elaskily et al., (2021) [11] shows how deep learning may be used in a novel method for detecting Copy Move Forgery. To implement the suggested model, both CNN and Convolutional Long Short-Term Memory networks are used. To detect copy-move frauds, this method first extracts image features using a sequence of Convolutions layers, Convolutional Long short-term memory (ConvLSTM) layers, and pooling layers, and then compares the features to find differences. Accessible public datasets were used to evaluate this model, including the MICC-F220, MICC-F2000, MICC-F600, and SATs-130.

Agrawal et al., (2021) [12] designed and implemented a method to translate numerical values into English to verify the check, a common cause of cheque bounce-off and the suspension of monetary transactions. The use of optical character recognition (OCR) to recognize the machine-printed numbers, and the resulting matching accuracy was 99.7%. First, the utilization of a particular database to teach the network, and after the target had reached accuracy, the use of additional data sets to put it to the test by trying to match numbers to words.

Hebbar, Nagaveni K., and Ashwini S. Kunte (2021) [13] stated a method via U-Net is proposed for identifying various forms of forgery and pinpointing the manipulated area down to the pixel. The forged images are subjected to Error Level Analysis (ELA), a technique that enhances the model's effectiveness by revealing the forged area with different levels of compression. When compared to CFA1, MFCN, RGB-N, Faster-RCNN_Edge Det, and DU-DC-EC Net, the proposed technique achieves 33% more precision and 27.8% greater F1_score.

Abhishek, and Neeru Jindal (2021) [14] suggested method uses semantic segmentation, colour lighting, and deep convolution neural networks to identify and locate image forgeries. After the pre-processing stage, the colour map is applied using colour illumination. The researchers use a deep convolutional neural network and the transfer learning technique to train the two classes of VGG-16. The experimental findings provide a detection accuracy of more than 98% for both forged and non-forged pixels.

Saber et al., (2020) [15] provide a detailed analysis of the research into numerous picture forgeries and forensic tools. Digital watermarking, digital signature, copy-move, picture retouching, and splicing are only a few of the methods presented in the literature to identify image counterfeiting. The research examination may provide light on the benefits and pitfalls of existing picture forensics technology, paving the way for the creation of more accurate algorithms for detecting forgeries. In addition, the research analyzes the advantages and disadvantages of various forgery detection systems, such as deep learning and convolution neural networks.

Gani, Gulnawaz, and Fasel Qadir (2020) [16] suggest a reliable strategy for identifying copy-move forgeries in the face of varying post-processing attacks. To get useful information out of each piece, we apply the DCT (Discrete Cosine Transform). The experimental results demonstrate that the suggested method outperforms the other state-of-the-art methods in the literature, especially when an image is severely degraded by post-processing attacks such as JPEG compression and additive white Gaussian noise.

Jaiprakash et al., (2020) [17] suggested a model based on low-dimensional features that can spot image counterfeiting caused by copying and pasting. We have used cross-domain feature extraction to detect signs of image modification of any kind. The detection accuracy for copy-move, as well as spliced images, has been demonstrated experimentally to be high. The proposed model achieves great accuracy across a wide variety of image formats.

Bappy et al., (2019) [18] discovered modified areas and labels them as such using resampling characteristics, Long-Short Term Memory (LSTM) cells, and an encoder-decoder network to achieve high-confidence manipulation localization. For example, Joint Photographic Experts Group (JPEG) quality loss, up sampling, down sampling, rotation, and shearing are all artifacts that may be captured using resampling features. The proposed network incorporates an encoder and LSTM network to examine the discriminative properties of manipulated and non-manipulated areas, capitalizing on bigger receptive fields (spatial maps) and frequency domain correlation. For accurate tamper detection in images, a decoder network must be trained to convert feature maps at lower resolutions into pixel-by-pixel predictions.

Wu et al., (2018) [19] A novel end-to-end deep neural network for predicting forgery masks was presented to address the issue of detecting picture copy-move fraud. To be more specific, that is employed a CNN to extract block-

like features from an image, calculate self-correlations between various blocks, employ a pointwise feature extractor to detect matching points, and rebuild a forgery mask using a deconvolution network. As opposed to conventional approaches, which need extensive training and parameter modification across the board, spanning feature extraction to post processing, the proposed strategy may be concurrently optimized to minimize forgery mask reconstruction loss.

Yao et al., (2017) [20] provide a method based on deep learning for identifying object-based video fraud. The proposed deep learning method employs a CNN to automatically extract high-dimensional characteristics from the input picture patches. The CNN model is distinct from others in the computer vision field because a video frame was fed to three levels of pre-processing. There is a high-pass filter layer to boost the remaining signal after video counterfeiting has been performed, a max pooling layer to lower the computational cost of image convolution, which is an absolute distinction layer to cut up temporal duplication between video frames.

Cozzolino et al., (2014) [21] introduced a novel approach to detecting picture forgeries by combining the results of two different tools. The steganalysis discipline has recently presented certain local descriptors that the machine learning tool uses to pick and integrate features. The patch match technique is used by the block-matching tool to facilitate the rapid search for potential matches. The two tools are fine-tuned to work together optimally, capitalizing on their combined benefits.

3. Proposed Methodology

In this paper, authors proposed a novel approach for detecting Copy-Move & Image splicing forgery. And for the implementation part we have used python programming language along with Google Colab. As a first step, a training set and a testing set are created from a dataset of images that have been altered. After the dataset was first tested, the hyper parameters were chosen. There are 32 filters chosen, the kernel size is (5, 5), and the dropout percentage is 0.25. The initial learning rate was set to 0.001 and an Adam optimizer was employed, with the loss function being binary cross entropy. The ReduceLRonPlateau callback was used to reduce the learning rate after the loss stopped getting better. 40% of the data is used for testing, while 60% is used for training. We obtained the data for the model's training via Tensorflow and Kaggle. Training and testing are conducted using the MICC-F2000 dataset and CASIA. Pre-processing such as normalization, de-noising, color enhancement, elimination of anomalies and the outlier detection are applied to the data, a threshold value is set to 3.5 after rigorous training done on the dataset to form unusual data points and one class SVM provides better result. After that, images are resized to the same proportions to ensure consistency. Using the Discrete Cosine Transform (DCT) on picture blocks helps with feature extraction and image is converted from spatial domain to frequency domain. The output of the DCT is a set of coefficients that represents the frequency content of the image. The equation used to calculate the 2-D coefficient matrix of DCT is given below:

$$B(p,q) = \propto (p) \propto (q) \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} A(m,n) \times \cos\frac{\pi(2m+1)p}{2M} \times \cos\frac{\pi(2n+1)q}{2N}, \quad (1)$$

$$where, \propto_p = \begin{cases} \frac{1}{\sqrt{M}}, & p = 0 \\ \sqrt{\frac{2}{M}}, 1 \le p \le M-1 \end{cases} \quad and \quad , \propto_p$$

$$= \begin{cases} \frac{1}{\sqrt{N}}, & q = 0 \\ \sqrt{\frac{2}{N}}, 1 \le q \le N-1 \end{cases}$$

The input image of size (M*N) is transformed by the 2D-DCT such that majority of the features is stored in the first few low frequency DCT coefficients. This allowed for the elimination of the image's high-frequency components with minimal loss of information. In feature selection, the Wrapping Subset Selection Method (WSSM) shown in Fig. 5 improves the classifier's performance while simultaneously decreasing the feature space's dimensionality. Here we took forward feature selection with 50 iterations. The Watershed Segmentation (WSM) method is applied to images to segment out Regions of Interest (ROI) that may include false material. It is a wrapper method that evaluates the performance of a machine learning algorithm using different subsets of features selected from the original dataset.

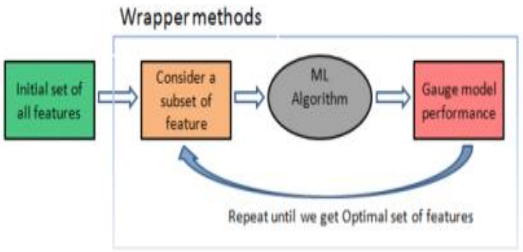


Fig. 5. WSSM Wrapper Method

The algorithm starts with an initial set of features and selects a subset of features based on their importance to the classification task. After that, the decided-upon features are put to use in the process of training an algorithm for machine learning, and the efficacy of the algorithm is assessed through the application of a cross-validation technique. The process is repeated with different subsets of features until the best subset of features is identified.

Then Watershed Segmentation (WSM) is used to segment images into basins. The goal of watershed segmentation in our work is to separate objects or regions that are visually

distinguishable in an image. We used bwareaopen = 20 and imdilate by 3 * 3. In the context of forgery detection, watershed segmentation is used to segment an image into regions of interest (ROIs) that are likely to contain forged or manipulated content. The gradient magnitude of the segmented image is calculated using

$$G_2 = \sqrt{(cH_2^2 + cV_2^2)} \qquad (2)$$

Where cH2 is the horizontal coefficients and cV2 is the vertical coefficients of the decomposed image.

Now, the Error Level Analysis (ELA) method is used to further improve ROIs. ELA is utilized in the process of ROI extraction and enhancement for forgery detection to identify portions of the image that have been modified or edited. Fig. 6 shows ELA used to extract regions of interest (ROI) from an image that is likely to have been edited. The bright areas in the ELA image is used to create a mask that highlights these regions. The mask can then be used to extract the ROI from the original image. Furthermore, ELA is used for image enhancement by increasing the contrast of the ELA image. This can make the inconsistencies in the image more visible, which can help to reveal areas of manipulation.

Fig. 6. Forged image & its ELA

Eight by eight blocks determine how much mistake there is in ELA. Two conditions can be used with ELA to explain JPEG images:

- If the mistake pattern is the same throughout all 8 × 8 blocks, then a JPEG is considered original. As a result, we can declare that the 8 × 8 pixel block has reached local minima.
- If an 8 × 8 pixel block is not at its local minimum and any 8 × 8 block has a larger error pattern, the JPEG is considered altered.

Each ROI is first classified as real or fake using an ensemble classifier, an ensemble classifier is the combination of three classifier i.e. Support Vector Machine (SVM), Random Forest (RF) and K-nearest neighbor (KNN). Choice for these three classifiers is very obvious as SVM is fast and best for outlier, RF works best with large or complex datasets whereas KNN is there to counter the limitation of SVM in terms of accuracy. If the classifier detect a fake image, a UNet classifier using DenseNet 121 as the encoder is utilized to differentiate between copy-move as well as spliced forgeries. Unet is very well known for handling large resolution pictures and as it reduce noise while preserving the original image format. DenseNet121 is a deep neural network architecture that has shown outstanding performance in various fields. It is a variant of the DenseNet family of networks, which are characterized by their 121 densely connected layers. DenseNet 121 can be used as a feature extractor to extract meaningful and discriminative features from the input images. We also tested it with Resnet50 but the overall performance is far better in Densenet121. We have used Unet Architecture for our proposed work but the results are not satisfactory, then we replaced its encoding layer with DenseNet121 that produced great results with the datasets used i.e. CASIA V2 and MICC-F2000. There are various deep learning models available for our work, we have tried CNN, DenseNet121, LeNet, GoogleNet, ResNet, VGGNet, and ConvNet out of these mentioned framework DenseNet121 produced best results metrics such as precision, recollection, and F1-score are applied to the test set to determine the model's efficacy. Hyper tuning of parameters such as epoch, batch size, learning rate etc. are done to achieve the best result of our proposed model.

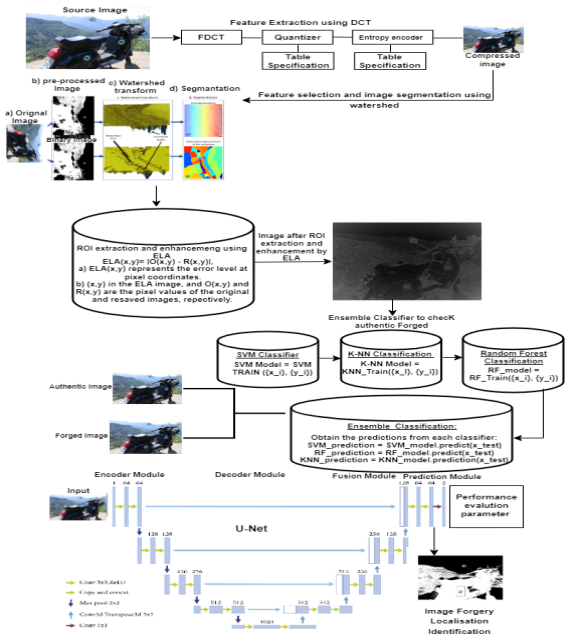


Fig. 7. Proposed methodology framework.

Fig. 7 is showcasing the steps and flow of controls in the framework proposed by us, this can also be represented in the form of algorithm. After this process, an output is generated detailing the classification outcomes for each region of interest (ROI) present in the input image. Classified ROIs are displayed in the input image with an indication of the sort of fraud (copy-move or spliced) for which they have been flagged.

A. Proposed Algorithm

Step 1: Dataset Acquisition: Collection of data from multiple resources i.e. CASIA V2 and MICC-F2000

Step 2: Pre-process data:

2.1 Outlier Removal

2.2 Noise Reduction

2.3 Color Correction

2.4 Resize Images

NumPy absolute vale is used for outlier detection, skimage is used for noise reduction, and image is converted to grayscale for color correction and image resized to 150 by 150.

Step 3: Perform feature extraction: Feature extraction from image with the help of Discrete Cosine Transform

$$F_{u,v} = \frac{1}{4} C_u C_v \left[\sum_{x=0}^{N_1-1} \sum_{y=0}^{N_2-1} f_{x,y} \cos\frac{\pi\mu(2x+1)}{2N_1} \cos\frac{\pi\mu(2y+1)}{2N_2}\right] \quad (3)$$

Where N1 x N2 are matrix, x, and y are pixels coordinates, u and v are the frequency coordinates, f(x, y) is the value of the pixel at position (x, y).

Step 4: Perform Feature Selection: Selecting the most useful features for detecting copy-move and spliced forgeries with the help of the Wrapping Subset Selection Method (WSSM) algorithm

Step 5: Image Segmentation: The Watershed Segmentation (WSM) algorithm is used to segment the input image into regions based on their similarity. This helps to isolate the regions of interest (ROIs) that may contain fraudulent content. The gradient magnitude of the segmented image is calculated using

$$G_2 = \sqrt{(cH_2^2 + cV_2^2)}$$

Where cH2 is the horizontal coefficients and CV2 is the vertical coefficients of the decomposed image.

Step 6: ROI Extraction and Enhancement: Error Level Analysis (ELA) is used to extract features

$$\begin{aligned} \overbrace{I_{A0}(i,j) - I_{B1}(i,j)}^{\text{Resavings}} &= \overbrace{ELA_1}^{\text{Recompres}} \\ I_{A1}(i,j) - I_{B2}(i,j) &= ELA_2 \\ I_{A2}(i,j) - I_{B3}(i,j) &= ELA_3 \\ &\vdots \\ &\vdots \\ I_{An}(i,j) - I_{Bn}(i,j) &= ELA_n \end{aligned}$$

Step 7: Forgery Classification: Collection of various classifiers SVM, Random Forest, KNN are being used with some parameters tuning i.e. Ensemble classifier. Applied for each ROI.

Step 8: Forgery Type Classification: For each classified ROI, if classified ROI is forged then enhanced UNet is used and encoding layers are replaced by DenseNet121. It will be used for training and testing as well. Computation of perceptual hash based on the average pixel by using

$$\text{Image\_hash} = \begin{cases} 1 & if\ pixel \geq average\_pixel \\ 0 & else \end{cases} \quad (4)$$

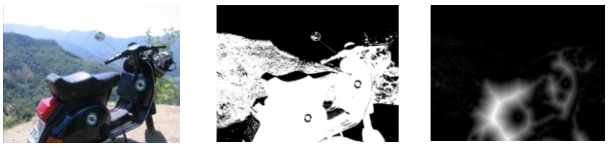
Threshold value = 20 is used to distinguished between copy-move or splicing. Check if the img_hash is below thereshold

$$\text{Image\_label} = \begin{cases} Copy\ Move & if\ Image\_hash < 20 \\ Splicing & else \end{cases} \quad (5)$$

Step 9: Outcome: display of input image with label and localization of forgery type i.e. Copy-Move forgery or Image Splicing forgery.

4. Simulation & Results

The studies used various datasets from different sources as inputs. The Media Integration Consortium's (MICC) F2000 & MICC-F220 datasets are a good example of copy-move forging. In MICC-F2000, there are 700 forgeries in this dataset created using copy-move techniques, and 1300 original photos. & In MICC-F220, there are 110 forgeries and 110 original photos. The Institute of Automation of the Chinese Academy of Sciences developed CASIA V2 to provide more information. A 60:40 split is employed to create sets of training and testing data for all datasets utilized for the research. About 60% of the data is used for training, whereas 40% is used for testing. Each dataset is trained to utilize the CNN model for 30 iterations (30 epochs) using 64-element batches. Fig. 8 and 9 shows various operation applied to an images from the datasets in order to process it and mark their true label i.e. fake or original.



(a) (b) (c)

Fig. 8. (a) Original Image (b) Binary Image & (c) Distance transform of a test image

(a) (b)

Fig. 9. (a) Region of interest (b) Error Level Analysis of a test image

A. Performance Evaluation

In this study, that is employed the accuracy calculated using Eq. 6 as a performance metric. Accuracy, recall, and F1-score are among the other metrics being investigated. When comparing the number of photos analyzed to the number of

samples, accuracy is the proportion of samples correctly identified as forgeries or originals.

Accuracy refers to a classifier's potential to identify genuine forgeries. As stated in Eq. 7, it represents the rate at which positive predictions are accurate. Recall, as demonstrated in eq. 8, measures how well a classification system can find all relevant information. Harmonically averaging the scores for memory and accuracy yields the F1-score. This is used in conjunction with the formula shown in Eq. 9 to get a single number those accounts for both accuracy and recall.

$$Accuracy = ((TP + TN))/(TP + TN + FP + FN) \quad (6)$$

$$Precision = TP/(TP + FP) \quad (7)$$

$$Recall = TP/(TP + FN) \quad (8)$$

$$F1 - Score = 2((Precision \times Recall)/(Precision + Recall)) \quad (9)$$

B. Performance Comparison Based on Dataset.

Table 1 displays comparison results across all datasets, including those containing copy-move forgery (MICC-F2000) and splicing forgery (CASIA V2).

Table 1. The Experimental Results

Framework Performance			
Model	**Accuracy**	**Precision**	**F1 Score**
U-Net [13]	-	82.9	68.6
15 Layer CNN-FIDAC [10]	84.85	87.76	85.15
DCT & DWT Based [17]	97	-	-
CNN-CovLSTM [11]	95	-	-
CNN-DenseNet based Segmentation for CASIA dataset [Proposed]	98	92	90
CNN-DenseNet based Segmentation for MICC-F2000 dataset [Proposed]	99	98	99

The performance of the proposed work is stated by 1254 total training samples and 537 total testing samples for CASIA V2 dataset. For MICC-F2000 dataset, total number of training samples are 1800 and total number of testing samples are 720 which is well above in comparision to the Elaskily et al. [11]. Table 1 demonstrate that the Hebbar, Nagaveni K., and Ashwini S. Kunte [13] model has U-Net feature achieved precision 82.9% with no reported accuracy and F1-score of 68.6%. And Pawar et al., [10] achieved an accuracy of 84.85% with 87.76% of precision and F1-score of 85.15%. In this table '-' indicates that the particular evaluation metrics is not mentioned in the respective literature.

Based on the provided table, the results can be summarized as follows:

For the CASIA-2 dataset:

- Jaiprakash et al. [17] achieved an accuracy of 97% with no reported precision and F1 score.
- The proposed model achieved a higher accuracy of 98% compared to Jaiprakash et al.'s model.
- The proposed model demonstrated a precision of 92% and an F1 score of 90%.

For the MICC-F2000 dataset:

- Elaskily et al. [11] achieved an accuracy of 95% with no reported precision and an F1 score.
- The proposed model outperformed Elaskily et al.'s model with a higher accuracy of 99%.
- The proposed model demonstrated a precision of 98% and an impressive F1 score of 99%.

The results are summarized in the table below, which shows that the CNN approach is superior to the state-of-the-art methods in differentiating between splicing and copy-move forgeries (90% vs. 96%). If authors look at the accuracy throughout the whole dataset, however, CASIA V2 dataset has the best results (97% accuracy). This can be explained by the increased number of training photos and the wider variety of forgeries included in the dataset. On the other hand, when tested on the merged MICC-F2000+CASIA V2 the data set, performance dropped to 98% accuracy.

Table 2. Training and validation parameters in UNET.

Dataset	**Training Loss**	**Validation Loss**
MICC-F2000	0.18	0.16
CASIA V2	0.15	0.19
Dataset	**Training accuracy**	**Validation accuracy**
MICC-F2000	95%	99%
CASIA V2	96%	98%

Table 2 presents the training and validation loss, as well as the training and validation accuracy for two datasets: MICC-F2000 and CASIA V2.

For the training and validation loss:

- For the MICC-F2000 dataset, the training loss is 0.18, while the validation loss is 0.16 percent.

• For the CASIA V2 dataset, the training loss is 0.15, and the validation loss is 0.19 percent.

The training loss represents the error or discrepancy between the predicted and actual values during the training phase as shown in Fig. 10 and 11, while the validation loss indicates the error during the validation phase, which is performed on a separate dataset as shown in Fig. 12 and 13.

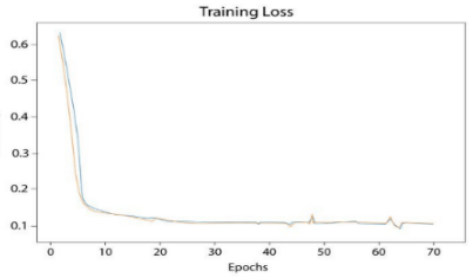


Fig. 10. Training loss of MICC-F2000 dataset

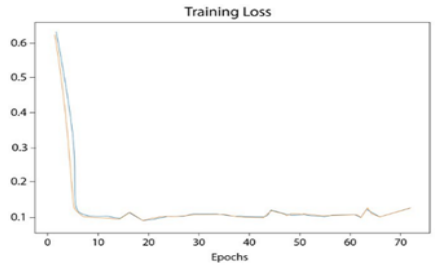


Fig. 11. Training loss of CASIA V2 dataset

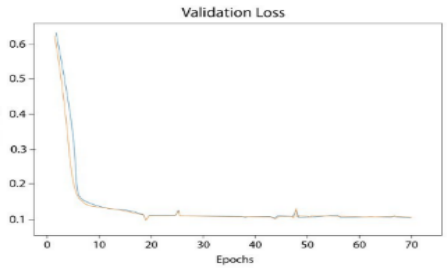


Fig. 12. Validation loss of CASIA V2 dataset

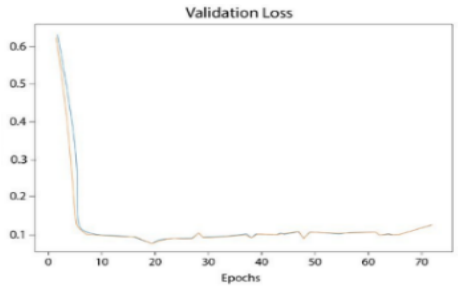


Fig. 13. Validation loss of MICC-F2000 dataset

For the training and validation accuracy:

- For the MICC-F2000 dataset, the training accuracy is 95%, and the validation accuracy is also 99%.
- For the CASIA V2 dataset, the training accuracy is 96%, and the validation accuracy is also 98%.

The training accuracy represents the proportion of correctly predicted samples during the training phase as shown in Fig. 14 and 15., while the validation accuracy indicates the proportion of correctly predicted samples during the validation phase as shown in Fig. 16 and 17.

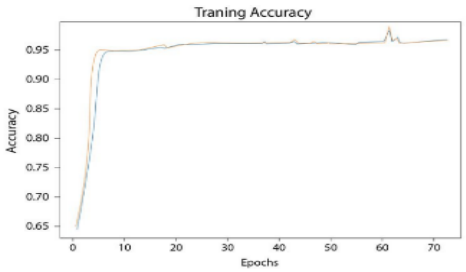


Fig. 14. Training Accuracy of MICC-F2000 dataset

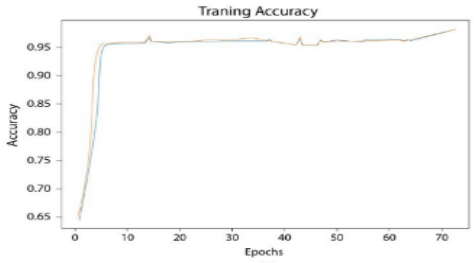


Fig. 15. Training Accuracy of CASIA V2 dataset

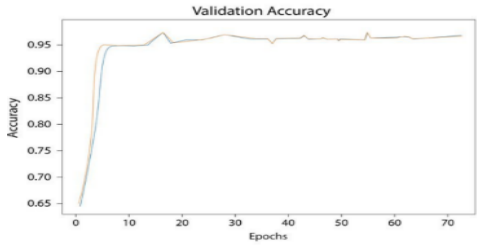


Fig. 16. Validation Accuracy of CASIA V2 dataset

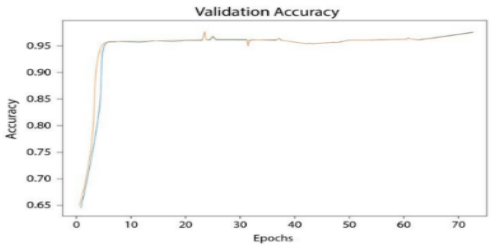


Fig. 17. Validation Accuracy of MICC-F2000 dataset

Fig. 18 depicts the accuracy graph of the compared work for the technologies ELA-Unet [13], 15 Layer CNN-FIDAC [10], DCT & DWT Based [17] and CNN-CovLSTM [11] on MICC-F2000 and CASIA V2 datasets and our proposed model clearly shows great accuracy to the technologies compared.

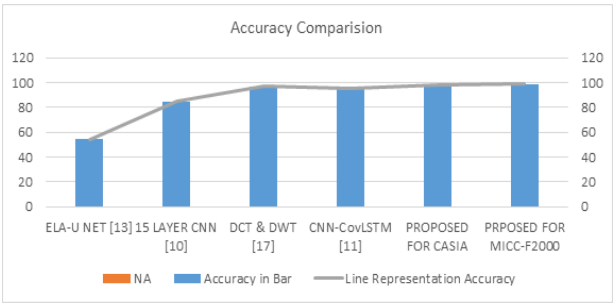


Fig. 18. Validation Accuracy of MICC-F2000 dataset

Fig. 19., shows a U-Net (Copy-Move) being used to detect and localize image forgeries. This technique makes use the U-Net deep learning architecture to detect copy-move deception, in which a section of a picture is copied and pasted to a new location. The technology employs robust image processing techniques to identify and locate these shifts. This technique makes a significant contribution to digital photo forensics and the identification of image alteration by efficiently detecting copy-move forgeries using the U-Net model.

Fig. 19. U-Net Based Image Forgery Localization Identification (Copy-Move)

Fig. 20. U-Net Based Image Forgery Localisation Identification (Image Splicing)

In Fig. 20., see an example of a method that uses a deep learning U-Net architecture to detect and localize picture splicing (image forgery). Using this technique, which may perform forensic analysis and verify the authenticity of an image to see if it has been tampered with in any way.

5. Conclusion

In conclusion, a two-stage approach is being used, beginning with coarse-grained segmentation using a modified version of the U-Net structure, and then moving on to fine-grained segmentation using a hybrid dense net. The suggested system is successful, exceeding the performance of DCT & DWT and CNN-CovLSTM detection algorithms that are currently available, according to extensive testing carried out on a diverse dataset. When compared to models developed by Jaiprakash et al. and Elaskily et al., the suggested model achieves better results on the CASIA-2 and MICC-F2000 datasets, as measured by accuracy, precision, and F1 score. On the CASIA-2 dataset, the suggested model outperforms Jaiprakash et al.'s model with an accuracy of 98%. It also demonstrates a precision of 92% and an F1 score of 90%. For the MICC-F2000 dataset, the proposed model achieves an accuracy of 99%, outperforming Elaskily et al.'s model. It exhibits a precision of 98% and an impressive F1 score of 99%. Overall, the proposed model demonstrates superior performance compared to previous models and shows promising results in accurately detecting image forgery. However, further analysis and evaluation are necessary to validate the proposed model's effectiveness in real-world scenarios and to ensure its robustness across diverse datasets.

References

[1] Xiao, B. Wei, Y. Bi, X. Li, W. Ma, J., “Image splicing forgery detection combining coarse to the refined convolutional neural network and adaptive clustering” Information Sciences 2020, 511, 172–191.

[2] Wu, Y. Abd Almageed, W. Natarajan P, “ManTra-Net: Manipulation Tracing Network for Detection and Localization of Image Forgeries with Anomalous Features” In Proceedings of the 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), Long Beach, CA, USA, 15–20 June 2019; pp. 9535–9544.

[3] Li X, Jing T, Li X., “Image splicing detection based on moment features and Hilbert-Huang Transform” In IEEE international conference on information theory and information security (ICITIS), 2010; Beijing, China; 1127–1130

[4] Castillo Camacho, I. Wang, K, “A Comprehensive Review of Deep-Learning-Based Methods for Image Forensics” J. Imaging 2021, 7, 69.

[5] Hsu J, Yu F. “Image tampering detection for forensics applications, Citeseer”, 2009.

[6] Liu G, Wang J, Lian S, Wang Z, “A passive image authentication scheme for detecting region-duplication forgery with rotation” Journal of

Network Computing & Application 2011; 34:1557–1565.

[7] Mushtaq S, Ajaz H, “Novel method for image splicing detection” In International conference on advances in computing, communications, and informatics (ICACCI); 2014; Delhi, India; 2398–2403.

[8] Krishnaraj, N., B. Sivakumar, Ramya Kuppusamy, Yuvaraja Teekaraman, and Amruth Ramesh Thelkar, "Design of automated deep learning-based fusion model for copy-move image forgery detection." Computational Intelligence and Neuroscience 2022 (2022).

[9] Koul, Saboor, Munish Kumar, Surinder Singh Khurana, Faisel Mushtaq, and Krishan Kumar, "An efficient approach for copy-move image forgery detection using convolution neural network." Multimedia Tools and Applications 81, no. 8 (2022): 11259-11277.

[10] Pawar, Shraddha, Gaurangi Pradhan, Bhavin Goswami, and Sonali Bhutad, "Identifying fake images through cnn based classification using fidac." In 2022 International Conference on Intelligent Controller and Computing for Smart Power (ICICCSP), pp. 01-06. IEEE, 2022.

[11] Elaskily, Mohamed A., Monagi H. Alkinani, Ahmed Sedik, and Mohamed M. Dessouky, "Deep learning based algorithm (ConvLSTM) for copy move forgery detection." Journal of Intelligent & Fuzzy Systems 40, no. 3 (2021): 4385-4405.

[12] Agrawal, Prateek, Deepak Chaudhary, Vishu Madaan, Anatoliy Zabrovskiy, Radu Prodan, Dragi Kimovski, and Christian Timmerer, "Automated bank cheque verification using image processing and deep learning methods." Multimedia Tools and Applications 80 (2021): 5319-5350.

[13] Hebbar, Nagaveni K., and Ashwini S. Kunte, "Image forgery localization using U-net based architecture and error level analysis." In 2021 3rd International Conference on Advances in Computing, Communication Control and Networking (ICAC3N), pp. 1992-1996. IEEE, 2021.

[14] Abhishek, and Neeru Jindal, "Copy move and splicing forgery detection using deep convolution neural network, and semantic segmentation." Multimedia Tools and Applications 80 (2021): 3571-3599.

[15] Saber, Akram Hatem, Mohd Ayyub Khan, and Basim Galeb Mejbel, "A survey on image forgery detection using different forensic approaches." Adv Sci Technol Eng Syst J 5, no. 3 (2020): 361-370.

[16] Gani, Gulnawaz, and Fasel Qadir, "A robust copy-move forgery detection technique based on discrete cosine transform and cellular automata." Journal of Information Security and Applications 54 (2020): 102510.

[17] Jaiprakash, Sahani Pooja, Madhavi B. Desai, Choudhary Shyam Prakash, Vipul H. Mistry, and Kishankumar Lalajibhai Radadiya, "Low dimensional DCT and DWT feature based model for detection of image splicing and copy-move forgery." Multimedia Tools and Applications 79 (2020): 29977-30005.

[18] Bappy, Jawadul H., Cody Simons, Lakshmanan Nataraj, B. S. Manjunath, and Amit K. Roy-Chowdhury, "Hybrid lstm and encoder–decoder architecture for detection of image forgeries." IEEE Transactions on Image Processing 28, no. 7 (2019): 3286-3300.

[19] Wu, Yue, Wael Abd-Almageed, and Prem Natarajan, "Image copy-move forgery detection via an end-to-end deep neural network." In 2018 IEEE Winter Conference on Applications of Computer Vision (WACV), pp. 1907-1915. IEEE, 2018.

[20] Yao, Ye, Yunqing Shi, Shaowei Weng, and Bo Guan, "Deep learning for detection of object-based forgery in advanced video." Symmetry 10, no. 1 (2017): 3.

[21] Anshul. K. Singh, Chandani Sharma and B. K. Singh, “A review on Automatic Image Forgery Classification Using Advanced Deep Learning Techniques”, ICDIS, Advances in Data & Information Sciences, Springer Nature, 25 Nov, 2022.

[22] Anshul. K. Singh, Chandani Sharma and B. K. Singh, “Image Forgery Localization and Detection using Multiple Deep Learning Algorithm with ELA” ICFIRTP 2022, IEEE, 2023.