FORGERY DETECTION IN DIGITAL IMAGES USING CNN

A Major Project II Report
Submitted in Partial Fulfilment of the Requirements
for the Award of the Degree Of

MASTER OF TECHNOLOGY

in

Information Systems

by

Shujaa Ahmad
(Roll No. 2K22/ISY/18)

Under the Supervision of

Dr Bindu Verma

Assistant Professor, Delhi Technological University

Department of Information Technology

DELHI TECHNOLOGICAL UNIVERSITY
(Formerly Delhi College of Engineering)
Shahbad Daulatpur, Main Bawana Road, Delhi – 110042, India

May, 2024

ACKNOWLEDGEMENTS

I express my gratitude to Dr. Bindu Verma, my supervisor and an assistant professor in the Department of Information Technology at Delhi Technological University, as well as all the other faculty members in the department. They were all helpful anytime I needed it. Without their guidance and support, none of this effort could have been done. In addition, I am thankful to the department and university for offering all of the resources needed to carry out the work.

SHUJAA AHMAD

DELHI TECHNOLOGICAL UNIVERSITY

(Formerly Delhi College of Engineering)
Shahbad Daulatpur, Main Bawana Road, Delhi-42

CANDIDATE'S DECLARATION

I **Shujaa Ahmad** hereby certify that the work which is being presented in the major project II report entitled **"Forgery Detection in Digital Images using CNN"** in partial fulfilment of the requirements for the award of the Degree of Master of Technology, submitted in the Department of Information Technology, Delhi Technological University is an authentic record of my own work carried out during the period from 2022 to 2024 under the supervision of **Dr. Bindu Verma**.

The matter presented in the thesis has not been submitted by me for the award of any other degree of this or any other institute.

Shujaa

Candidate's Signature

Date: 11/07/2024

DELHI TECHNOLOGICAL UNIVERSITY

(Formerly Delhi College of Engineering)
Shahbad Daulatpur, Main Bawana Road, Delhi-42

CERTIFICATE BY THE SUPERVISOR

Certified that **Shujaa Ahmad** (2K22/ISY/18) has carried out their search work presented in this major project II report entitled **"Forgery Detection in Digital Images using CNN"** for the award of **Master of Technology** from the Department of Information Technology, Delhi Technological University, Delhi, under my supervision. The report embodies results of original work, and studies are carried out by the student himself and the contents of the thesis do not form the basis for the award of any other degree to the candidate or to anybody else from this or any other University/Institution.

11/7/24

Signature

Dr. Bindu Verma

(Assistant Professor)

Department of Information Technology

Delhi Technological University, Delhi

Date: 11/07/2024

Forgery Detection in Digital Images using CNN

Shujaa Ahmad

ABSTRACT

With the widespread availability of smartphones equipped with digital cameras and other inexpensive imaging equipment, as well as the low cost of internet connectivity, there is a growing interest in capturing and sharing images online. Many image-editing and enhancement applications are freely available for smartphones and computers, and while they are frequently used to improve and enhance image quality, they are also occasionally used to make alterations to the image in order to propagate false information or slander someone.

Image forging is the process of concealing details or adding something that was not o riginally part of the image, causing the integrity of the image to be compromised. For example, a forged driving licence provides false information about a person's ability to drive. Also, digital images are now often utilised as evidence in the media and courts. Therefore, it becomes extremely important to spot these tampered and forged images.

In this work, the various types of image forgeries are explored. This study includes a review and comparison of several detection methods for image forgeries. To solve this issue, a CNN network has been built based on previous research and its performance was compared on two different datasets. A data augmentation strategy, as well as multiple hyperparameter tuning, are also evaluated in terms of classification accuracy. Our results suggest that the outcomes are highly dependent on the difficulty of the dataset. In this research, we proposed a Convolutional Neural Network (CNN) model for detecting image forgery.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1

INTRODUCTION

1.1 Overview

Digital image forging is the act of altering a digital image's contents in such a way that the message or meanings of the image are changed without any traces being left behind. The availability of inexpensive and high-quality digital imaging equipment in the pockets of individuals, combined with mobile internet and social networking applications such as Instagram, WhatsApp, etc. has increased the utilization and generation of digital photos in this internet world. Moreover, image-editing applications like Adobe Photoshop, Snapseed, Picsart are commonly employed by social media users for some enhancements and beautification. Although this use may be accepted, it is also not uncommon to manipulate the photographs such that they provide misleading information or defame someone.

The main problem here is that if the forgery is done professionally, it becomes very difficult for us humans to tell the difference between the real and the counterfeit. During the COVID-19 pandemic, it was noticed that some people employed image editing applications to forge their medical reports that contained their test results from positive to negative in order to travel freely. This demonstrates the sensitivity of the matter; it may be a very hazardous scenario not just for those engaging in such activities, but also for the rest of us. Having said that, digital images are frequently utilized as evidence in the media and courts, making it crucial to distinguish between the real and the counterfeit.

1.2 Forgery Detection Techniques

Forgery detection methods are classified into two categories [1]. Fig. 1.1 shows the pictorial representation.

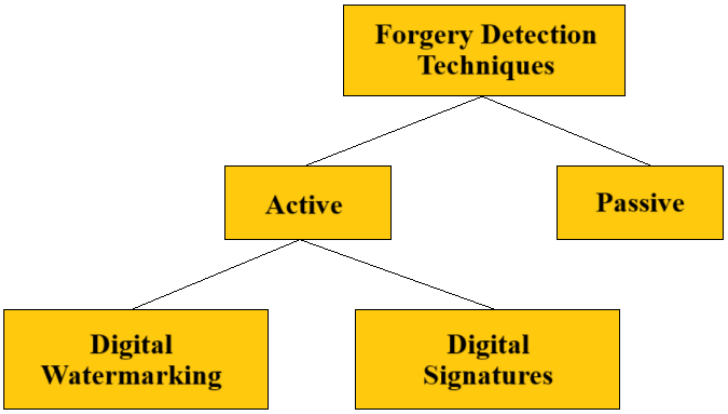


Fig. 1.1 Forgery Detection Techniques

1.2.1 Active Techniques

The first is an active process that seeks to extract hidden information from a picture as it goes through its life cycle before being formed. Digital watermarks and digital signatures are used as concealed information. Digital watermarking entails embedding a watermark using a secret key, and because images are frequently changed, this watermark is vulnerable to accidental exploitation. In order to detect any changes, the watermark is retrieved from the picture and compared to the original watermark. Similarly, digital signatures are used to identify changes to an image. They work by generating a hash of the image that is unique for an image, and if the image is altered in any way, the same hash cannot be obtained again.

1.2.2 Passive Techniques

The passive forgery detection approach, unlike the active technique, deals only with the counterfeit picture and no additional hidden information. The passive technique examines the statistics and semantics of the image to detect alteration. As illustrated in Fig. 1.2, it falls into one of the following categories: camera-based, pixel-based, physical environment-based, geometry-based and format-based. **[2]** Fig. 1.2 shows the pictorial representation.

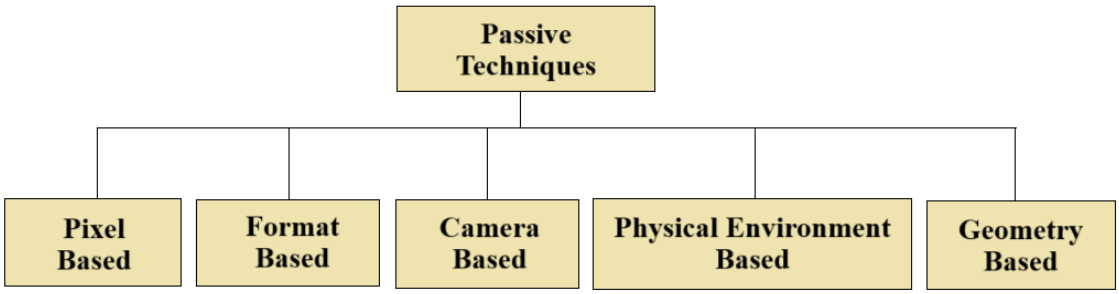


Fig. 1.2 Passive Techniques

- **Pixel Based:** Pixel-based techniques try to detect forgery by identifying statistical inconsistencies at the pixel level or arithmetic changes in pixel values.

- **Format Based:** Format-based refers to one of the picture formats, typically JPEG. JPEG artifact-based techniques identify forgeries on images based on the JPEG image format. The artifacts that remain in the image after JPEG compression or JPEG blocking are inspected, and any variations in the artifacts indicate a forgery

- **Camera Based:** In the camera-based technique, any inconsistencies in the digital footprints introduced by camera components such as the sensor, camera lens, and so on **[3]** are analysed to identify forgery.

- **Physical Environment Based:** Physical Environment-based approaches look at lighting variations, shadows, and reflections in images to detect forgeries **[4]**. It involves finding inconsistencies in the lighting of the environment in which the image was taken. It is simpler to detect forgery using this method since it is difficult to match the lightening, shadows, and reflection of the original image once it has been obtained and modified.

- **Geometry Based:** Geometric-based techniques involve detecting irregularities in an object's geometric measurements and relative position in relation to the camera.

1.3 General Forgery Detection Procedure

The following steps serve as the foundation for all digital image forgery detection techniques [5] in general:

- **Pre-Processing:** This is an optional step; however, it is essential for some models. This stage involves converting the photos from colour to greyscale or some other colour model like YCbCr or HSV. It speeds up processing while reducing the image dimensionality.

- **Feature Extraction:** Several features from the image are extracted to help with forgery detection. “An image can be transformed from spatial to frequency domain.” Because the relevant information is conveyed by a few coefficients, certain high frequency components, such as edges, can be eliminated while low frequency components, which are necessary for detection, can be emphasised.

- **Feature Matching:** The extracted features are compared in order to determine similarity between them. To make searching easier, we sort all of the features and use a distance metric such as Euclidean distance to find similar features. Some models may also employ clustering. Thresholding is used to deal with some false positives that occur in this step.

- **Forgery Detection and Localization:** Various algorithms, including DCT coefficient analysis and convolution neural networks, are employed to locate and detect the forgery.

CHAPTER 2

LITERATURE REVIEW

2.1 Traditional Methods

The traditional methods can be broadly classified as block-division-based or key points extraction-based. Alkawaz et al., 2018 in [6] proposed a methodology for the detection of image cloning (copy-move) forgery using DCT coefficients, as well as investigating the effect of block size on the performance of a forged portion by implementing the approach with block sizes ranging from 4x4 to 8x8. Furthermore, it was discovered that a block size of 8x8 provided greater accuracy than a block size of 4x4. Based on a similar methodology, Dua et al., 2020 in [7] proposed a method for identifying both image-cloning and image-splicing forgeries. “First, the RGB image is transformed to YCbCr colour format. Image Y in a YCbCr represents the luminance component, whereas Cb and Cr represent the chroma channels. Since most forgery traces are concealed in the chroma channels and may be missed by the human eye, all three are used in this method. The majority of information that leads to improved luminance component detection by human eyes is carried by the Y channel.” To identify forgeries, they used an “Support Vector Machine” or SVM classifier. With CASIA 1.0 and 2.0 datasets, this technique achieved average identification rates of more than 93% and 98%, respectively.

Rathore et al., 2021 [8] in proposed a method for detection of copy-move image tampering based on singular value decomposition, which begins with converting the input image to grayscale with overlapping blocks of fixed size. They used SVD (Singular Value Decomposition) for dimensionality reduction since it can keep relevant features while requiring less storage space, and BWT (Biorthogonal Wavelet Transform) for feature extraction on each block. The extracted features are arranged lexicographically and for improved accuracy, a threshold parameter established using Minkowski distance is employed. Tampering detection and classification are carried out using IRVM (Improved Relevance Vector Machine). When compared to existing

schemes, the suggested IRVM technique attains an accuracy rate of 92.22% on the "CoMoFoD dataset". The CoMoFoD dataset is a small dataset that came in 2013 to help in the understanding of snippet transformations like as translation, rotation, and scaling. [9]

Vega et al., 2020 in [10] proposed a JPEG compression-based technique that focuses on areas with varied levels of compression within the same image. "A compressed JPEG image must have roughly the same level throughout its content." Any variation in an area's error level suggests a high probability of manipulation. The proposed approach produced a 73.3% accuracy rate on the CASIA v1.0 dataset, and it performs better for images with higher resolution. Hegazi et al., 2021 in [11] presented a model based on key-point extraction for copy-move forgery detection. "The key-points extraction is done using Scale Invariant Feature Transform or SIFT as it is invariant to scaling and rotation and is proved to be the most accurate and stable feature detector". They created a two-stage outlier detection mechanism using the GORE (Guaranteed Outlier Removal) and RANSAC (Random Sample Consensus) algorithms and it has been shown to bring down the false negatives. They tested the proposed model on the MICC-F220 dataset and achieved F1 score equal to 97.56%.

2.2 Deep Learning Based Methods

Deep leaning is a popular concept that is applied in many fields. Deep learning mainly depends on CCN. CNN stands for convolution neural network. CNN has recently gained popularity due to its capacity to understand image attributes. They are commonly used in image recognition and image analysis activities [12]. The main aim of a CNN is to extract relevant features from an image. CNN has multiple layers: convolution layers, maximum pooling layers, flattening-layer, and fully-connected or dense layers [13].

Ali et al., 2022 in [14] presented a CNN model for the detection of image splicing and image cloning tampering by recompressing the image. Because the source of the tampered portion and the background picture differ, when the forged image is recompressed, the forged part compresses differently because of the compression

difference. The given tampered image is recompressed, and the difference between the tampered and recompressed images is computed, resulting in an image with the forged region highlighted. A CNN-based network is trained using the computed difference as an input feature. On the CASIA 2.0 dataset, they attained an accuracy of 92.23%. Sudiatmika et al., 2019 in [15] presented a novel approach for tackling the issue of distinguishing genuine from counterfeit images. They did this by combining VGG16 and ELA. They were able to attain an accuracy of 88.46% on CASIA 2.0 dataset.

Kaul et al., 2022 in [16] presented a deep learning-based CNN model with multiple convolution and pooling layers that has achieved high performance with low computational cost for copy-move image forgery. They used "three convolutional layers, three max pooling layers, four dropout layers, two dense layers and one flattering layer". They tested it on the MICC-F2000 dataset and achieved an accuracy of 97.5%. Doegar et al., 2019, in [17] proposed a deep learning approach based on pre-trained existing AlexNet architecture. The AlexNet model has 25 layers, with the key layers being convolution, pooling, fully connected, softmax, and the ReLU activation function. The proposed method is carried out in two stages: first, an SVM classifier is trained using a pre-trained AlexNet model based on deep features, and second, images are supplied to the classifier to tell fake from real. It was tested using the publicly accessible dataset MICC-F220 and achieved an accuracy of 93.94%.

Elaskily et al., 2020 in [18] proposed another deep learning technique in which a CNN model consisting of six convolution layers, each succeeded by a maximum pooling layer, was constructed. The "Global Average Pooling or GAP layer is used to cut down the number of parameters in the network and reduce the risk of overfitting". Since MICC F-600, MICC F-220, MICCF-2000, and SAT 130 are all very tiny datasets, they combined them to train a CNN. They achieved very high accuracy, but only tested a small portion (10%) of the composed dataset. Because of this, there may be a strong correlation between test and training images. Ouyant et al., 2017 in [19] proposed a CNN model to predict a picture as either tampered or authentic. They used transfer learning to predict if an image is fabricated or legitimate by using AlexNet as the base model with a modified classification layer. The ImageNet dataset was used to initialize

the model weights. They created artificial copy-move forgery by randomly copying rectangles from a photo and pasting them in multiple regions on the same photo in order to counteract overfitting. It appears that the model did well on the custom datasets but poorly on a real-world dataset known as CMFD. “It is composed of 48 source images in which a total of 87 regions (referred by the authors as “snippets”), with different sizes and content (from smooth areas, (e.g.), the sky, to rough ones, (e.g.), rocks, to human-made, (e.g.), buildings) are manually selected and copy-moved.” [20]

Majumder et al., 2018 in [21] advocated a shallow CNN model with only two convolution layers without any pooling layer. The first and second convolutional layers were given higher filter sizes of 32 × 32 and 20 x 12, respectively. The strategy relies on the idea that, in CNNs, simpler visual structures, like edges and corners, are learned in the initial layers, while more intricate high-level features are learned at deeper ones. Hence, low-level traits are more likely to be helpful in detecting the artifacts generated by forging processes. This decision results in a limited amount of network parameters, lowering the risk of over-fitting during training. They attained an accuracy of approximately 79% without any dataset pre-processing. Kwon et al., 2021 in [22] proposed CAT-Net (Compression Artifact Tracing Network), a CNN capable of detecting image splicing forgery. It used a fusion of two different streams where one stream learns visual artifacts, while the other stream, which has been pre-trained with double JPEG compression detection, learns compression artifacts. This model when retrained on images containing splicing forgery from CASIA v2 dataset attained an accuracy of 87.29%.

Tyagi et al., 2023 in [23] proposed a CNN-based model called MiniNet capable of detecting both image-splicing and image-cloning forgeries. Their primary goal was to create a lightweight CNN model with high accuracy that could also be employed on slower devices. They trained and tested their model on a variety of image forgery datasets, achieving an accuracy of around 93% on the CASIA dataset. Wu et al., 2019 in [24] presented ManTra-Net, a fully convolutional network capable of handling a wide range of forgery types and sizes, including splicing, removal, copy-move,

augmentation, and even unknown forms of forging. Retraining and testing them on the CASIA 2.0 dataset yielded an accuracy of roughly 56.15%.

In conventional methods, feature extraction is done manually by “block-based methods” or “key-point-based methods” whereas in the deep learning approaches the features are learned automatically. In block division methods computation cost is high and they have trouble with images having large-scale distortion. Key-point extraction methods incur a lower computational cost but are generally inefficient for images which are smoothened. Some hybrid techniques use both block division and key-point extraction to balance the trade-off. Deep learning approaches with their automatic feature learning are very popular in multiple domains. Most of the researchers are now focusing on deep learning methods. However, “the performance of the deep-learning models relies solely on the amount and quality of the training data” and in some cases, it has been seen that they perform worse than block and key point-based methods due to less amount of data. It is also worth noting that a model's performance is heavily dependent on the dataset employed. If the dataset is too easy or contains unrealistic forgery, a model will be able to detect it more accurately. For instance, compared to the CASIA datasets, the MICC datasets contain easier forgeries. A dataset can contain one or more types of picture forgery, such as the CoMoFoD dataset, which contains just copy-move image forging but also includes post-processing procedures such as brightness, contrast modification, noise addition, blurring and compression.

CHAPTER 3

BACKGROUND STUDY

3.1 Digital Image Forensics

“Digital image forensics” is an area of study whose aim is to identify forged images. DIF primarily addresses two issues: determining the origin of an image and ensuring the image's integrity. To tackle them, we must first understand the “digital image life cycle” and the various types of footprints that are left on the image prior to alteration.

3.1.1 Digital Image Life Cycle

The three phases of the digital image life cycle are image acquisition, image coding, and image editing [25]. Fig. 3.1 displays the pictorial representation.

Fig. 3.1 Life Cycle of a Digital Image

A digital image goes through a series of stages starting from the moment it is acquired before it is fabricated. The image is left with certain digital footprints from each phase that can be examined to identify the counterfeit. The first stage is image acquisition, which is the process of capturing an image with any instrument such as a camera or a scanner. The device used leaves a digital footprint on the image. The camera's lens, sensor, digital-to-analog (D/A) converter, etc. can all have an impact on the footprint. The second stage, image coding, involves saving the image and compressing it, which in turn leads to the formation of a digital footprint. The third stage of the process, image editing, includes post-processing of the image to enhance its visual representation. These post-processing techniques frequently serve only as a platform to perform forgery.

3.2 Types of Forgery in Digital Images

Mainly there are three types of image forgeries [2] which are discussed below:

3.2.1 Image Cloning

Copy-move or image cloning forgery entails copying a portion of a picture into another place inside the same image. In this case, the source and destination images are the same. It could be done to conceal or duplicate some portion. Since the forged region is just a portion of the actual image, it is harder to spot forgeries because the properties of the duplicated portion and the original image match identically. Fig. 3.2 is an example of image cloning.

a.) shows the actual image, b.) shows the cloned image.

Fig. 3.2 Image Cloning (Copy Move) Forgery Example [26]

3.2.2 Image Splicing

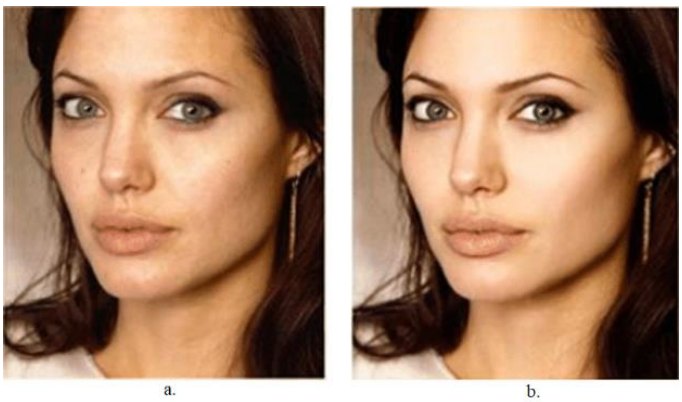
Image splicing is the process of inserting a part of a donor image into a region of a forged image. Multiple portions can be taken from multiple donor images. The attributes of the original portion and the forged portion in this instance are not exactly the same, which can aid in the forgery detection process. Fig. 3.3 shows an example of copy-paste forgery (image splicing).

a.) shows the unforged image, b.) shows the donor image and c.) shows the forged image.

Fig. 3.3 Image Splicing (Copy Paste) Forgery Example [27]

3.2.3 Image Retouching

Image retouching involves modifying image properties through operations such as rotation, flipping, scaling, and so on to improve the overall appearance of the picture. These are often used by the mass media on items like magazine covers, posters, etc., although social media users frequently use them as well. This sort of image modification, while less harmful, can be ethically misleading. Fig. 3.4 shows an example of image retouching forgery.



a.) shows the actual photo, b.) shows the retouched photo

Fig. 3.4 Image Retouching Forgery Example [28]

CHAPTER 4

PROBLEM STATEMENT

The main objective of this study is to evaluate how the performance of a "Convolutional Neural Network" or CNN in detecting image forgery varies based on the complexity of the dataset. To fulfil that purpose a "CNN based deep-learning model" is presented to detect both copy-move or image-cloning and copy-paste or image-splicing forgeries. Two popular datasets, MICC F-2000 and CASIA 2.0, were chosen for this study to examine the performance of a CNN model on both.

While many authors have achieved very high accuracies on MICC datasets, we assume that their performance will degrade significantly when tested on the CASIA v2 dataset because MICC datasets do not contain very realistic image forgeries. We believe that the classification performance decreases with increasing data difficulty.

The 'CASIA v2 dataset' is recommended over the 'CASIA v1 dataset' because of its higher sample count. This holds great significance as prior studies have demonstrated that an enormous quantity of training data is necessary for a somewhat well-trained CNN. In model evaluation, the main metric is the classifier's accuracy during testing. This choice was made because it is the main metric used in earlier deep-learning studies to evaluate the ability to identify picture forgeries.

CHAPTER 5

PROPOSED METHOD

"CNNs are designed to be non-linear, networked neurons that are modelled after the human visual system". Their remarkable potential has already been proven in several 'computer vision applications', including 'object detection' and 'image segmentation'. They might also be useful for a number of other uses, such as image forensics. Image forgery can be easily accomplished with the many tools accessible today, and since it is so harmful, it is important to identify it when it occurs. Because of the disparate origins of the photographs, a variety of artifacts arise when an image fragment is moved from one location to another. CNNs have the ability to identify these artifacts in manipulated images even if they might not be visible to the unaided eye. While the compression levels over the entire image must be the same for a non-altered image, they are different for an altered image Using the above concept and taking inspiration from [14], a CNN-based deep learning model is developed that can detect both copy-move (image cloning) and copy-paste (image splicing) picture tampering. The Fig. 5.1 displays the flow chart of the proposed method.

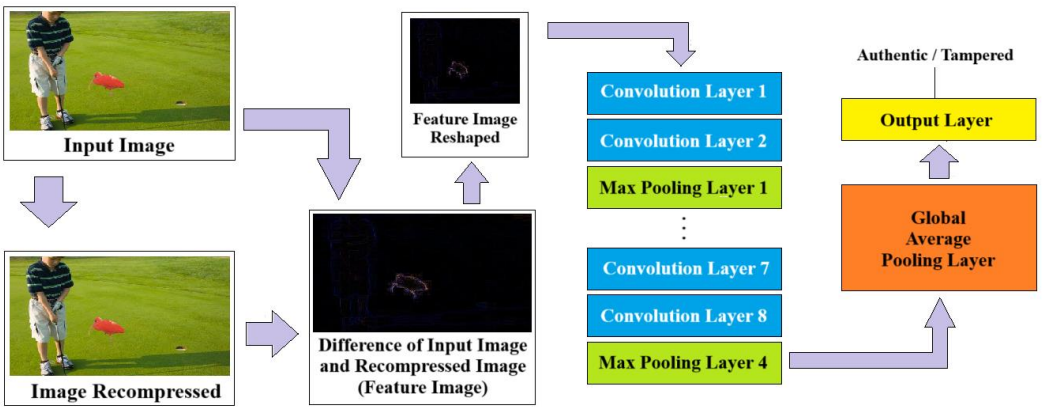


Fig. 5.1 Flow Chart of the Proposed Method

5.1 Network Architecture

"Convolutional Neural Network or CNN which is a deep-learning neural network mainly used for image-based problems". A basic CNN model is made up of many 'convolutional layers', followed by 'fully-connected layers' and a 'classification layer'. The convolution layer is the core layer that performs the convolution operation on an image using a kernel that is significantly smaller in size than the real image. This produces a 2D matrix known as an activation map, which illustrates how the kernel reacts at every spatial location within the picture. The non-linear activation function is another crucial component of this layer which is applied to the output of linear processes such as convolution. "Rectified Linear Unit or ReLU is a prominent non-linear activation function. The polling layer is responsible for feature extraction and dimension reduction. 'Average pooling' and 'max-pooling' are the two main types of pooling. The flattening layer turns the inputs from the pooling layer into a one-'dimensional array', and the 'fully-connected layer' connects all of the neurons.

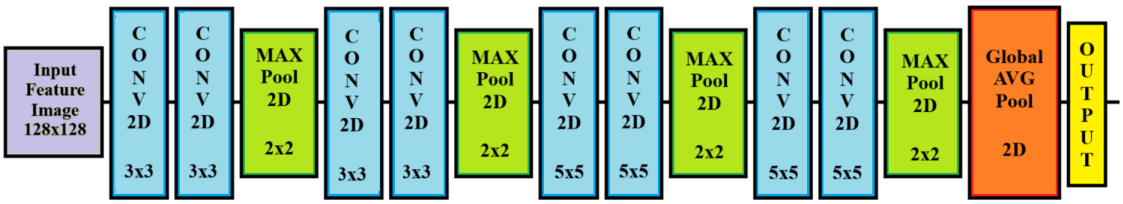


Fig. 5.2 Proposed Network Architecture

Fig. 5.2 shows the architecture of the proposed network. The network is made up of eight convolutional layers organized in pairs of two, followed by a max-pooling layer with a pool size of 2x2. The first four convolutional layers have a 3x3 kernel size, whereas the subsequent layers are 5x5. The convolution layer is fed a 128x128x3 picture, with 3 representing the RGB colour channels. A global average polling layer is added instead of a fully connected layer as it reduces the number of parameters and preserves spatial information. Then finally a dense layer with sigmoid activation and containing a single neuron is used for classification of an image into forged or authentic.

As mentioned earlier, if a picture contains a forgery, the forged section will compress differently from the rest of the image when it is recompressed since the source of the forged portion differs from the source of the actual image. This significantly highlights the altered component when comparing the original image to its recompressed counterpart. We thus use this to train our model for image tampering detection. Algorithms 5.1 enlists the steps involved in forgery detection in the proposed model.

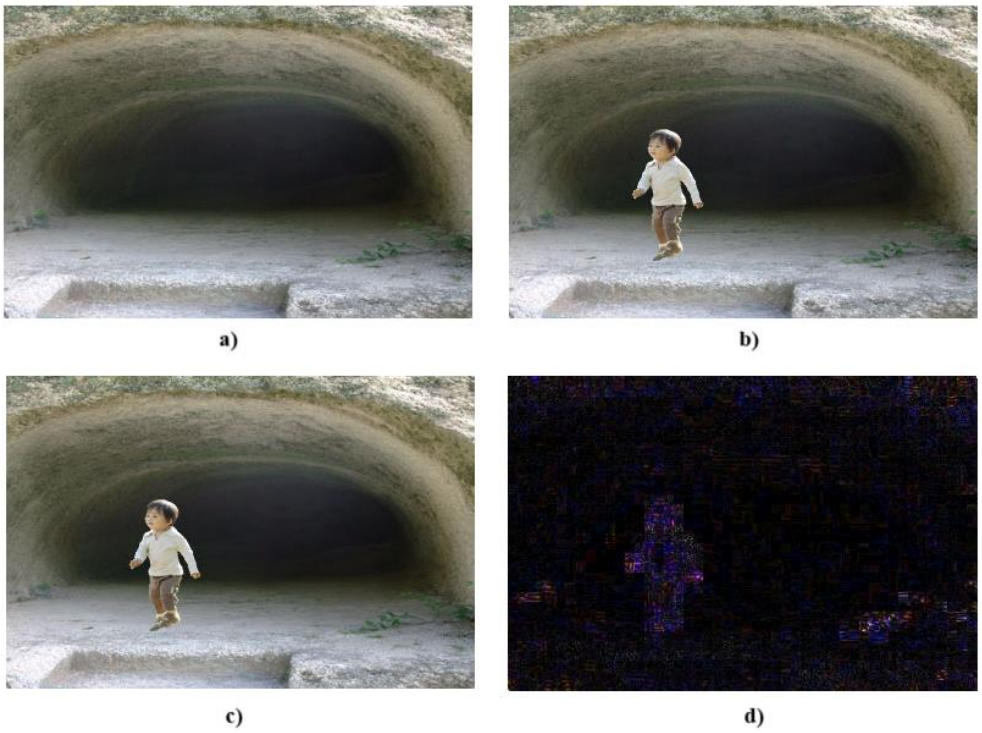
Algorithm 5.1

/* ***Model Training*** */
Input Image 'X_i' (i = 1 to n), with labels 'Y_i' (Y_i=0 if X_i is forged, else Y_i = 1)
Output: Trained Model: Prediction_Model()

/* ***Model Description*** */
Prediction_Model(image with size 128x128x3){
    First convo. layer 32 filters (size 3x3, stride size 1x1, activation: "relu")
    Second convo. layer 32 filters (size 3x3, stride size 1x1, activation: "relu")
    Max-pooling of size 2x2
    Third convo. layer 32 filters (size 3x3, stride size 1x1, activation: "relu")
    Fourth convo. layer 32 filters (size 3x3, stride size 1x1, activation: "relu")
    Max-pooling of size 2x2
    Fifth convo. layer 32 filters (size 5x5, stride size 1x1, activation: "relu")
    Sixth convo. layer 32 filters (size 5x5, stride size 1x1, activation: "relu")
    Max-pooling of size 2x2
    Seventh convo. layer 32 filters (size 5x5, stride size 1x1, activation: "relu")
    Eighth convo. layer 32 filters (size 5x5, stride size 1x1, activation: "relu")
    Max-pooling of size 2x2
    Global AVG Pooling
    Output layer (one neuron) using "sigmoid" activation}
for *epoch* = 1 to *total_epochs* do
    train_error = 0
    for i = 1 to n do
        $A_{recompressed\_i}$ = JPEG_Compression(A_i, quality)
        $A_{diff\_i} = A_i - A_{recompressed\_i}$
        $A_{reshaped\_diff\_i} = reshape(A_{diff\_i}, (128, 128, 3))$
        train_error = (Y_i - Prediction_Model($A_{reshaped\_diff\_i}$)) + *train_error*
    end for
    update_model(*train_error*, Prediction_Model(), Adam_optimizer)
end for

/* ***Prediction*** */
Input: *'Image_Input'*
Output: *'Image_Input'* labelled as authentic or tampered
$Image\_Input_{recompressed}$ = JPEG_Compression(*Input_Image*, quality)
$Image\_Input_{diff} = Image\_Input - Image\_Input_{recompressed}$

26: *Image_Input*$_{reshaped\_diff}$ = *reshape*(*Image_Input*$_{diff}$, (128, 128, 3))
27: *Predicted_label* = Prediction_Model(*Image_Input*$_{reshaped\_diff}$)
28: /* If Predicted_label = 0 then Image_Input is forged
29: /* else the Input_Image is authentic



a) displays the actual image b) displays the image with forgery
c) displays the recompressed forged image
d) displays the difference of the forged image and recompressed image

Fig. 5.3 Example Showing Processing of a Forged Image

Given a forged image A (shown in Fig. 5.3 b), we recompress it using JPEG compression and call it $A_{recompressed}$ (shown in Fig. 5.3 c). Next, we compute the difference between the provided and the recompressed one, which we refer to as A_{diff} (shown in Fig. 5.3 d). Because of the disparities between the forged region of the actual image and other regions, the forged portion gets highlighted in A_{diff}. This process of recompressing and computing the differences of both images is called ELA (Error Level Analysis). A_{diff}, an image that is referred to as a feature image for the proposed model. When trained on the feature images, the suggested model uses the artifacts left over from image tampering to determine whether altered regions exist in a given image.

CHAPTER 6

EXPERIMENTAL ANALYSIS AND RESULTS

This chapter covers how the proposed model would be trained and tested. In addition, we'll analyze and compare its performance to that of other deep-learning techniques.

6.1 Hardware Resources

The proposed model has been evaluated on a system with specific hardware specifications, “Windows 11 64-bit, Intel Core i7 processor, 16 GB of RAM, and an NVIDIA TITAN RTX graphics card with 8 GB of RAM”. The model implementation is done using ‘Python 3’ with the following libraries: Keras 3.3.3, Tensorflow 2.16.1, and Pillow 10.3.3.

6.2 Datasets Used

The datasets used for model training and testing are discussed below.

6.2.1 CASIA Datasets [29]

The CASIA datasets were produced to meet the demand for a dataset with large and realistic photos. In 2009, two datasets were released: ‘CASIA v1’ and ‘CASIA v2’. ‘CASIA v1’ has 1721 photos in total, 800 original and 921 fake. The image format is JPEG, and the image size is set at 384x256 pixels. The ‘CASIA v2’ was designed primarily to examine post-processing operations, and it includes 12,614 photos, 7,491 of which are genuine and 5,123 are tampered. It includes photos ranging in size from 240x160 to 900x600. Fig. 6.1 shows tampered images from the ‘CASIA v1’ and ‘CASIA v2’ dataset respectively. “Crop and paste operations were used on authentic images to create the tampered images using Photoshop software” [30].

Fig. 6.1 Tampered Images from CASIA Datasets

A modified CASIA dataset comprising 400 genuine and 400 modified photos was presented by Zheng et al. [31] in the year 2020. It also contains 3600 authentic and tampered images that were made by applying single content preserving manipulations to unforged and forged photographs. They used manipulations like JPEG compression, motion blur, Gaussian noise, salt & pepper noise, scaling, and rotation.

6.2.2 **MICC Datasets** [32]

The MICC team made available two datasets containing copy-move image forgeries. The MICC F-220 dataset contains 220 images, 50% of which are genuine and the other 50% are modified using copy move forgery. The photo size varies from 722x480 to 800x600. The MICC F-2000 comprises 2000 photos, 1300 of which are genuine and 700 of which are tampered. The total modified region is under 1.12% of the overall image. No attempt was taken in selecting realistic tampering, i.e. the tampered region doesn't match well with the neighbouring region. Fig. 6.2 shows a tampered image from the MICC F-2000 dataset.

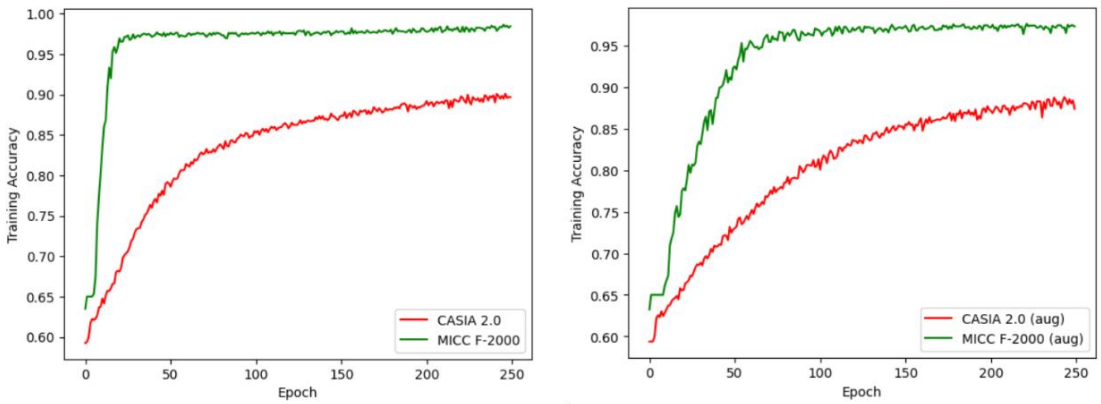
In the year 2013, Amerini et al. [33] released a new dataset with improved forgery effect and post-processing operations on some images. It has 440 authentic and 160 tampered images with sizes ranging from 800x533 to 3888x2592. The photos are available as a JPEG or a PNG. For localization of forged region, ground truth images were also provided. There is no fixed tampered region.

Fig. 6.2 Tampered Images from MICC F-2000 Dataset

For initial training and experimentation, the model has been trained and examined using the ‘MICC F-2000’ and ‘CASIA 2.0’ datasets while the final implementation of the model is done with the pictures from the ‘CASIA 2.0’ dataset after they are processed with the ELA (Error Level Analysis) algorithm. ‘MICC F-2000’ dataset lacks realistic image tampering. It is simpler to locate tampered regions because they do not align well with nearby regions. ‘CASIA 2.0’ has superior and realistic picture tampering and includes photographs from a number of genres, including animals, articles, architecture, scenery, characters, nature, plants, textures and indoor images.

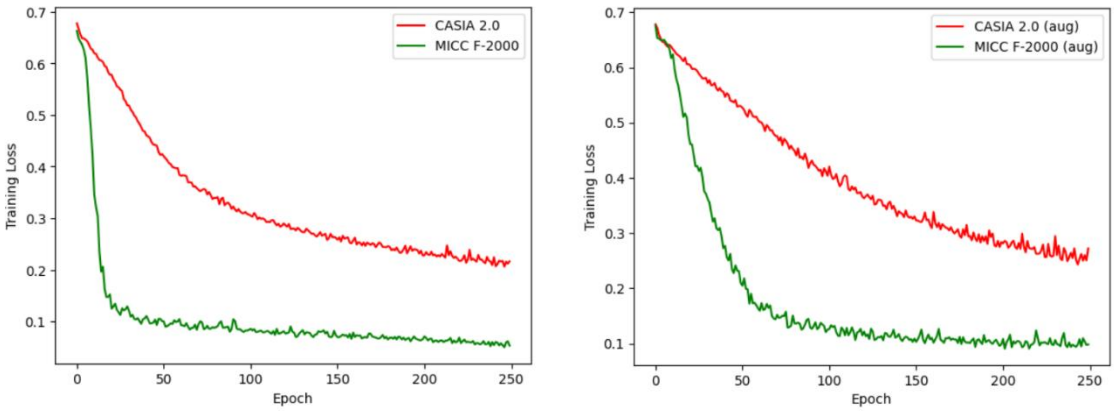
6.3 Model Training and Analysis

Four distinct networks were trained with the proposed architecture to compare their performance. Fig. 6.3 (a) depicts the training accuracies for the CAISA 2.0 and MICC F-2000 datasets, without any data augmentation. The model obtained 98.62% training accuracy on the MICC F-200 dataset and 90.00% on the CASIA 2.0 dataset. The model attained 88.78% training accuracy on the CASIA 2.0 dataset and 97.62% training accuracy on the MICC F-2000 when trained on the augmented dataset (horizontal and vertical flip). Fig. 6.3 (b) depicts the training accuracies for the CAISA 2.0 and MICC F-2000 datasets, involving data augmentation. The loss function for training the model is “binary_crossentropy” loss. Fig. 6.4 (a) displays the training loss for both datasets without data augmentation, while Fig. 6.4 (b) displays the training loss for both networks with data augmentation.



(a) Training Accuracy Non-augmented CASIA 2.0 vs MICC F-2000
(b) Training Accuracy Augmented CASIA 2.0 vs MICC F-2000

Fig. 6.3 Training Accuracies Comparison Non-Augmented and Augmented



(a) Training Loss Non-augmented CASIA 2.0 vs MICC F-2000
(b) Training Loss Augmented CASIA 2.0 vs MICC F-2000

Fig. 6.4 Training Loss Comparison Non-Augmented and Augmented

6.3.1 CASIA 2.0 (Non-Augmented) vs MICC F-2000 (Non-Augmented)

The CNN was trained on the CASIA 2.0 dataset with an “initial learning rate” of 0.0001 for the “Adam optimizer” and a “batch size” of 64. The training accuracy recorded was 90%, and the validation accuracy was 64.92%. Table 6.1 displays the confusion matrix for CASIA 2.0 Non-augmented dataset.

Table 6.1 Confusion Matrix Non-Augmented CASIA 2.0

	Predicted Authentic	**Predicted Tampered**
Actual Authentic	617	408
Actual Tampered	477	1021

The second CNN was trained using the MICC F-2000 dataset, with an "initial learning rate" of 0.0001 for the "Adam optimizer" and a "batch size" of 64. The training accuracy was 98.62, while the validation was 96.24%. Table 6.2 displays the confusion matrix for MICC F-2000 Non-augmented dataset.

Table 6.2 Confusion Matrix Non-Augmented MICC F-2000

	Predicted Authentic	**Predicted Tampered**
Actual Authentic	135	5
Actual Tampered	10	250

6.3.2 CASIA 2.0 (Augmented) vs MICC F-2000 (Augmented)

The CASIA 2.0 and MICC F-2000 datasets were augmented with horizontal and vertical flip transformations. Using the CASIA 2.0 augmented dataset, the CNN was able to attain training accuracy of 88.78% and validation accuracy of 84.97%. Similar findings were obtained while training using the MICC F-2000 augmented dataset, yielding 97.62% training accuracy and 95.99% validation accuracy. The comparison between the two is displayed in the Table 6.3.

Table 6.3 Accuracy Comparison Augmented CASIA 2.0 and MICC F-2000

Dataset	**Training Accuracy**	**Validation Accuracy**
CASIA 2.0 (Augmented)	88.78%	84.97%
MICC F-2000 (Augmented)	97.62%	95.99%

6.4 Hyperparameter Tuning

The model attained a validation accuracy of around 96% on the MICC F-2000 dataset, which is much greater than the accuracy achieved on the CASIA 2.0 dataset. The MICC datasets only contain copy-move image forgeries and the forged portion can usually be seen with the unaided eye. In contrast, the CASIA 2.0 dataset contains both copy-move and copy-paste image forgeries, as well as images contain more realistic forgeries that are easily missed by the naked eye. The MICC F-2000 dataset did not improve when processed with the ELA technique; the model attained a validation accuracy of just 93.5%, compared to approximately 96% with only data augmentation and no ELA. For this reason, additional training and tuning of the model is done with a primary focus on the CASIA 2.0 dataset with Error Level Analysis. Experimentation is done with three hyper-parameters to find out how they affect the model performance.

6.4.1 Learning Rate

To examine the impact of varying learning rates on model performance, three networks were trained. The CASIA 2.0 dataset's ELA-processed images were used to train the networks. The results for multiple learning rates are laid out in Table 6.4.

Table 6.4 Effect of Learning Rate on CASIA 2.0 ELA

Learning Rate	Training Accuracy	Validation Accuracy
0.00001	90.35%	91.00%
0.0001	94.42%	93.22%
0.001	99.97%	91.20%

6.4.2 Batch Size

The Batch size is considered one of the key hyperparameters that affect the model performance. The model performance for various batch sizes is displayed in the Table 6.5.

Table 6.5 Effect of Batch Size on CASIA 2.0 ELA

Batch Size	Training Accuracy	Validation Accuracy
32	96.29%	93.65%
64	94.42%	93.22%
128	93.58%	92.74%

6.4.3 Number of Filters

The number of filters simply increases the count of parameters in the CNN layer. It provides for the extraction of a differing feature than other filters in the same layer. The Table 6.6 shows the model performance for different numbers of filters.

Table 6.6 Effect of Number of Filters on CASIA 2.0 ELA

Number of Filters	Training Accuracy	Validation Accuracy
32	94.42%	93.22%
64	98.16%	92.39%

6.5 CNN Training with ELA

The proposed model flow chart is displayed in Fig. 5.1. The first step is the pre-processing of the dataset while the second is ELA conversion followed by model training which is discussed below.

- **Pre-processing**

During the pre-processing stage, images are normalised. Ensuring a consistent data distribution across every image is the primary objective of normalisation. The dataset is resized to 128x128 pixels for consistency.

- **ELA Conversion**

Images are converted to ELA by reading an image from the dataset and compressing it again, after which the difference in the pixel values of the actual (pre-compressed) and recompressed photos are calculated. The brightness of the image obtained after

this process is enhanced. Each pixel value is divided by 255 in order to normalize the CNN training. The next step is to classify each image, where 1 denotes an authentic image and 0 denotes forged image.

- **Model Training**

The convolutional layer represents the features and acts as a feature extractor. Furthermore, the pooling layer decreases the convolution layer's output map while decreasing the chances of overfitting. The maximum pooling layer has a pool size of 2x2. Multiple convolutional layers are used with different kernel sizes (3x3 and 5x5) and the number of filters is set to 32 for each convolutional layer.

Following the ELA conversion of every image in the CASIA 2.0 dataset, the dataset was randomly partitioned into two groups ("training and validation sets"). The split was in the 80:20 ratio. The model was trained using ELA-converted images from the CASIA 2.0 dataset, with varied image quality levels. The model was trained using "binary cross-entropy" loss and "Adam optimizer" for 100 epochs to optimize the network. A block size of 64 with an "initial learning rate" of 0.0001 was employed.

Using various levels of image quality, the model was found to achieve varying levels of accuracy for ELA processed images. Thus, experiments were carried out using ELA at various image quality levels. Fig. 6.5 displays an ELA image at different levels of quality for an input image.

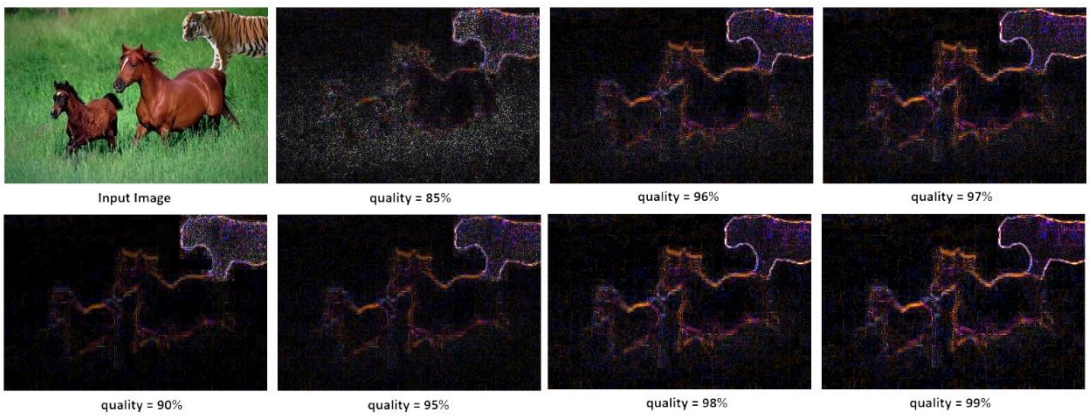


Fig. 6.5 ELA Image at Different Levels of Image Quality

6.6 Results Obtained

Table 6.7 displays the accuracies of training and validation for different image quality levels. The model worked best at an image quality level of 99%, resulting in a validation accuracy of 93.38%. Table 6.8 shows the confusion matrix. Fig. 6.6 displays the model's training and validation accuracy, as well as the training and validation loss. Finally, Table 6.9 displays the accuracy results of the model for different approaches. Table 7 displays the comparison of the model accuracy with deep-learning techniques proposed by other authors.

Table 6.7 Effect of Image Quality on Accuracy

Image Quality Level	Training Accuracy	Validation Accuracy
85%	91.71%	89.10%
90%	92.80%	88.78%
95%	93.86%	93.14%
96%	94.62%	90.44%
97%	94.67%	92.82%
98%	94.42%	93.22%
99%	**95.10%**	**93.38%**
100%	94.11%	93.14%

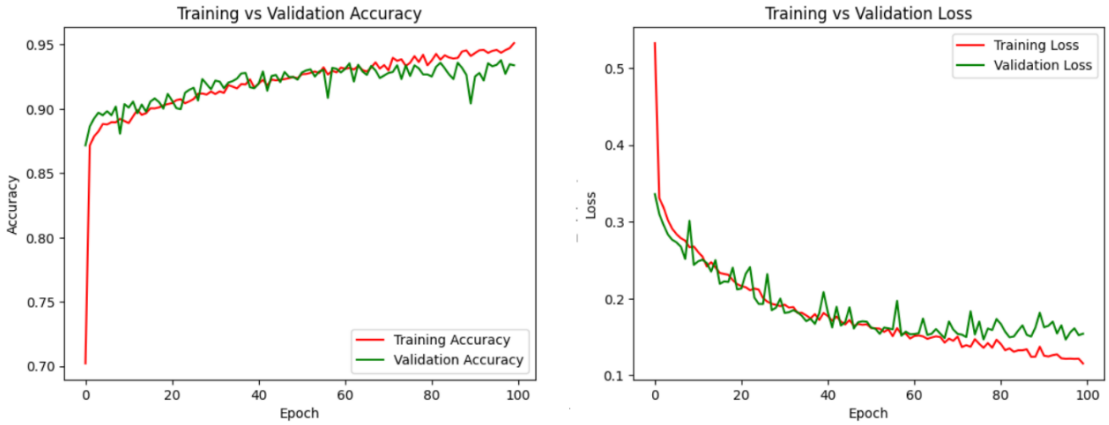


Fig. 6.6 Model Accuracy and Loss

Table 6.8 Confusion Matrix CASIA 2.0 ELA

	Predicted Authentic	**Predicted Tampered**
Actual Authentic	984	41
Actual Tampered	126	1372

Table 6.9 Model Accuracy Comparison on CASIA 2.0

Model	**Training Accuracy**	**Validation Accuracy**
CNN	90%	64.92%
CNN + Dataset Augmentation	88.78%	84.97%
CNN + ELA	**95.10%**	**93.38%**

Table 7 Accuracy Comparison with Deep-Learning Techniques on CASIA 2.0

Technique	**Accuracy**
Mantra-Net	56.14%
CAT-Net	87.29%
VGG16 + ELA	88.46%
CNN + Recompression	92.23%
Mini-Net	**93.75%**
Proposed Technique	93.38%

CHAPTER 7

CONCLUSION, FUTURE SCOPE AND SOCIAL IMPACT

In this study, we looked into how a CNN might be used to identify photos that have been manipulated. More specifically, we used ELA (Error Level Analysis) in conjunction with a CNN network to identify image forgeries that involved copy-move and copy-paste (image splicing). Using the CASIA 2.0 and MICC F-2000 datasets, we ran several experiments. While we were able to achieve a validation accuracy of about 96.24% on the MICC F-2000 dataset without making any changes to the dataset, we only ended up achieving 93.38% of validation accuracy on the 'CASIA 2.0 dataset' by using the proposed CNN model combined with the ELA technique. In contrast to MICC F-2000, which only includes copy-move image forgery and lacks realistic image forging, the 'CASIA 2.0 dataset' includes both copy-move and image-splicing picture forgeries. These results support our theory that the more challenging the data, the lower the classification performance.

Furthermore, our research discovered that even when performed by specialists, image alterations may be identified with greater than 90% accuracy. However, there is still a lot of work to be done in this field of "image forgery detection"; the final objective is to create a neural network that can accurately recognize forgeries based on the complexity of the imagery. There is still room for development in the CNN+ELA training model, including expanding the dataset and using powerful computer devices.

The social impact of an image tampering detection application is significant, reflecting its crucial role in modern society. It enhances trust in digital media by verifying image authenticity and combating misinformation. These applications support the integrity of news and social media, protect against fraud, and contribute to the credibility of visual content across various platforms.

REFERENCES

[1] W. D. Ferreira, C. B. Ferreira, G. d. C. Júnior and F. Soares, "A review of digital image forensics," *Computers & Electrical Engineering,* vol. 85, no. 106685, 2020.

[2] S. Pradhan, D. U. Chauhan and D. S. Chauhan, "A Review on Digital Image Forgery Detection," *2022 2nd International Conference on Innovative Practices in Technology and Management (ICIPTM),* pp. 1-6, 2022.

[3] R. Rani, A. Kumar and A. Rai, "A Brief Review on Existing Techniques for Detecting Digital Image Forgery," *2021 Sixth International Conference on Image Information Processing (ICIIP),,* pp. 533-538, 2021.

[4] B. S. Kumar, R. Cristin, K. Karthick and T. Daniya, "Study of Shadow and Reflection based Image Forgery Detection," *2019 International Conference on Computer Communication and Informatics (ICCCI),* pp. 1-5, 2019.

[5] D. P. S. Gayathri K S, "AN OVERVIEW OF COPY MOVE FORGERY," *Computer Science & Engineering: An International Journal (CSEIJ),* vol. 12, 2022.

[6] M. H. Alkawaz, G. Sulong, T. Saba and A. Rehman, "Detection of copy-move image forgery based on discrete cosine transform," *Neural Comput & Applic,* vol. 30, p. 183–192, 2018.

[7] S. Dua, J. Singh and H. Parthasarathy, "Image forgery detection based on statistical features of block DCT coefficients," *Procedia Computer Science,* vol. 171, no. ISSN 1877-0509, pp. 369-378, 2020.

[8] N. K. Rathore, N. K. Jain, P. K. Shukla, U. Rawat and R. Dubey, "Image Forgery Detection Using Singular Value Decomposition," *National Academy Science Letters,* vol. 44, no. 4, p. 331–338, 2021.

[9] D. Tralic, I. Zupancic, S. Grgic and M. Grgic, "CoMoFoD -New Database for Copy-Move Forgery Detection," *Proceedings Elmar - International Symposium Electronics in Marine. ,* 2013.

[10] E. A. A. VEGA, E. G. FERNÁNDEZ, A. L. S. OROZCO and A. L. J. G. VILLALBA, "Passive Image Forgery Detection Based on the Demosaicing Algorithm and JPEG Compression," *IEEE Access,* vol. 8, pp. 11815-11823, 2020.

[11] A. Hegazi, A. Taha and M. M. Selim, "An improved copy-move forgery detection based on density-based clustering and guaranteed outlier removal," *Journal of King Saud University –Computer and Information Sciences,* vol. 33, pp. 1055-1063, 2021.

[12] R. Agarwal, D. Khudaniya, A. Gupta and K. Grover, "Image Forgery Detection and Deep Learning Techniques: A Review," *2020 4th International Conference on Intelligent Computing and Control Systems (ICICCS),* pp. 1096-1100, 2020.

[13] K. M. Hosny, A. M. Mortda, M. M. Fouda and N. A. Lashin, "An Efficient CNN Model to Detect Copy-Move Image Forgery," *IEEE Access,* vol. 10, pp. 48622-48632, 2022.

[14] S. S. Ali, I. I. Ganapathi, N.-S. Vu, S. D. Ali, N. Saxena and N. Werghi, "Image Forgery Detection Using Deep Learning by Recompressing Images," *Electronics,* vol. 11, no. 3, p. 403, 2022.

[15] I. B. K. Sudiatmika, F. Rahman, T. Trisno and S. Suyoto, "Image forgery detection using error level analysis and deep learning," *TELKOMNIKA,* vol. 17, no. 2, pp. 653-659, 2019.

[16] S. Koul, . M. Kumar, S. S. Khurana, F. Mushtaq and K. Kumar, "An efficient approach for copy-move image forgery detection using convolution neural network," *Multimedia Tools and Applications,* vol. 81, p. 11259–11277, 2022.

[17] A. Doegar, M. Dutta and K. Gaurav, "CNN Based Image Forgery Detection Using Pre-trained AlexNet Model," *International Journal of Computational Intelligence & IoT,* vol. 2, no. 1, 2019.

[18] M. A. Elaskily, H. A. Elnemr, A. Sedik, M. M. Dessouky, G. M. E. Banby, O. A. Elshakankiry, A. A. M. Khalaf, &. H. K. Aslan, O. S. Faragallah and F. E. A. El-Samie, "A novel deep learning framework for copy-moveforgery detection in images," *Multimedia Tools and Applications,* vol. 79, p. 19167–19192, 2020.

[19] J. Ouyang, Y. Liu and M. Liao, "Copy-Move Forgery Detection Based on Deep," *10th International Congress on Image and Signal Processing, BioMedical Engineering and Informatics,* 2017.

[20] M. Zanardelli, F. Guerrini, R. Leonardi and N. Adami, "Image forgery detection: a survey of recent," *Multimedia Tools and Applications,* vol. 82, p. 17521–17566, 2023.

[21] M. T. H. Majumder and A. B. M. A. A. Islam, "A Tale of a Deep Learning Approach to Image Forgery Detection," *5th International Conference on Networking, Systems and Security,* pp. 1-9, 2018.

[22] M.-J. Kwon, I.-J. Yu, S.-H. Nam and . H.-K. Lee, "CAT-Net: Compression Artifact Tracing Network for Detection and Localization," *IEEE Winter Conference on Applications of Computer Vision (WACV),* pp. 375-384, 2021.

[23] S. Tyagi and D. Yadav, "MiniNet: a concise CNN for image forgery detection," *Evolving Systems,* vol. 14, p. 545–556, 2023.

[24] Y. Wu, W. AbdAlmageed and P. Natarajan, "ManTra-Net: Manipulation Tracing Network For Detection And Localization of," *IEEE/CVF Conference on Computer Vision and Pattern Recognition,* pp. 9535-9544, 2019.

[25] G. Kaur, N. Singh and M. Kumar, "Image forgery techniques: a review," *Artificial Intelligence Review,* vol. 56, 2022.

[26] Y. Lai, T. Huang, J. Lin and H. Lu, "An improved block-based matching algorithm of copy-move forgery detection," *Multimedia Tools and Applications,* vol. 77, 2018.

[27] L. Alamro and N. Yusoff, "Copy-Move Forgery Detection using IntegratedDWT and SURF," *Journal of Telecommunication, Electronic and Computer Engineering (JTEC),* vol. 9, pp. 67-71, 2017.

[28] S. Alzahir and R. Hammad, "Image forgery detection using image similarity," *Multimedia Tools and Applications,* vol. 79, 2020.

[29] J. Dong, W. Wang and T. Tan, “CASIA Image Tampering Detection Evaluation Database,” *IEEE China Summit and International Conference on Signal and Information Processing,* pp. 422-426, 2013.

[30] S. Tyagi and D. Yadav, “A detailed analysis of image and video forgery detection techniques,” *The Visual Computer,* 2022.

[31] Y. C. C.-H. C. Yue Zheng, “A PUF-Based Data-Device Hash for Tampered Image Detection and Source Camera Identification,” *IEEE Transactions on Information Forensics and Security,* vol. 15, pp. 620-634, 2020.

[32] I. Amerini, L. Ballan, R. Caldelli, A. D. Bimbo and G. Serra, “A SIFT-Based Forensic Method for Copy–Move Attack Detection and Transformation Recovery,” *IEEE Transactions on Information Forensics and Security,* vol. 6, pp. 1099 - 1110, 2011.

[33] I. Amerini, L. Ballan, R. Caldelli, A. D. Bimbo, L. D. Tongo and G. Serra, “Copy-move forgery detection and localization by means of robust clustering with J-Linkage,” *Signal Processing: Image Communication,* vol. 28, no. 6, pp. 659-669, 2013.

LIST OF PUBLICATIONS

1. Shujaa Ahmad and Bindu Verma, “Forgery Detection in Digital Images Using CNN and ELA” accepted in 1st IEEE International Conference on Advances in Computing, Communication and Networking- ICAC2N-24, (16th - 17th December 2024), Greater Noida, India.
2. Shujaa Ahmad and Bindu Verma, “A Review of Deep Learning Techniques for Image Forgery Detection” accepted in 1st IEEE International Conference on Advances in Computing, Communication and Networking- ICAC2N-24, (16th - 17th December 2024), Greater Noida, India.