DIGITAL IMAGE FORGERY DETECTION USING DEEP LEARNING

Kaki Saikumar[1] , Dr.GNV Vibhav Reddy[2]

[1]PG Scholar , Department of Computer Science and Engineering , Sreedattha Institute Of Engineering And Science Sheeriguda, Ibrahimpatnam Hyderabad , Telangana,India.

[2]Associate Professor, Department of Computer Science and Engineering ,Sreedattha Institute Of Engineering And Science Sheeriguda, Ibrahimpatnam Hyderabad , Telangana,India.

ABSTRACT:

The identification and apprehension of image forgeries represents a pivotal domain of study within imaging forensics. Thanks to the accessibility of state-of-the-art technology, powerful image editing tools, and software packages, photos can be effortlessly altered or manipulated. There are countless images all around us; however, some of them may be phoney. While there are many different kinds of picture forgeries, copy-move forgery is a significant technique that has gained popularity recently. In copy movement forgeries, a portion of an image is copied and pasted at a different spot inside the same image, reflecting disparate interpretations of the image. Three types of solutions—block level, key point based, and deep learning based—have been employed to address this issue. There are several block-level approaches, including local binary patterns, stationary wavelet transforms, scale invariant feature transforms, speedup robust feature transforms, and deep learning-based methods like mobile nets and The debut nets. In contrast to block level and key point-based techniques, deep learning models have recently attracted the attention of researchers. Deep learning models have a way to autonomously acquire and extract features from the training dataset. We have carefully examined and researched the many methods found in the literature to enhance the efficiency of image forgery detection and classification. We have developed improvised deep neural network models to save computational expenses and obtain higher accuracy. In this thesis, we introduced a streamlined and tailored Convolutional Neural Network (CNN) model with its objective of achieving enhanced classification accuracy for manipulated images. This work serves as an initial step towards this goal. Moreover, we expand upon this methodology and address its constraints by introducing a novel hybrid approach, namely VI-NET, which is an ensemble deep neural network. This algorithm has been able to achieve a maximum accuracy of 98%. In addition, we conducted a comparison of our outcomes using supplementary metrics, including precision, recall, false positive rate (FPR), and false negative rate (FNR).

Key words: Convolutional Neural Network (CNN), false positive rate (FPR), and false negative rate (FNR), Deep Learning.

I. INTRODUCTION

With the progress of technology, humans are surrounded by a multitude of images. It might refer to photographs that are posted on social media or shared through other digital channels with friends and family. Images have a significant influence on how people perceive themselves in the modern world. Can we trust the authenticity of the visuals that we are receiving or sharing? Do we have certainty that we are transmitting accurate information to the global internet community? We lack confidence in providing answers to these questions. Also, if these pictures are used as proof in the court system, it could be unfair to the people who are being tried. In the present era, where even our mobile gadgets possess great ability to alter the

appearance of photos, there is another reason to question the images. It is sometimes difficult to discern whether there has been any tempering because of the way the image's texture might be altered. Therefore, image forging (IF) is the application of a collection of techniques to genuine photographs as a way to conceal their true interpretations and meanings. [1].

The manipulated digital images sourced from social media, newspapers, healthcare documents, and websites lose their authentic meaning [2]. In 2008, Iran released manipulated photographs that depicted four missile launches, thereby disseminating inaccurate data regarding the nation's military prowess [3]. Forged healthcare and diagnostic photos transmitted online can result in faulty diagnoses [4]. Consequently, there is an evident need for efficient techniques to identify image forgeries.

The photos are susceptible to many forms of counterfeiting. Copy-move forgery is the most prevalent form of forgery. Copy-move forgery involves duplicating an image patch and pasting it elsewhere within the same image. There are arguments that several deep learning approaches [7] are superior to non-AI-based approaches like block-level and key point-based approaches [8] for the detection of image forgeries. Newer AI systems, like deep learning, can learn model features from training photos, while earlier methods, like artificial intelligence, rely on created features [9]. A two-stage deep neural network with an encoder-decoder architecture achieved 98% reliable classification [10], while a generative adversarial network (GAN) with a convolution neural network (CNN) achieved 95% classification and localization accuracy for fabricated images [11]. Deep learning methods acquire non-linear characteristics and generate more generalised models in both healthcare and non-healthcare domains.

This chapter provides a comprehensive overview of picture forgeries, including their different kinds, existing gaps in the literature, research objectives, and contributions. The issue of potential bias in a deep learning solution designed to detect image forgery has also been previously deliberated.

II. LITERATURE SURVEY

The internet has evolved into an important tool in everyday life for obtaining information and extracting pertinent information from a website. Digitized content is readily accessible and easily disseminated, whether through legal or unlawful means. Users, for example, copy, resize, and/or manipulate photos available on the Internet before publishing them as their own, resulting in forged images. This chapter examines conventional algorithms used for detecting forgeries.

A. DIVERSE TECHNIQUES EMPLOYED FOR DETECTING COUNTERFEIT IMAGES

First, two photographs that need to be identified are extracted from the database, and after that, feature vectors for the two pictures are created. The features are compared, and the photographs are classed as duplicate pairs or not based on the degree of resemblance. Figure 2.1 illustrates the primary steps involved in identifying duplicate images.

- Pair Images
- Compare Feature Vectors
- Estimate similarity vector feature matching
- More Similar
  - Forged Image
- Less Similar
  - Non-Forged Image

B. UTILIZATION OF DEEP LEARNING TECHNIQUES IN IMAGE FORGERY DETECTION

Image manipulation techniques, such as shearing, rotating, and scaling, were used to assess the quality of the applied photos [12]. Experimental results show that the suggested transfer learning technique successfully accelerates CNN model convergence without increasing image quality or revealing image manipulation. The convolutional neural network (CNN) is an advanced deep learning method that efficiently recovers high-level features from

a large collection of tagged images. In document image processing, ink analysis enables the detection of ink age and forgery, as well as the identification of the pen or writer. The paper titled "Ink spectral information in hyperspectral document photos" was authored by Khan et al in 2018.

[13] showed an instances of copy-move forgeries have been identified. The initial attention mechanism in the generator detects the exact location of the copy-move operation, whereas the subsequent attention mechanism analyses the co-occurrence of patches. The affinity matrix is utilised to generate attention maps that integrate both co-occurrence and location-aware attributes. The discriminator network enhances the precision of the localization findings.

In their study, [14] introduced two methods to identify image tampering by combining re-sampling features with deep learning. These algorithms aim to accurately pinpoint any modifications made to the images. Initially, The Random transform of resampling features is calculated for overlapping picture patches. A heat map is generated by employing deep learning classifiers, and a Gaussian conditional random field model is acquired to identify tampered zones. To detect forgeries, the Random Walker segmentation approach is applied. Using overlapping data, the second technique computes resampling characteristics. Image patches are used to classify and localise using a Long Short-Term Memory (LSTM) network. The detection and localization capabilities of both approaches for identifying and finding digital picture frauds are compared by the authors.
A deep learning-based image fraud detection system employs a convolutional neural network (CNN) to learn hierarchical representations from RGB colour photos. [15] conducted the study. Initializing the first layer of the network's weights with the basic high-pass filter set used for calculating residual maps in the spatial rich model (SRM) works as a regularize to effectively block the effect of image contents and capture the subtle artefacts that were introduced by the tampering operations, rather than a random strategy. Using the pre-trained CNN as a patch descriptor, dense features are extracted from the test images. The final discriminative features for SVM classification are then produced using a feature 16 fusion procedure.

In their study, [16] provided a comprehensive compilation of deep learning methods that can be employed for analysing huge amounts of data as a means to detect forgeries. The researchers investigated a range of prevalent deep learning architectures and their pragmatic implementations in this investigation. An analysis is conducted on four topics: autoencoder, convolutional neural network, deep belief network, and constrained Boltzmann machine.

In their 2018 paper, [17] introduced BusterNet, a novel deep neural network structure designed specifically for detecting instances of image copy-move forgeries. The system features a dual-branch design, which is thereafter succeeded by a tri-branch architecture. A module for fusion of two branches employs visual artefacts to identify potential manipulation places and visual similarities to identify copy-move regions. [18] present a colour filter array (CFA)-based convolutional neural network (CNN) technique for precisely localising tampered regions. The CFA interpolation approach incorporates pixel correlation and ensures consistency. The proposed CNN method effectively differentiates between traces generated by copy-move forgeries and different post-processing activities.

III. CNN BASED FORGERY CLASSIFICATION

The purpose of the study outlined in this chapter is to evaluate the efficacy of a tailored Convolutional Neural Network (CNN) model in accurately categorizing manipulated

photographs. The categories of photos encompass copy move forgeries as well as post-processed photographs that involve scaling, rotation, and other degrees of alteration. Patches undergo compression. To address this problem, we have created a novel customised lightweight CNN model that can achieve approximately 95% accuracy for both standalone datasets and the integration of several datasets.

A. TAILORED CONVOLUTIONAL NEURAL NETWORK (CNN) STRUCTURE

Upon familiarizing oneself with the pertinent literature, it becomes evident that despite the existence of several techniques, detecting copy move fraud in digital photographs continues to be a difficult task. However, none of them has resistance against the countless possible variants and are lacking in precise depiction. In locations characterized by minimal or absent texture, a technique focused on critical points would not yield satisfactory results. The block level techniques are not efficient in handling images that have undergone geometric modifications. Thus, the techniques that have been used up to now are not perspective-change-invariant nor mirror reflection-invariant. Only a limited number of studies have employed deep learning to address this problem, which has been identified as a significant requirement. Thus far, the localization process has received minimal attention. This CNN architecture uses deep neural networks to classify forged images in various manipulations, including copy move forging.

The idea of a convolutional neural network is used in this work as a first step towards sorting the fake and real photos into different groups. A model with pooling and convolution layers offers higher accuracy compared to other solutions. While there exists a wide array of beforehand trained models that can be employed for some point initialization, it seems that these models impose a significant computational burden and fail to deliver satisfactory accuracy across all 26 datasets. Various layer combinations and parameter settings were tested and refined to achieve an accurate prediction model.

Our primary goal is to simplify the model and create layers that require minimal computer power.

The input image is initially converted to a grayscale image to reduce the number of parameters. Figure 5, which shows the recommended architecture, illustrates this. To reduce the complexity of the parameters and training time, all input images are resized to dimensions of 32 by 32 pixels.

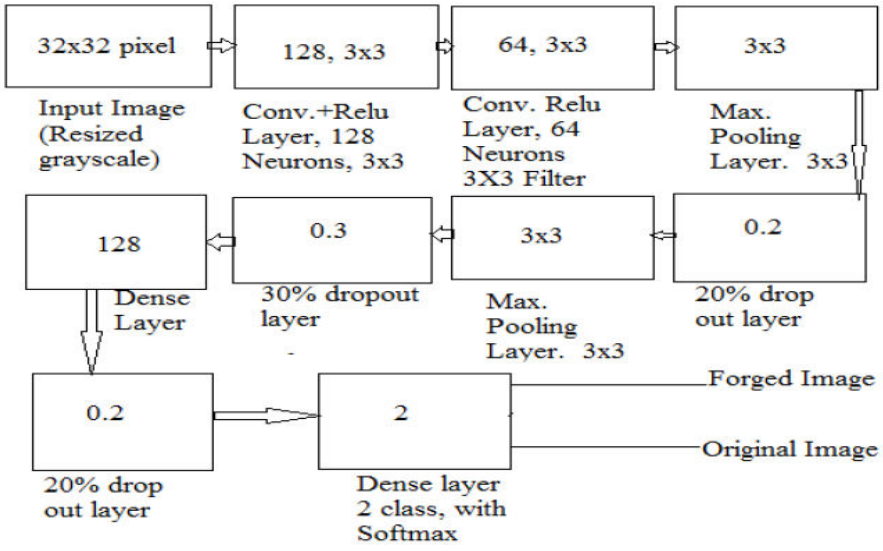

Figure 1. CNN based Forgery Classification

After the input layer comes the convolution layer, which has 128 neurons and uses a 3x3 filter to execute convolution operations. Subsequently, there is an additional convolution layer of 64 neurons utilising the relu activation function. Figure 1 displays the suggested design for the Convolutional Neural Network (CNN).

The accuracy of the method under discussion is then compared to that of alternative techniques, which essentially entail the manual extraction of characteristics. This section will focus on constructing a model that utilises a convolutional neural network to automatically extract local feature vectors. This will enable precise identification of photos as either counterfeit or authentic.

To achieve this objective, various datasets such as COMOFOD, DIID, and the Image Manipulation dataset, among others, have undergone experimentation.

The primary experiment contributes to originality by testing the precision of the outcome when applied to many datasets. Experiments are conducted and results analysed in all of these circumstances.

IV. FINDINGS AND ANALYSIS

Approximately 6050 cleaned images from the COMOFOD dataset are being utilised to train the suggested model. A train test with an 80/20 split is conducted in order to validate the outcome. Considering this, the confusion matrix illustrating the results of testing 1210 images on a trained model is provided below. The model learned from these pictures about 50 times, until it got to an accuracy of 93.2 percent. To test the model's resilience to different types of invariance, weights from the COMOFOD [118] dataset were applied to additional datasets. The DIID [11] dataset comprises images that will be utilised for additional training.

These images consist of a variety of examples, including basic copy-move, rotated images, and scaled images. This signifies that all of the photos have been effectively merged to produce a resilient model. The benchmark dataset, which was used to evaluate the model's resilience and make sure it generates reliable results for the benchmark images, is also included in the training image collection in Figure 2. Once these photos are integrated, 8424 of them are used for training, and 2106 of them are used for validation. This yields a 94.77% accuracy rate, which is far higher and more reliable than the block level and key point based approaches.

Original Label	COMOFOD Dataset Test Result		COMOFOD+DIID Dataset		COMOFOD+DIID+ Image Manipulation Dataset	
0	228	87	438	59	1229	38
1	0	895	24	971	72	767
Predicted label→	0	1	0	1	0	1

Figure 2. Confusion Matrix

During the training, each original image is labelled as "0" and each fraudulent image is labelled as "1." The model combines gradient descent and the Adam optimizer to achieve optimal performance. To mitigate the risk of overfitting, the 'Adam' optimizer is being employed in this context. Accuracy at the picture level was also assessed and compared in this work. To validate results, evaluate the accuracy with which images are identified as forgeries or originals. Given this information, the metrics of precision and recall are considered. The degree to which your result matches the actual result is known as precision, and the degree to which it accurately reflects overall performance across all of the data is known as recall in Figure 3 and 4.

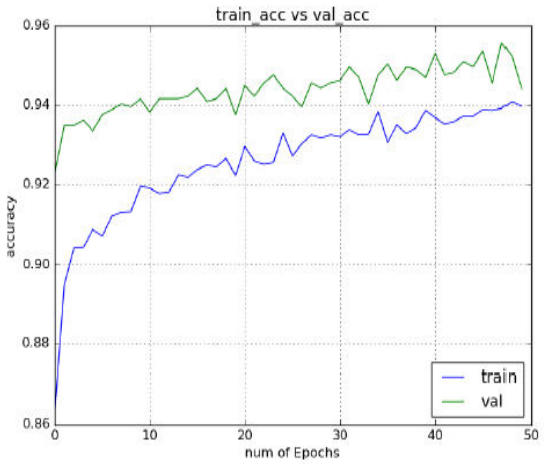

Figure 3. Accuracy based on Dataset

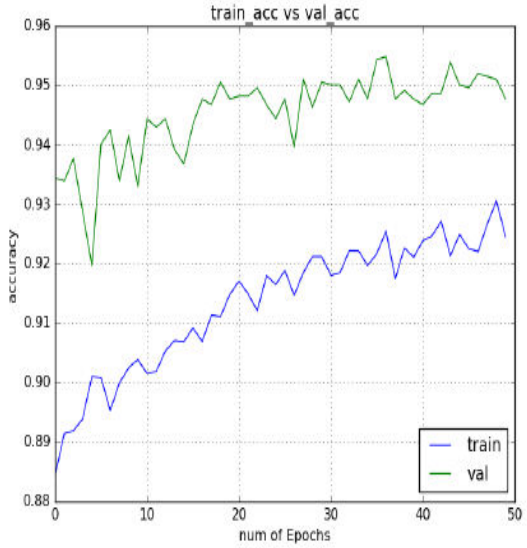

Figure 4. Accuracy based Image Manipulation

```
import os
import sys
from tempfile import NamedTemporaryFile
from urllib.request import urlopen
from urllib.parse import unquote, urlparse
from urllib.error import HTTPError
from zipfile import ZipFile
import tarfile
import shutil
```

The above are the packages being used for the process of classifying and detecting the images for the procedure of forgery.

```
CHUNK_SIZE = 40960
DATA_SOURCE_MAPPING = 'casia-dataset:https%3A%2F%2Fstorage.googleapis.com%2Fkaggle-data-sets
```

The Dataset is identified and its path is defined and located based on 40960 chunks of data size.

```
#import necessary libraries
import numpy as np
import matplotlib.pyplot as plt
%matplotlib inline
np.random.seed(2)
from sklearn.model_selection import train_test_split
from sklearn.metrics import confusion_matrix
```

To specify the learning process, random seed function is initiated and it adds training the data and determine the confusion matrix.

CNN based layer is added with categorical and sequential factor. Then various layer specified as

```
from keras.utils.np_utils import to_categorical
from keras.models import Sequential
from keras.layers import Dense, Flatten, Conv2D, MaxPool2D, Dropout
from keras.optimizers import Adam
from keras.preprocessing.image import ImageDataGenerator
```

Dense, flatten, conv2D, max pool and dropout later for the process of CNN Model.

```
image = Image.open(path).convert('RGB')
image.save(temp_filename, 'JPEG', quality = quality)
temp_image = Image.open(temp_filename)

ela_image = ImageChops.difference(image, temp_image)

extrema = ela_image.getextrema()
max_diff = max([ex[1] for ex in extrema])
if max_diff == 0:
    max_diff = 1
scale = 255.0 / max_diff

ela_image = ImageEnhance.Brightness(ela_image).enhance(scale)
```

It makes to fetch the real images and convert them into ela_image.

IV. CONLCUSION

We need to study whether there is a method that can endure repeated transformations and maintain excellent accuracy, as evidenced by the literature and data. Even while the literature has demonstrated a considerable degree of accuracy, this precision comes with a cost: it is susceptible to specific manipulations. Nevertheless, we have

noted a satisfactory level of accuracy. Using The suggested adjusting technique can be categorised as perfect given the difficulties presented by sensitivity to reflection, compression of images, rotation, and scaling as a whole. The CNN technique used in this paper achieved an accuracy of around 95% using a model built from scratch. This model possesses the capability to accurately recognise photographs, even in cases when the images have undergone various treatments and post-processing techniques, such as blurring, altering the lighting, and diminishing the colour intensity. Further accuracy increase could be achieved as a future project using a CNN model that encompasses the whole process from start to finish. We have endeavoured to enhance the precision of categorization by using a novel hybrid Convolutional Neural Network (CNN) model.

REFERENCES

[1]. Alberry Hesham, A, Abdelfatah A Hegazy & Gouda I Salama 2018, 'A fast SIFT based method for copy move forgery detection', Future Computing and Informatics Journal, vol. 3.2, pp. 159-165.

[2]. Amerini Irene, et al. 2011, 'A SIFT-based forensic method for copy-move attack detection and transformation recovery', IEEE Transactions on Information Forensics and Security, vol. 6.3, pp. 1099-1110.

[3]. Ardizzone Edoardo, Alessandro Bruno & Giuseppe Mazzola 2015, 'Copy–move forgery detection by matching triangles of key points', IEEE Transactions on Information Forensics and Security, vol. 10.10, pp. 2084-2094.

[4]. Babu, SBG Tilak & Ch Srinivasa Rao 2021, 'An optimized technique for copy–move forgery localization using statistical features', ICT Express.

[5]. Baohua Zhang, Lu Xiaoqi & Jia Weitao 2013, 'A multi-focus image fusion algorithm based on an improved dual-channel PCNN in the NSCT domain', Optik, vol. 124.20, pp. 4104-4109.

[6]. Bappy Jawadul, H et al. 2019, 'Hybrid LSTM and encoder–decoder architecture for detection of image forgeries', IEEE Transactions on Image Processing, vol. 28.7, pp. 3286-3300.

[7]. Barad Zankhana, J & Mukesh M Goswami 2020, 'Image forgery detection using deep learning: A survey', 6th International Conference on Advanced Computing and Communication Systems (ICACCS), IEEE.

[8]. Barni Mauro, Quoc-Tin Phan & Benedetta Tondi 2019, 'Copy move source-target disambiguation through multi-branch CNNs', arXiv preprint arXiv:1912.12640.

[9]. Bayram Sevinc, Husrev Taha Sencar & Nasir Memon 2008, 'A survey of copy-move forgery detection techniques', IEEE Western New York Image Processing Workshop, IEEE.

[10]. Bi Xiuli & Chi-Man Pun 2017, 'Fast reflective offset-guided searching method for copy-move forgery detection', Information Sciences, vol. 418, pp. 531-545.

[11]. Bi Xiuli & Chi-Man Pun 2018, 'Fast copy-move forgery detection using local bidirectional coherency error refinement', Pattern Recognition, vol. 81, pp. 161-175.

[12]. Christlein Vincent, et al. 2012, 'An evaluation of popular copy-move forgery detection approaches', IEEE Transactions on Information Forensics and Security, vol. 7.6, pp. 1841-1854.

[13]. Costanzo Andrea, et al. 2014, 'Forensic analysis of SIFT keypoint removal and injection', IEEE Transactions on Information Forensics and Security, vol. 9.9, pp. 1450-1464.

[14]. Cozzolino Davide, Giovanni Poggi & Luisa Verdoliva 2015, 'Efficient dense-field copy-move forgery detection', IEEE Transactions on Information Forensics and Security, vol. 10.11, pp. 2284-2297.

[15]. Cui Kebin, et al. 2014, 'An improved unit-linking PCNN for segmentation of infrared insulator image', Applied Mathematics and Information Sciences, Vol. 8.6, p. 2997.

[16]. Davarzani Reza, et al. 2013, 'Copy-move forgery detection using multiresolution local binary patterns', Forensic Science International, vol. 231.1-3, pp. 61-72.

[17]. Diallo Boubacar et al. 2020, 'Robust forgery detection for compressed images using CNN supervision', Forensic Science International: Reports, vol. 2, p. 100112.
[18]. Dua Shilpa, Jyotsna Singh & Harish Parthasarathy 2020, 'Detection and localization of forgery using statistics of DCT and Fourier components', Signal Processing: Image Communication, vol. 82, p. 115778.