Integrity Analysis and Detection of Digital Forensic Evidences

by

Eshrak Ahmed
19301003

Department of Computer Science and Engineering
School of Data and Sciences
Brac University
Summer 2023

Declaration

I, the author of this thesis guarantee that this is my original work and it has not been submitted for any academic purposes. I certify that this work has completely done by me and reflects my own ideas, opinions and hard work. No matter has been purposefully hidden and all the sources that I have used have been cited correctly. I furthermore confirm that all research and experiments were carried out in compliance with the rules regulations, guidelines and ethics that were established by Brac University's research ethics council. The data that are collected and have been used are managed with highest care to maintain privacy and confidentiality. Furthermore, I confirm that neither I made-up data by my own nor manipulated the results. The conclusions and suggestions that are made in this thesis are solely based on my research's result and is unaffected by external factors. I am fully aware that I might face disciplinary action if any violation of standards in academic honesty and moral conduct in research occurs. In my future endeavors I promise to upload the highest level of academic honesty and moral research practices.

Student's Full Name & Signature:

Eshrak

Eshrak Ahmed
19301003

Approval

The thesis titled “Integrity Analysis and Detection of Digital Forensic Evidence” submitted by

1. Eshrak Ahmed (19301003)

Of Summer, 2023 has been accepted as satisfactory in partial fulfillment of the requirement for the degree of B.Sc. in Computer Science on September 2023.

Examining Committee:

Supervisor:
(Member)

Farig Yousuf Sadeque

Dr. Farig Yousuf Sadeque

Assistant Professor
Department of Computer Science and Engineering
Brac University

Co-Supervisor:

Rubayat A. Khan

Mr. Rubayat Ahmed Khan
Senior Lecturer
Department of Computer Science and Engineering
Brac University

Program Coordinator:
(Member)

Dr. Md. Golam Rabiul Alam

Associate Professor
Department of Computer Science and Engineering
Brac University

Head of Department:
(Chair)

Sadia Hamid Kazi, PhD

Chairperson, Associate Professor
Department of Computer Science and Engineering
Brac University

Ethics Statement

I promise to maintain the highest standard of ethics while performing the thesis by getting permission from each and every participant legally, be conducting my thesis honestly and maintaining impartiality at all times. I was very honest and transparent with my research and I also recognized my lacking and uncertainties. In my thesis I tried to avoid plagiarism by citing every source that I have used. I feel my work will benefit the society and people especially in those areas where maintaining integrity of image is very crucial.

Abstract

Technology has improved people's day to day activities like how we communicate and access information. These days people are equipped with a digital camera and mobile phone and they tend to record almost everything happening around them like capturing food they are having or capturing beautiful sceneries around them. Maintaining image integrity is not crucial in informal situations but it is very important to maintain for forensics scientists who are dealing with digital forensic evidence. Recently in our country, a new law has been passed which states that from now on digital proofs can be used in court as evidence. As we know, digital files can be modified; hence the authentication of each and every piece of digital evidence has to be verified manually by experts. There are some researches on this, but could not find any feasible publicly available datasets to work on tools that can detect tampered automatically. My target is to build a dataset consisting of copy-move and cut-paste image forgeries created from the original images; and build a system with CNN models that will detect and automatically exclude photographs that have obvious, and medium levels of modification which will ease the pressure on digital forensics scientists.

Keywords: ; digital proofs; digital files; digital evidence; copy-move; cut-paste; dataset; CNN

Dedication

I dedicate this thesis to my parents as they have been my greatest motivation, encouragement and supported me unwaveringly. I also dedicate this thesis to my supervisor and co-supervisor because without their guidance I wouldn't have been able to complete my thesis. Lastly I would like to dedicate this to digital forensic scientists.

Acknowledgement

Firstly, praise be to the Great Allah for enabling the uninterrupted completion of my thesis. Secondly, I am very grateful to my advisor, Dr. Farig Yousuf Sadeque and my co-advisor Mr. Rubayat Ahmed Khan who motivated me to do thesis in this topic and supported me in every possible way throughout one year and were available whenever I needed help. I am also grateful to Brac University Computer Science and Engineering department for providing me enough resources and facilities. Lastly, my parents for their immense support and prayer throughout my journey. I am currently just a few days away from graduating thanks to their thoughtful support and prayers.

Table of Contents

List of Figures

List of Tables

Chapter 1

Introduction

1.1 Motivation

Beyond our common people's imagination, technology has enhanced our lives and permanently altered the way we perceive the world. In today's world, practically everyone is a mobile journalist equipped with a sound and video system, and everything is continuously being recorded. People photograph almost everything around them, including picturesque fields, calm highways, blue skies, and mesmerizing rural areas. People who live in cities also photograph structures, foods, and other things. The majority of the time, these media assets are utilised for purely private purposes, such as sending images of several shirts to friends so they may select one, discussing events or situations with friends, etc. The media's integrity is not necessarily necessary to be upheld in these informal situations. It's acceptable if someone changes the media. The integrity of the media, however, takes ultimate priority over all other considerations in some professions, and the original form of the media must always be provided. One image may have the ability to choose between a lifetime behind bars and a life of freedom in the ”digital forensics of evidence” profession. Integrity is crucial in these situations because of this. In other words, before presenting a media file as proof, we must make sure that it hasn't been altered. There are several algorithms available for identifying numerous types of alteration to achieve so in the instance of images. But as media manipulation technology has advanced, it has also gotten to the point where it is quite challenging to tell whether or not a picture has been altered. Along with that there are not many publicly available datasets to work on image forgeries. Only two or three are available but they have their drawbacks; like one not having the original images of the tampered images and another one 90 percent images are greyscale images so the models don't train properly. So, I am trying to build up a dataset and then work on models that would work on multiple image forgery types to detect whether it is tampered or not.

1.2 Objective and Aim

I want to build a system with my own dataset that would try to detect whether any image has been tampered or not and the models chosen would work on multiple type image forgeries. The system would detect through the regions of the image inconsistencies in the pixel distribution, color, or texture of an image that indicate manipulation and in lastly it would show whether the image is tampered or not.

1.3 Method

My primary objective is to detect any kind of fake or forgery image. To begin with, I divided my working approach into modules so that I could focus on my dataset. Firstly, I gathered my own 1000 images and then tampered them with two types of forgeries to create my dataset. Since my dataset consisted of raw images so I pre-processed them using pre-processing techniques, resized my images to work with the models and then trained and the models with training portion of my dataset. These were evaluated by accuracy and loss curves for models by importing Python libraries such as matplotlib, Numpy, pandas and also using forged images and finally testing to detect whether the images that were provided were forged or not.

1.4 Research Problem

Due to advancement in technology, image forgery has increased drastically within past few years. Forging images is fine as long as people uses it for personal use but when it come to digital forensics it is important to find out whether pictures are real or modified so it is important to detect tampered images. A lot of research has been conducted in this field using CNN models, statistical models but the only problem is the availability of publicly available datasets. The publicly available datasets that are available, each has its own flaws. MIC-f220 contains only 100 real images and 100 forged images. CG-1050 v2 contains mostly greyscaled images and MSCOCO doesn't have the authentic images of the forged images. So it is important to create a dataset with significant amount of original images and its forged images so that more extensive research can be done in this field.

1.5 Thesis Orientation

Firstly, I discussed about the related that have been done related to image forgeries. It includes varies techniques that have been applied to detect image forgery. Then I presented a outline of my system. After that I discussed the components and my dataset like how I build up my dataset that contributed to my output and implementation is discussed. Lastly the accuracy of the models and testing are discussed.

Chapter 2

Literature Review

2.1 Morphological Filter Detection

The knowledge and methods for authenticating multimedia artefacts and reconstructing their processing histories have significantly advanced over the past ten years among scientists and practitioners in the field of multimedia forensics. Paste detection, resizing, re-compression, picture augmentation, discrepancies in image geometry and illumination owing to suspected manipulation, and many forms of nonlinear filtering are all included in this large category of particular manipulation detectors. Very little research has been done on morphological filters, which are often employed in image processing for artefact removal and picture enhancement, in the context of nonlinear filter detection. The detection of this kind of filtering is of interest in the context of image phylogenesis and in the identification of specific manipulations in legal scenarios. However, it could also be very helpful to identify potential counter-forensic attacks based on morphology, which is very effective in removing local noise and could be used at the end of the image editing process to cover other types of traces. By precisely recognising the use of morphological filters on both uncompressed and compressed photos, the suggested extension operates with grayscale images. The suggested approach makes a distinction between two scenarios: morphological filtering in raw pictures and post-processing such as compression, noise addition, and filtering. The tests made use of the UCID, DRESDEN, and RAISE datasets, which are all openly accessible. [8].

2.2 Watermarking Approach

Watermarking based approach is one of the most popular method to find forgery image. According to [16] a novel watermarking technique that detects manipulation on paper. The approach makes use of the Code Division Multiple Access (CDMA) technique. It is necessary to encode the building blocks required to produce the Common Code (watermark) data before using a Walsh table. Every block in this embedding technique contains the LSB. Throughout the detection and verification procedures, the encode (data representation) matrix, which forms the foundation for the semi-blind detection step, is crucial. The proposed strategy was evaluated against various types of attacks and changed procedures. The recommended approach offers high PSNR and SSIM values. The quality of watermarked photos is quite great. Even when the watermark is incorporated into the original image, there is no deterioration.

2.3 Related Works

[4] In this paper they used Alexnet model in MICC-F220 dataset, a publicly accessible benchmark dataset containing 110 non-forged and 110 forged pictures with 3 channels (colour images), sizes ranging from 722*480 to 800*600 pixels, was utilised in the study which consisted mainly copy-move forgeries. They achieved very good accuracy using this model.

This research paper worked on making dataset which is known as CG1050v2 Original and Tampered datasets [9]. Here they created a dataset from various objects and scenarios with 1050 images and then creating their corresponding tampered images. But almost 90 percent of the images here are greyscaled so using this dataset the models don't train properly. Another dataset [11] which is MICC-f220 consists of original and tampered images but it consists of only 220 images and not all original images have been tampered in this dataset.

From [3]. In this paper, they worked on detecting copy-move forgery. For the purpose of detecting copy-move forgeries, the study suggests a model dubbed BusterNet . The three primary parts of BusterNet, an end-to-end deep learning system, are CNN Feature Extractor, Similarity Detection (Simi-Det), and Manipulation Detection (Mani-Det).The VGG16 architecture, which was pretrained on the ImageNet dataset, forms the foundation of the CNN Feature Extractor. From the supplied photos, it extracts high-level characteristics.Simi-Det is in charge of spotting comparable areas in the picture. To compare and forecast the similarity between pairs of picture patches, it employs a Siamese network architecture. A similarity map that emphasises areas with high similarity scores is the result of Simi-Det.Mani-Det is made to find areas in a picture that have been altered. Utilising a convolutional neural network (CNN), manipulation detection is carried out utilising the similarity map from Simi-Det. Mani-Det creates a manipulation map after estimating the likelihood that each pixel will be altered. The final forgery detection map is created by fusing the similarity map with the manipulation map. The possibility of each pixel being a component of a copy-move fake is shown on the forgery detection map. A synthetic dataset is used to train BusterNet for the identification of copy-move forgeries. Each component is trained individually, the branches are frozen, the fusion module is trained, and ultimately the complete network is fine-tuned from end to end.They used CASIA CMFD dataset and CoMoFoD dataset by customizing models in old dataset. They improved the accuracies of the dataset by about 4 percent and 7 percent respectively.

In the paper [5] the researchers used CASIA v2 dataset which consists of copy-move and cut-paste forgery. Here, the model is a convolutional neural network (CNN) called VGG-16. It comprises of two fully connected blocks and two convolutional blocks, and it takes as input picture patches of size 40x40x3. Two convolutional layers with the ReLU activation function are present in each convolutional block, followed by a pooling layer. Overfitting is prevented by the use of dropout layers. The supplied data is adjusted to fit inside the range. There are 869,154 parameters in the network as a whole. They have achieved good accuracy as they used this model. In [12] researchers have used Alexnet model in CASIA dataset consisting of copy-move and cut-paste image forgery and they obtained very good accuracy with that model.

In this [6] paper the researchers worked with copy-move, splicing, object-removal, and morphing forgeries together. Here, they used models on their own dataset like ELA (Error Level Analysis), NOI1 (Noise-based method using high-pass wavelet coefficients), CFA1 (Camera's filter array patterns), and DCT (JPEG Blocking artifacts) and a learning based method Bayar et al.'s Constrained Convolutional layer to detect these forgeries. The Dataset that they used DEFACTO dataset which is a novel dataset created especially for the research and development of picture modification detection techniques. The dataset produced semantically meaningful forgeries by using the Microsoft Common Object in Context database (MSCOCO). It has more than 200,000 photos, and each one has multiple comments that help identify forgeries and provide details on how they were altered. Their highest accuracy was 55 percent and lowest was 20 percent.

Chapter 3

Methodology

3.1 Workflow

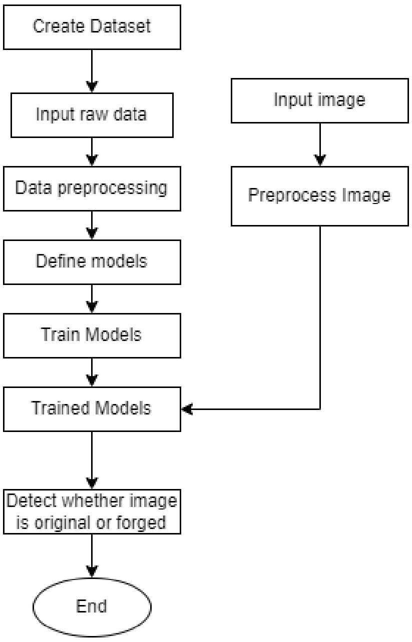


Figure 3.1: A flow chart of integrity image detection system

First of all, I begin with raw images from my dataset. Then the raw images are pre-processed using pre-processing techniques. Then the images would be divided into two parts that is testing and training. After that, the models would be trained on the training set of images. After the models such as Alexnet, Vgg16 and Vgg19 are trained the testing set of images would be used to find out how well the models are able to detect image forgeries. With the help of models, we would know whether the image if forged or not. This workflow has been designed to detect whether image has been forged or not and create a efficent method to detect copy-move and cut-paste image forgeries Finally the accuracy of the models would be determined by loss and accuracy graphs and then testing would be done.

3.2 Dataset

Every day, image alteration tools and techniques improve in complexity, making it nearly difficult to tell whether a picture has been altered or not. The variety of methods that people can edit photos makes it more difficult to identify tampering. However, if a picture is changed, there are typically still some footprints there, and I am looking for evidence of that. As photographs may be edited in many different ways, multiple types of image flaws is needed to be looked upon in order to determine whether or not the image has been altered. First of all I started to look for suitable dataset for my research but most of the datasets are not publicly available. The datasets that were available didn't seem well suitable to me to work on those. MSCOCO dataset has many images but the dataset does't have the original images of the tampered ones. MICC-F220 dataset contains only 220 images and not all original images have been tampered in this dataset. Another publicly available dataset CG-1050 contains mostly greyscale images; that is about 92 percent images are greyscaled so the models are not able to train properly to detect colourful images since these days mostly coloured images are used. So I started collecting my own dataset. I took pictures as I went to various places in Bangladesh like Cox Bazar, my village, various places in Dhaka city like Dhanmondi, Bijoy Sarani etc. and also outside of Bangladesh like India, Dubai, USA. In total the dataset contains 3000 images. The number of original images are 1000. These images were manipulated with Adobe Photoshop and Canva. Each original images were modified by two types of image forgeries which are copy-move and cut-paste.

Copy- move is a type of forgery where a certain region of the image is copied and then it is pasted into another region within the same image. It is done to create identical objects, regions, people within the same image to make it seem that the duplicated region is original. It is done firstly by selecting a region that would be copied and then from original location the copied portion is pasted in another in the same image and it is carefully done so that it seems real.

In cut-paste forgery certain part of one image is cut and then pasted onto another image. To create this type of forgery first an image is selected where the certain portion is cut. Then the forger selects another image where the portion would be pasted.

The number of forged images are 2000; that is copy-move forgery has 1000 images and cut-paste forgery has 1000 images. This is done to create a more diverse dataset so that if researchers use this dataset for one of either type of forgeries they can use it and also this dataset would be useful if people are designing models for these two type of forgeries together. The forgeries were done in such a way that they are not detectable by naked human eyes. These images were also being converted to 128*128 to work with models as different images consisted of different dimensions. The images below show some original images and their corresponding tampered images.

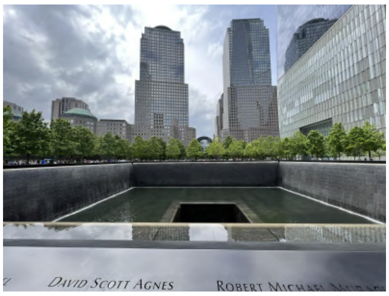


Figure 3.2: Original image 1

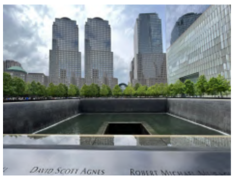


Figure 3.3: Copy-Move

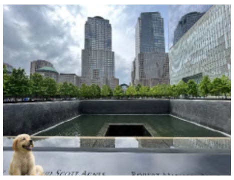


Figure 3.4: Cut-paste

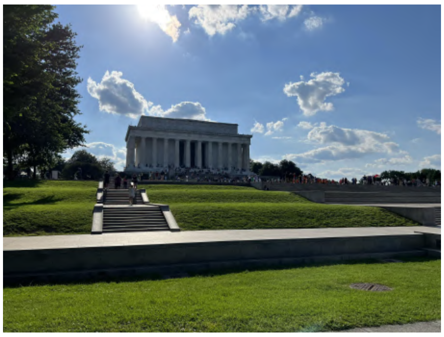


Figure 3.5: Original image 2

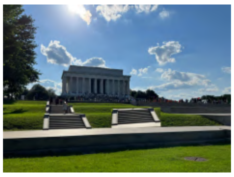


Figure 3.6: Copy-Move

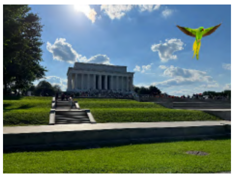


Figure 3.7: Cut-paste

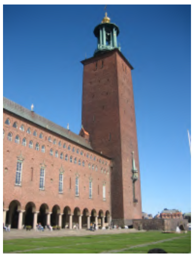


Figure 3.8: Original image 3

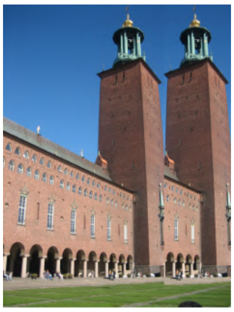


Figure 3.9: Copy-Move

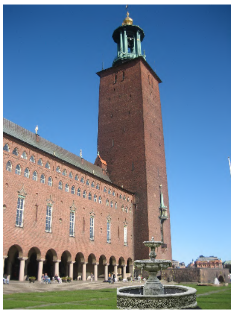


Figure 3.10: Cut-paste

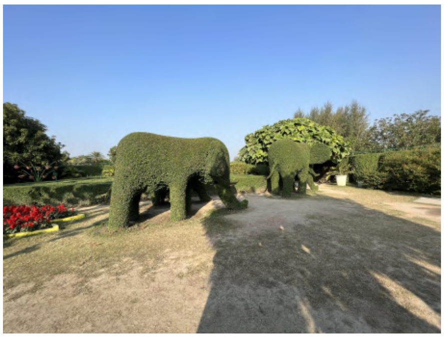


Figure 3.11: Original image 4

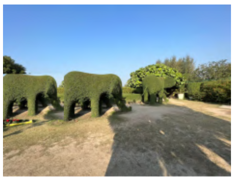


Figure 3.12: Copy-move

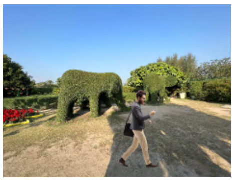


Figure 3.13: Cut-paste

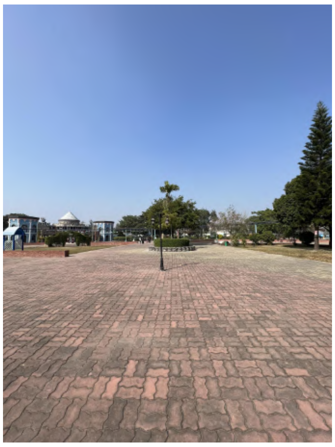


Figure 3.14: Original image 5

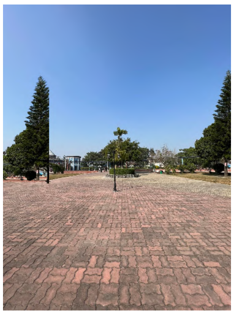


Figure 3.15: Copy-Move

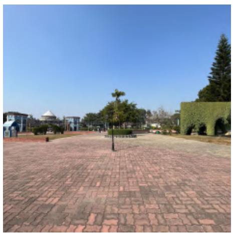


Figure 3.16: Cut-paste

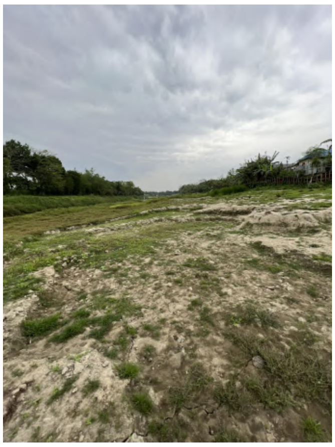


Figure 3.17: Original image 6

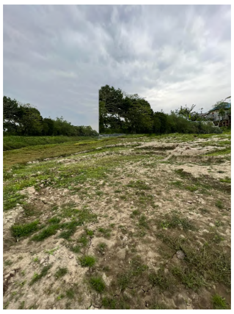


Figure 3.18: Copy-Move

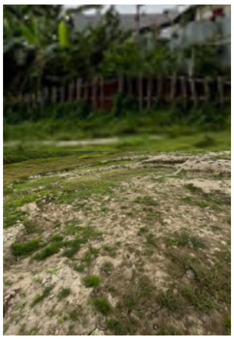


Figure 3.19: Cut-paste

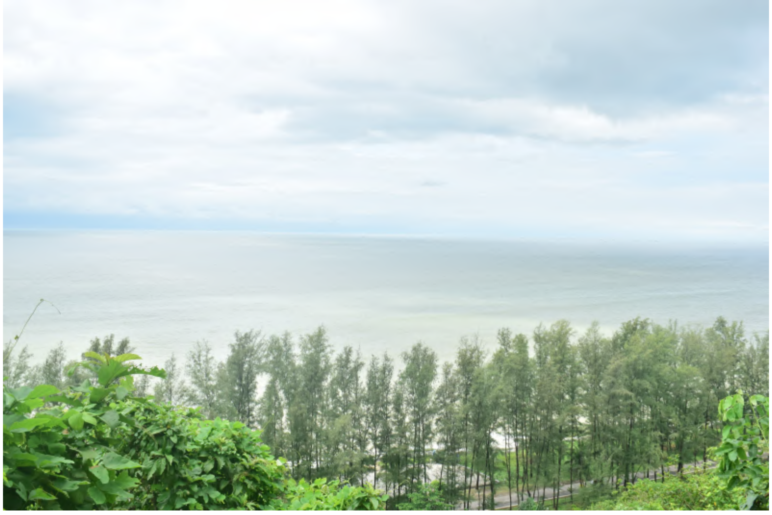


Figure 3.20: Original image 7

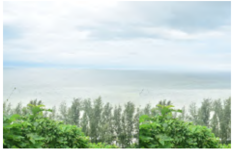


Figure 3.21: Copy-move

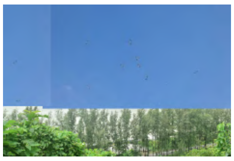


Figure 3.22: Cut-paste

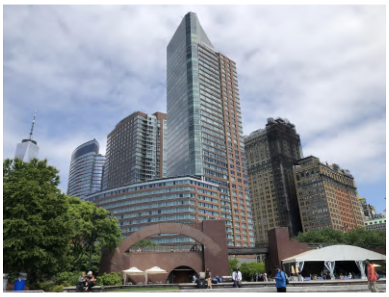


Figure 3.23: Original image 8

.

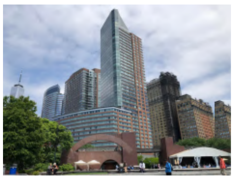


Figure 3.24: Copy-Move

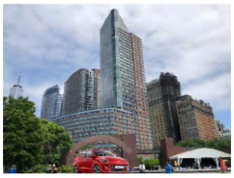


Figure 3.25: Cut-paste

A model needs to be trained and after training a model must be tested on a different set of data known as the testing set once it has been trained. The model is run once on the testing set once the training phase is finished to show it can reliably predict results based on recent, unused data. The model should be evaluated using comparable data. Consequently, the testing set should be an accurate representation of what it would see in real-world data. So the dataset is split into 90 percent for training and 10 percent for testing.

3.2.1 Data Augmentation

Data augmentation is a technique used in computer vision and machine learning to produce extra copies of existing data in order to artificially expand a dataset. This effect has been achieved by cropping, rotating, flipping, altering brightness and contrast, adding noise, and making other adjustments to the original data. Data augmentation may be used to increase the dataset's variety and improve the performance of machine learning models when the original dataset is limited or unbalanced. We can produce fresh variations of the original data to enhance the model's capacity to generalize to novel, unseen data. Data augmentation may also benefit music, video, and images. For example, in image classification tasks, data augmentation techniques can involve randomly cropping or rotating the photos, flipping them horizontally or vertically, changing the colour or brightness, and adding noise or blue. It is crucial that the transformations employed in the augmentation process precisely reflect the changes in the data that occur in the actual world in order to guarantee that the improved data appropriately represents the original data. Additionally, regularisation should be used together with data augmentation to prevent the model from being overfit to the augmented data. In order to augment my dataset, I have used ImageDataGenerator in google collab. I have used techinques such as rotation, shift,flip to augment my dataset.

3.3 Models

3.3.1 Alexnet

One Convolutional Neural Network Architecture (CNN) utilised for image processing is Alexnet.[2] claims that it has an eight-layer CNN network with 60 million parameters and 650,000 neurons. It has five convolutional 3 Fully Connected Layers and Layers. The output of the last fully connected layer is sent into a 1000-way softmax, which creates a distribution across the 1000 class labels. The multinomial logistic regression goal, which is comparable to maximising the average of the log-probability over training examples of the predicted distribution for the correct label, is what the network aims to maximise. The second, fourth, and fifth convolutional layer kernels are only connected to those kernel mappings in the previous layer's kernel. The third layer's kernels are connected to each of the second layer's convolutional layer's kernel mappings. The neurons in the fully connected layers are connected to all the neurons in the layer above. Response-normalization layers come after the first and second convolutional layers. Max-pooling layers come before both the response-normalization layers and the fifth convolutional layer. The output of each fully connected and convolutional layer is affected by the ReLU (Rectified Linear Units) nonlinearity. In order to filter the 224*224*3 input image, the first convolutional layer uses 96 kernels of size 11*11*3 with a 4-pixel stride, which is the distance between consecutive neurons' receptive field centres in a kernel map. The second convolutional layer receives the output of the first layer and filters it using 256 kernels of size 5*5*48. The third, fourth, and fifth convolutional layers are connected to one another without any pooling or normalising layers. The third convolutional layer has 384 kernels with a size of 3 * 3 * 256 that are coupled to the outputs of the second convolutional layer. In comparison to the fourth convolutional layer, which has 384 kernels of that size, the fifth convolutional layer has 256 kernels of size 3* 3* 192. There are 4096 neurons in each of the entirely connected layers. Advantages of the model includes: compared to models like VGG16, the model is shallower, but it provides equivalent or, in certain circumstances, substantially more accuracy than VGG16 or other models. Using convolution layers, the edges of the images may be automatically retrieved, and fully connected layers can learn these traits. Performance is improved by allowing direct picture input to the classification model as well as scale pooling, spatial pyramid pooling, and side supervision using AlexNet. The architecture of the AlexNet model is depicted in the image below.

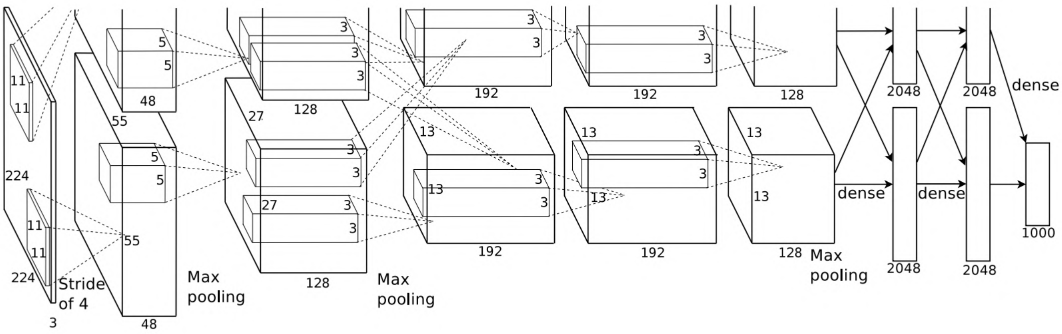


Figure 3.26: Alexnet architecture

3.3.2 Vgg16

Vgg16 is a convolutional neural network design and an image processing model. The Vgg-16 architecture is recognised as one of the greatest vision model architectures to date.[15] . The most notable aspect of VGG16 is that it consistently used the same padding and maxpool layer of 2x2 filters with a stride 2 and prioritised having convolution layers of 3x3 filters with a stride 1. Convolution and max pool layers are set up in the same way across the whole design. Two completely linked layers and a softmax for output are included in the conclusion. The number 16 in VGG16 stands for the 16 layers with weights. This network is a huge network with about 138 million parameters. The model's architecture includes a single maxpool layer with a 2x2 pool size and a 2x2 stride, two convolution layers with the same padding and 128 channels each, and two convolution layers with the same padding and 64 channels each. 3 x 256-channel convolution layers with the same padding and 3x3 kernal structure; 1 maxpool layer with a 2x2 pool size and 2x2 stride; a single maxpool layer with a 2x2 pool size and a 2x2 stride; three convolution layers with the same padding and 512 channel 3x3 kernal; a single maxpool layer with a 2x2 pool size and a 2x2 stride; and one maximum pool layer with a 2x2 pool size and a 2x2 stride. [14]. Some advantage of using the model is- The VGG 16 model may contain a large number of weight layers because to the convolution filters' low size; more layers naturally result in greater performance. The model's top-5 test accuracy in ImageNet, a dataset of more than 14 million images separated into 1000 classes, is 92.7 percent.
The figure below shows the architecture of Vgg16.

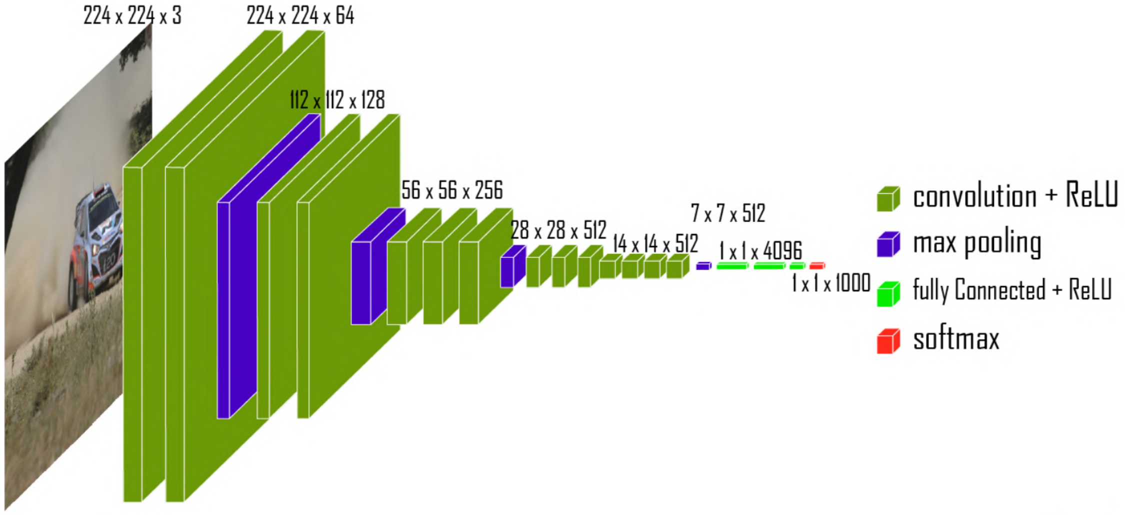


Figure 3.27: Vgg16 architecture

3.3.3 Vgg19

The Visual Geometry Group (VGG) has created a deep convolutional neural network model called VGG19 [18]. The VGG16 model, which was the 16-layer version of the VGG architecture, is an expansion of that model. The number "19" in VGG19 refers to the total number of layers, which includes 3 fully linked layers and 16 convolutional layers. The VGG19 model is well renowned for its consistent architecture and simplicity, making it simple to comprehend and use. However, it is computationally costly, especially for training on big datasets, because to its depth and amount of parameters. Firstly in input layer he model requires a fixed-size input picture with RGB channels that is generally 224x224 pixels in size. The network receives the input picture and runs it through a number of convolutional and pooling layers in order to learn hierarchical features. Then in Convolutional Blocks each of the five convolutional blocks in VGG19's architecture consists of a number of convolutional layers, followed by a max-pooling layer. As we delve farther into the network, there are more filters (or channels). To preserve the spatial dimensions, the convolutional layers employ tiny (3x3) receptive fields with a stride of 1 and a padding of 1. After that in Activation Function, except for the last convolutional layer, each convolutional layer is followed by the ReLU (Rectified Linear Unit) activation function. Then there is Max Pooling layer each pair of convolutional layers is followed by the application of a max-pooling layer with a 2x2 window and a stride of 2. The most crucial traits are preserved while the spatial dimensions are reduced with the aid of max pooling. Then in FULLY CONNECTED LAYER The output is flattened and fed through three fully connected layers after the last convolutional block. The number of output classes for the ImageNet dataset, which served as the foundation for VGG19's training, is represented by the number of neurons in the first two completely connected layers—4096 each—and the third fully connected layer—1000. In SOFTMAX LAYER a softmax activation function, which transforms the raw scores into probabilities for each class, comes after the last fully connected layer. As a result, the model is able to forecast the class with the highest likelihood as its output.

The figure below shows the architecture of Vgg19.

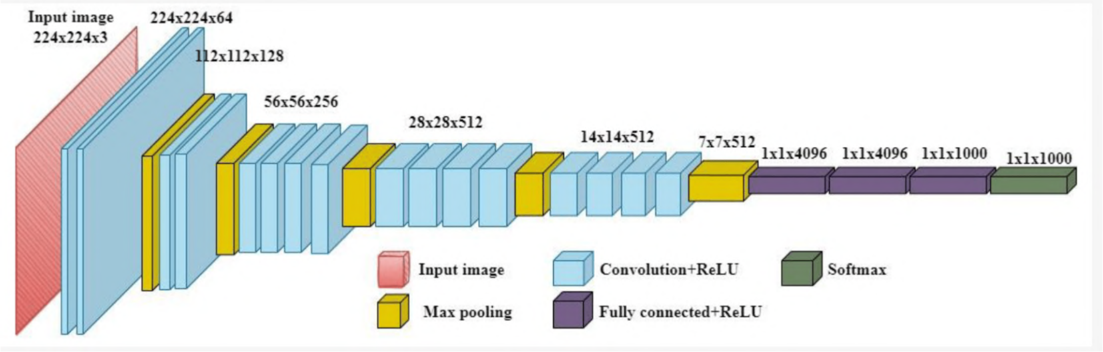


Figure 3.28: Vgg19 architecture

Chapter 4

Implementation

To detect whether any image has been forged or not. I made my own dataset consisting of 1000 original images and then creating their tampered 2000 images with photo editing software. After that I used image augmentation for better training of my models. I used CNN models to detect image forgery as these models have shown great results in this field according to previous researches done. Cnn models work when a dataset contains proper amount of original and its forged pictures. CNNs are quite good at automatically identifying important elements in photos. The network gains the ability to distinguish between patterns, textures, and structures that are present in both real and fake pictures throughout training. The convolutional layers of the network are used to teach these characteristics. The network learns to distinguish between authentic and forged images by adjusting its internal parameters. So it can be said CNNs have the ability to recognise a variety of picture forgeries, such as copy-move, cut-paste, retouching, and more. CNNs improve the precision and automation of forgery detection systems by taking use of their capacity to learn complicated visual patterns. CNN models like Alexnet Vgg16 and Vgg19 were implemented; trained on my dataset and tested to detect whether the images have been modified or not.
I started my work by creating my own dataset by collecting it and then editing it on photo editor softwares to create their tampered version.Then I imported libraries necessary for my work in google collab. Then the first thing I did was to pre-process the data since my dataset consisted of raw images of various forgeries like copy-move, cut-paste images. Pre-processing converted the images into numpy arrays then the arrays were normalized.

The images were also converted into 128*128 to work perfectly with the models used here.
After that models ALexnet Vgg16 Vgg19 were trained with my data and 30 epochs were used .These models work by feature extraction, fine tuning and loss function. In Alexnet model feature extraction works by consisting of multiple layers that can learn hierarchical features from images and then the final layer is removed to retain the convolutional layers.

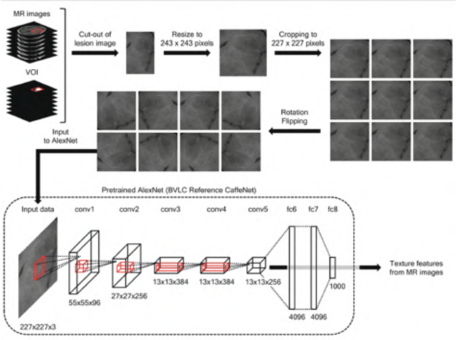


Figure 4.1: Feature extraction of Alexnet

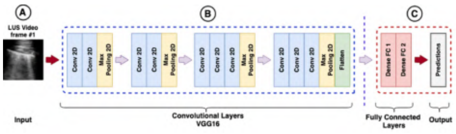


Figure 4.2: Feature Extraction of VGG16

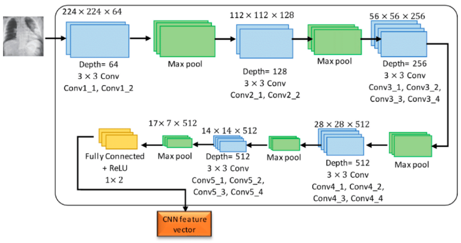


Figure 4.3: Feature Extraction of VGG19

Every layer has distinct input and filter sizes, as well as the matching numbers of filters and output maps. For both models fine tuning works by adding new layers on top of the feature extractor to adapt the model to the image forgery detection task. The added layer captures specific characteristics of the forged regions in the images.

Then appropriate loss function measures the difference between the predicted and originals. Then the models were trained. After that the models were tested to see whether the models are able to detect image forgery or not on the forged images. After that graphs of accuracy vs epoch and loss vs epoch were found out. The reports were produced by importing python libraries such as Mathplotlib. In general, the report produced by my study is built to be user-friendly and intuitive, making it simple for people to understand and analyse the information. The report's visual depiction of the data offers an easy-to-read and understandable summary of the ability to detect image tampering. Then the models were validated with the validation portion of the data and finally tested to see whether the models are able to detect image forgery or not on the forged images.

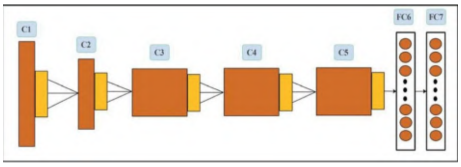


Figure 4.4: Fine Tuning of Alexnet

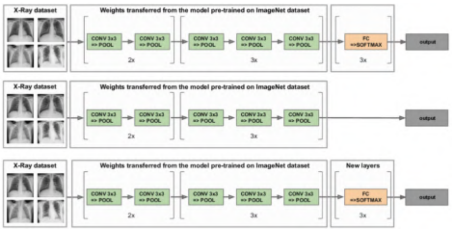


Figure 4.5: Fine Tuning of VGG16

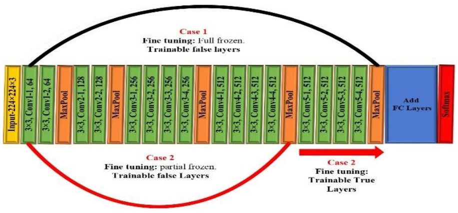


Figure 4.6: Fine Tuning of VGG19

Here, Alexnet model consists of 8 layers (5 convolutional layers and 3 fully connected layers) here. Kernel size is (11*11) with stride (4*4) in the first layer and then smaller kernels (5*5) and (3*3) is used with strides (1*1). Max-pool size of (3*3) is used with stride (2*2). In VGG16 16 layers are used (13 convolutional layers and 3 fully connected layers.) Kernel size is (3*3) with stride (1*1). Max-pooling size of (2*2) is applied with stride of (2*2). In VGG19 19 layers are used (16 convolutional layers and 3 fully connected layers). Kernel size is (3*3) with stride (1*1). Max-pool size of (2*2) with stride (2*2) is used.

Chapter 5

Results

The performance of the work was evaluated by processing the raw images from the combined dataset then augmenting images. After that training and testing the models. After finding the results the accuracy of the work was found out.The results of these works and some testing whether the system detects forged images or not are shown below.

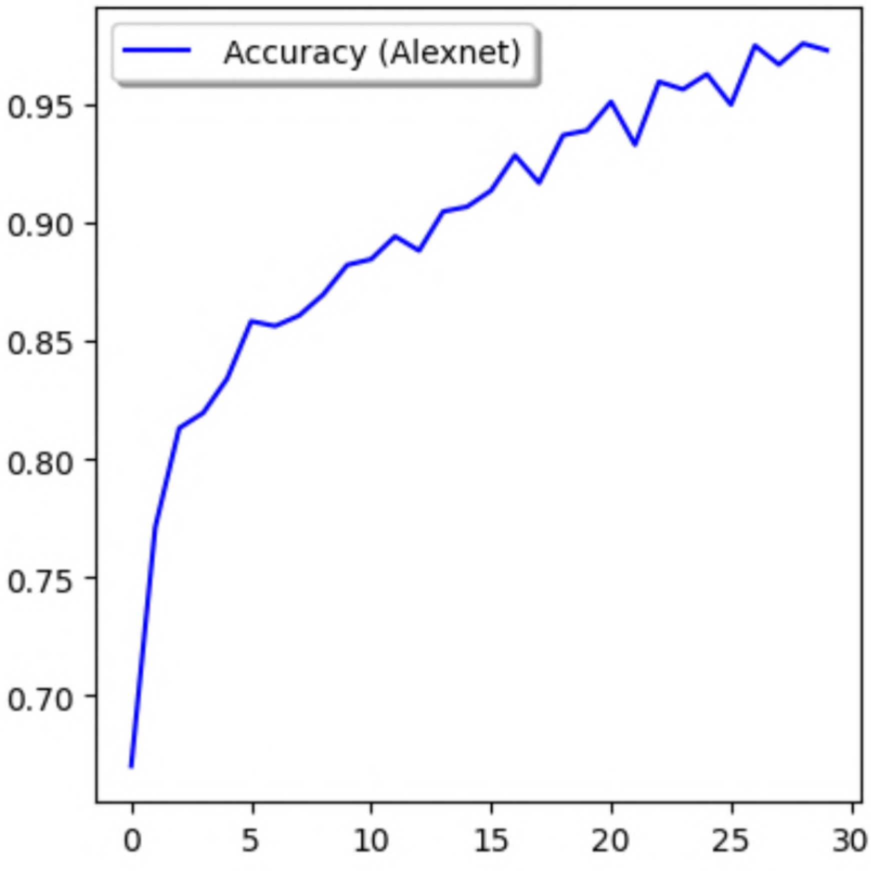


Figure 5.1: Accuracy Graph of Alexnet

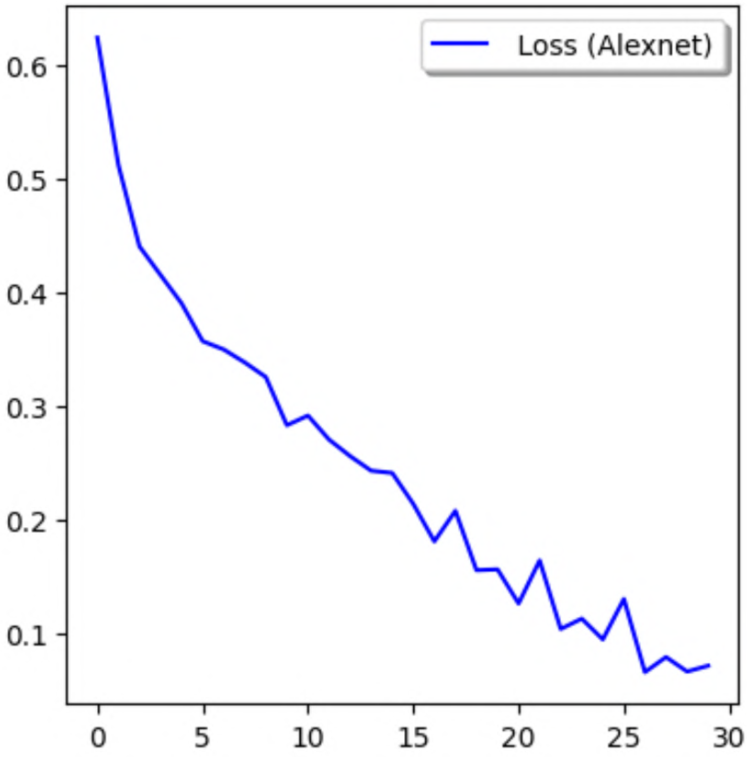


Figure 5.2: Loss Curve of Alexnet

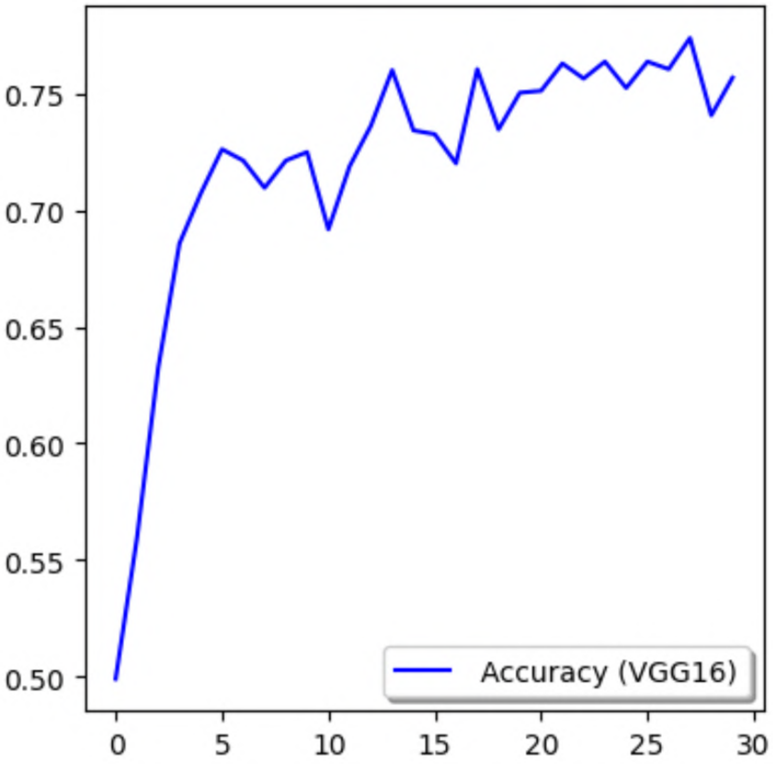


Figure 5.3: Accuracy Graph of Vgg16

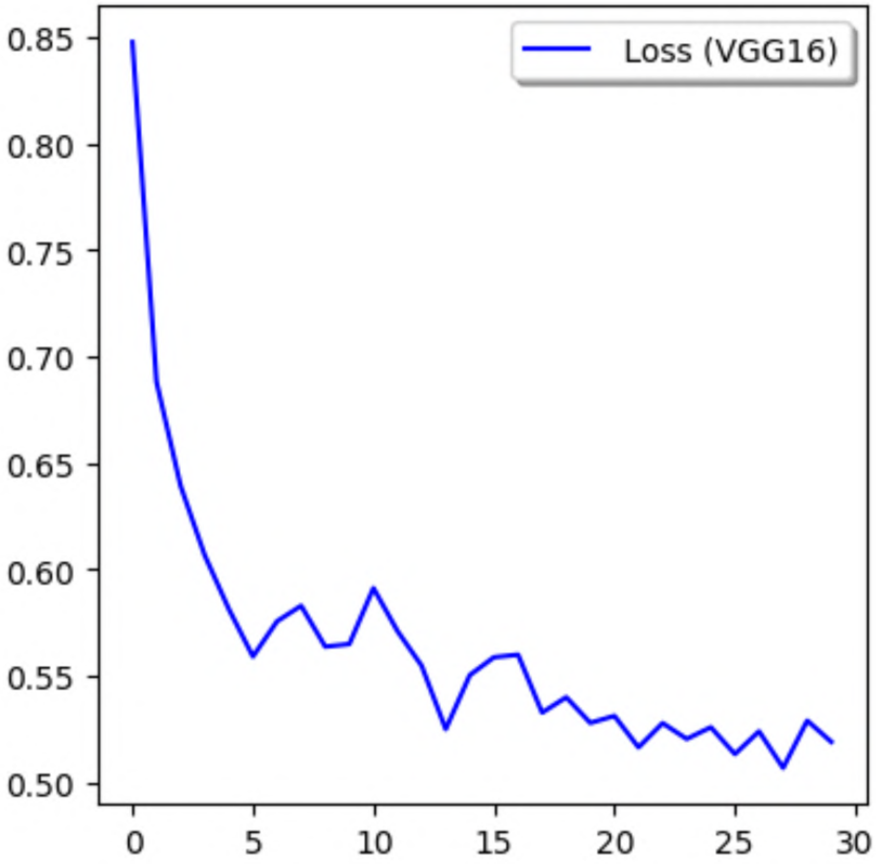


Figure 5.4: loss curve of Vgg16

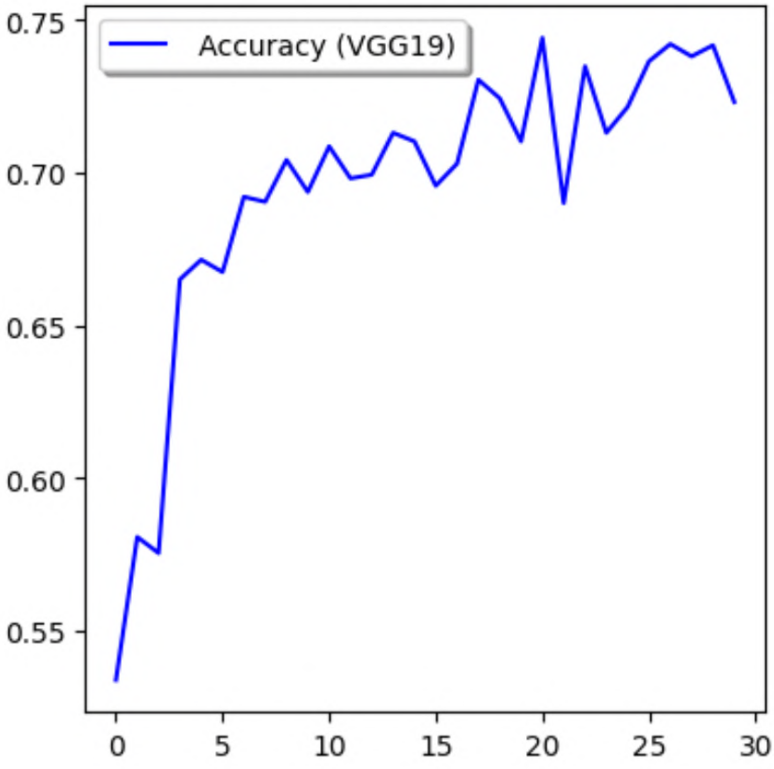


Figure 5.5: Accuracy Graph of Vgg19

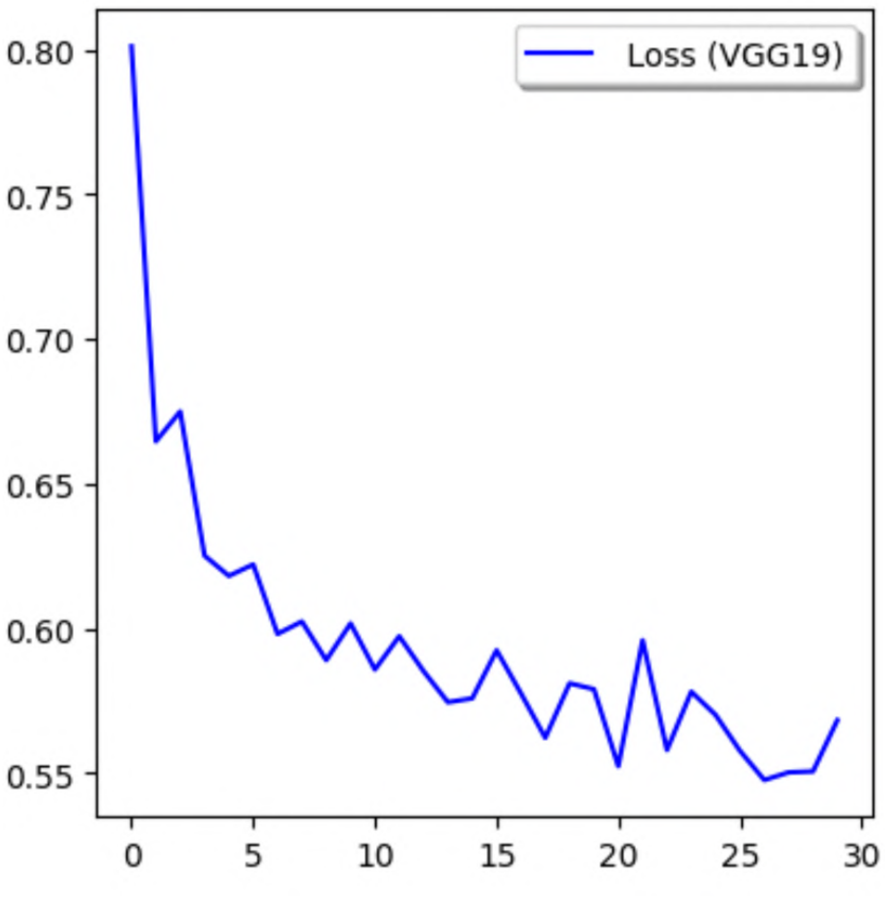
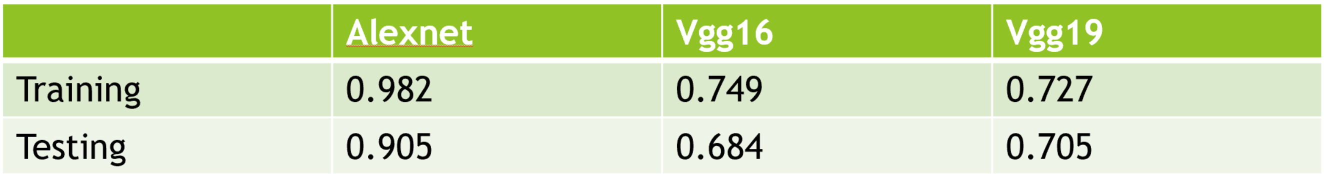


Figure 5.6: Loss curve of Vgg19

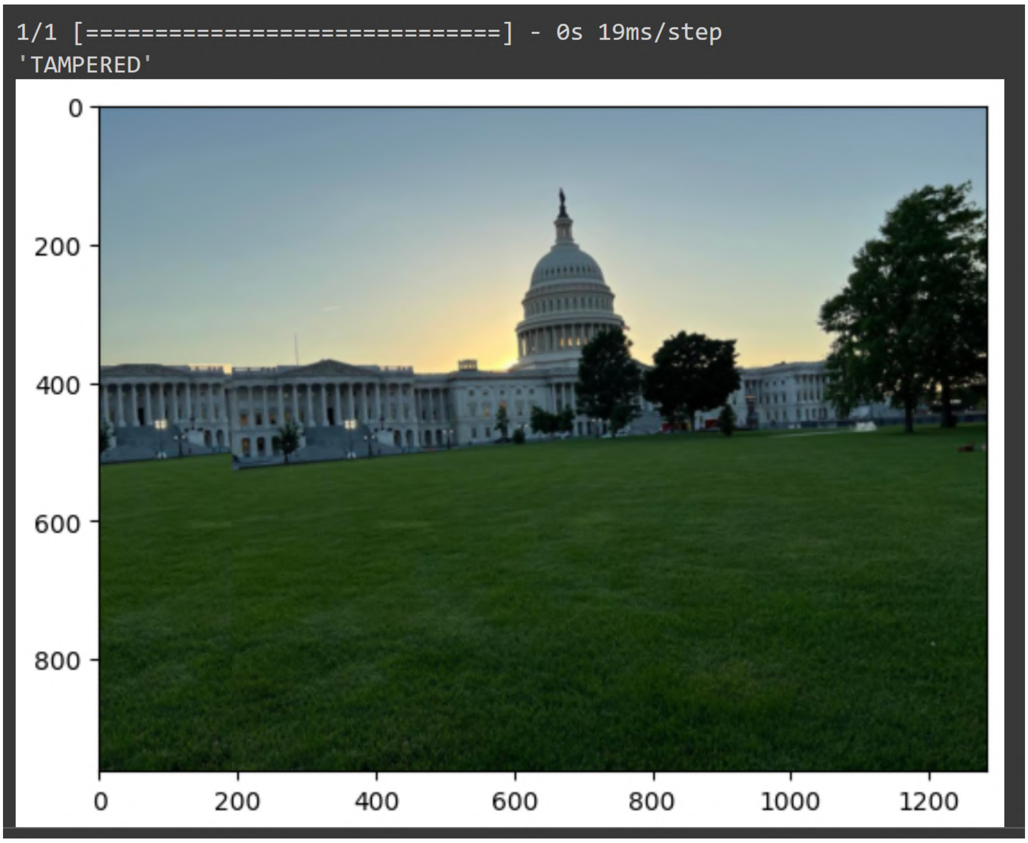
The table below summerizes the testing and training accuracies-



	Alexnet	Vgg16	Vgg19
Training	0.982	0.749	0.727
Testing	0.905	0.684	0.705

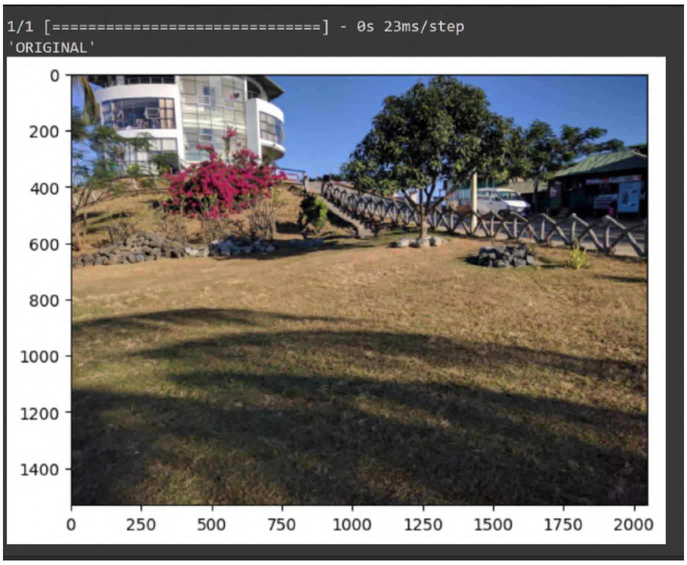
Table 5.1: Summerization of Scores for Training and Testing

Some testing are shown below.



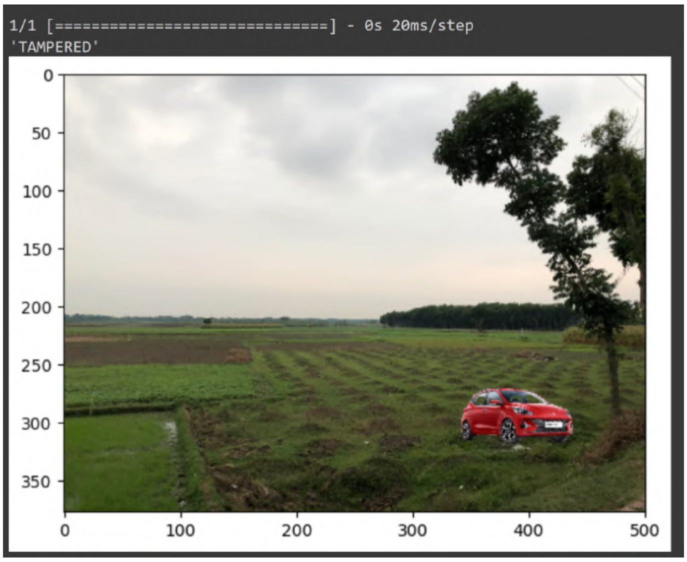
```
1/1 [==============================] - 0s 19ms/step
'TAMPERED'
```

Figure 5.7: Testing1



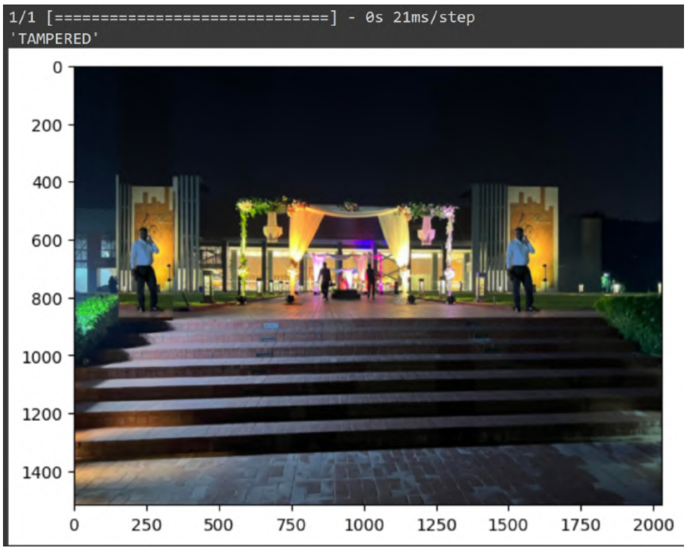
```
1/1 [==============================] - 0s 23ms/step
'ORIGINAL'
```

Figure 5.8: Testing 2



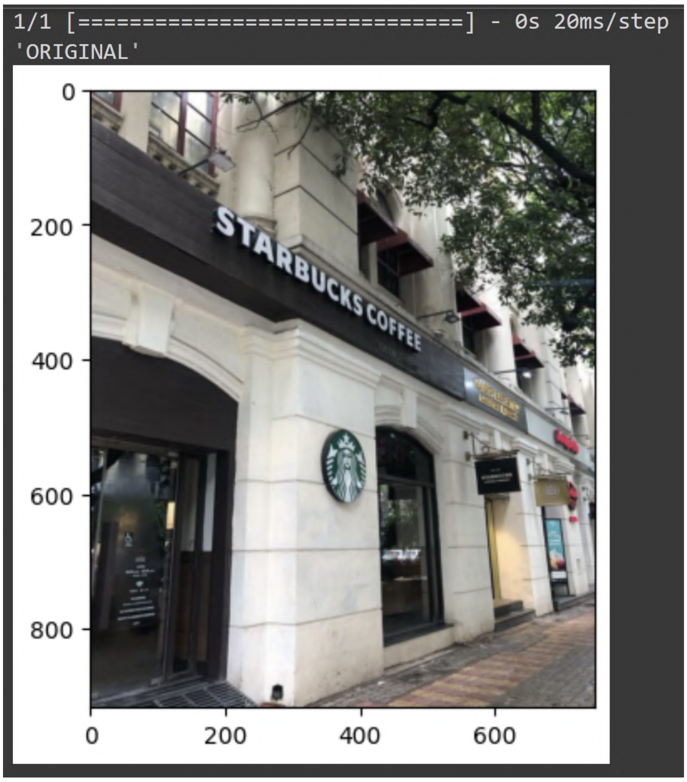
```
1/1 [==============================] - 0s 20ms/step
'TAMPERED'
```

Figure 5.9: Testing 3



```
1/1 [==============================] - 0s 21ms/step
'TAMPERED'
```

Figure 5.10: Testing 4



```
1/1 [==============================] - 0s 20ms/step
'ORIGINAL'
```

Figure 5.11: Testing 5

Chapter 6

Conclusion

Over the past few years, technology has advanced significantly. In a short amount of time, computers, phones, and cameras are being upgraded, making them more potent and superior. People are supposed to benefit from these improvements, yet some people try to abuse this development. One of them entails image modification.Using the combination of image processing and CNN models such as Alexnet and Vgg16 my system is able to detect whether the image is forged or not. The use of libraries such as Numpy, ImageDataGenerator, matplotlib, allows to implement the work in a very well-organized way. The system's ability to detect manipulated image can be helpful to digital forensic scientists as it is crucial for them to find out whether image is real or forged. This system can also be extended to scenarios where image forgery detection is important like evidences presented in court, medical imaging, insurance claims, and journalism, personal reasons etc. and the dataset made by me which consists of original images and its copy-move and cut-paste image forgery can be used further to work on by other researchers and improve research on image forgery.

Therefore, the system designed can result in benefit to decrease use of forged images. This can indicate that less people would use forged images to get away with crimes or any illegal works that they are trying to do. Hence, if any image is forged appropriate measures can be taken against the people according to law. In general, this system provides a viable means to stop using forged images for unfair advantages and my dataset provides original and forged images of the original images with two forgery categories which is used in my research and can be further used in future researches to design more efficient and better models since most of the datasets are not publicly available which makes it very difficult to obtain those datasets.

6.1 Challenges

During this research of detecting image forgeries I faced multiple challenges using image processing and CNN models. One of the main challenges was to find a dataset with significant number of original and its corresponding tampered images that is suitable for my research. Most of the datasets either don't fulfil my requirements or they are not suitable to work with. I also faced challenges creating a diverse dataset which contains various type of pictures. It is difficult to create forged pictures that are not detectable through naked eyes.

At first, I wanted to work with medical images but there are no suitable available data for this. For example- [10] contains images of forged and original X-rays to detect covid-19 but since the images are bit unclear and they contain only a square making it unsuitable to work with. I found another medical dataset [19] but that dataset was very difficult to work with mainly beacuse the dataset was extremely large and were saved in DICOM format which needed to be converted to pictures first with 3-D numpy arrays. I tried to do it in google collab but the format was not coverting to pictures with the code that they provided. Then some datasets I found are very good and matches well with my requirement like [17] but these are not open datasets and requires subscription and even permission to unlock it. Then I decided to create my own dataset with significant number or original images with its corresponding forgeries.

Few datasets such as defacto-copymove, defacto-splicing contains only the forgered images not the original ones. The dataset CG1050v2 [9] contains mostly greyscaled images so the models would not train properly and also presently coloured images are used greyscaled images are not used much. Furthermore, though the dataset [7] is enriched with original and tampered images but working with it won't be fruitful as it contains AI images not of any real life images like objects, people, scenarios etc.

There were also challenges finding which models would be suitable for work. Since there are numerous algorithm available for forgeries some work on single forgeries and some models can be used on multiple forgeries as proven by previous researches. Like in [1] they used SIFT and SURF models for their dataset. This model can only be used in copy-move not other type of forgeries. In [13] they worked on splicing image. They used RTAG model which can only be used for cut-paste. CNN models like Alexnet, VGG16, VGG19 can be used for both copy-move and cut-paste forgery.

Although there is one dataset [6] where the researchers worked on combining multiple forgeries like my research but they used copy-move, splicing and inpainting and used statistical models to train the dataset which is different from my research. For this I looked for which models have been used for detecting image forgeries in more than one type. Like in this paper [20] they used ALexnet to detect copy-move forgery and in the paper [4] they used Alexnet to detect both copy-move and cut-paste but the copy-move dataset only had copy-move type forgery and splicing dataset had splicing type forgery so they trained the models separately on different datasets. It is also difficult to train models with very quality image or very large datasets as it requires very high computational power.

To conclude, my research has bought challenges and limitations with CNN methods to detect image forgeries. Despite the difficulties, I have shown that CNN techniques may be used to create accurate and dependable models for detecting image tampering. My contribution and also making a dataset consisting of original and its corresponding manipulated images. I intend to overcome the shortcomings of my model and dataset in further work and create more precise and reliable approaches for spotting image forgery and try to increase my dataset with more diversity of forgeries.

6.2 Contribution

Through my thesis work, I have significantly impacted the field of image forgery. One of my main contribution is in the area of image forgery detection. The system that is designed in my research work can detect tampered images which is useful in areas where it is important to find out whether images are modified or not like in court or when digital forensic scientists are dealing with images for various purposes. I also created my own dataset with copy-move and cut paste forgeries from the original images; since there is very less publicly available datasets each having their own flaws. Hence, it can be said my system offers effective solution in the field of image forgery and dataset can be used for further research since it is difficult to get a suitable dataset which is available publicly.

My work in the field of image forgery is also supported from previous research work that have been done. For example the MICC-F220 datatset also has dataset consisting of copy-move and cut-paste forgeries but the number of images are very few. CG-1050 v2 also has various types of forgeries with significant number of images but the problem is most of the images are greyscaled so the models give very less accuracy. Though, there have been only one research work combining various type of forgeries. This work has been done on the DEFACTO MSCOCO dataset where the researchers combined few forgeries which is copy-move, morphing, inpainting and splicing. This paper encouraged me to make dataset combining few type of forgeries to create an efficient system instead of just working with one category. This is also suppported from previous research as the models used here have been used before to detect copy-move and cut-paste forgeries. There are other algorithms For instance- for copy-move SIFT and SURF algorithm is used. For splicing DCT is used but these algorithm works for one type of forgery. I think since my work is completely new as it as no dataset has this significant amount of images with two equal corresponding copy-move and cut-paste forgeries. In future there would be further works improve the accuracies that my model gave for this research. It can also be noted that from my dataset which would be publicly available there can be further research in this field as it is difficult to get a suitable dataset due to not being publicly available.

To conclude in terms of image forgery , my thesis work has significantly advanced the subject of image forgery detection. Image forgery detection using my system is a dependable and effective option, enhancing to stop unfair use of manipulated image. This research work indicated that this work is bit different because till now each original image was tampered with one type of forgery but here I made two types of forgeries with significant number of original images. Finally, it can be said that this work made significant contribution to the field of image forgery and make people's life easier and open doors for further research and enhancement in this field by creating an efficient system and a dataset for image forgery detection and research.

6.3 Limitations

My thesis has some limits, just like any other research study, and they must be acknowledged. I would acknowledge about the limits of my study in this part. My thesis's dependence on the quality and accessibility of datasets is one of its main weaknesses. Finding the right datasets for my study was difficult, and the quality of the datasets I found wasn't sufficient. According to Akhtar and Mian (2018), the quality of the training data has a significant impact on how well an image forgery model performs, and using low-quality data can have this effect. Also the forgery detection models require good quality pictures not any sort of blur pictures or very bad quality pictures which is another limitation of my research; the models would not train properly as it depends on the quality of the input picture. As my system is only trained with a mixture of few categories so it would not work with other type of image forgeries that are present. This is one of the common limitations of image forgery models that have been trained so far. Training forgery detection models requires large and diverse datasets that contain various types of original and forged pictures of people is difficult to get due to privacy, not being a free source and ethical concerns. Some datasets are either too large which requires high computational power and many days to train or the datasets are too small that is training with those datasets doesn't give reliable results. When it comes to identifying complex forgeries, such as Deep Fakes, they frequently perform badly. During the analysis phase, the choice of initiation method and detection location (pixel/region) becomes incompatible. As a result, perfect automation for this phase is not possible. Instead, they virtually always require a specialist to get involved in the procedure. It has been discovered that CNN-based image forgery detections are only effective in a small number of situations, such as device detection and copy-moving identification. Also, models that have been used in research till today mainly focuses on improving the existing accuracies of models that is trained to detect only a particular category of forgeries but no work is done much and not given much importance to detect multiple forgeries together which is a limitation on the whole image forgery research. Moreover till now and including my research medical images forgeries are not given importance although its also important to detect whether images of X-rays or lung cancer are forged or not.

The dataset used here tried to cover various scenarios but still there are plenty of objects and scenarios in the world that have been not covered till now.

In conclusion, my thesis has shown promising results in identifying tampered picture which include the reliance on the reliability and accessibility of primary datasets, the restrictions on the types of forgeries to be identified, the impact of image quality on the accuracy of the image forgery models, the reliance on reliable input data for the models, and the model scalability. For my used models to continue to be effective in detecting picture forgeries, I think it will be essential to solve these constraints.

6.4 Future Work

I have noticed a number of potential areas for further research as I try to enhance and improve my thesis model.

One of the most important area for future work is to improve my existing models by increasing the accuracy of the models. While my model has shown significant result but still this is not enough and there is huge room for improvement. I also want to explore more CNN models such as Inceptionv3, Efficientnet to check whether the accuracy increases or decreases or remains the same. Another area of future work is to include more type of image forgery such as morphing, inpainting along with my included forgeries.

The forgeries discussed above contains only images of various objects and people but there are signature based forgeries should also be combined with those type of forgeries. Detecting whether an image containing a fake signature or not very crucial in this modern age as number of frauds with signature is increasing. Although there are few researches regarding it but this category more efficient models should be designed and dataset should be made to train models to detect it. In this Digital age fingerprint forgery is also increasing so there should be more research regarding it. There is also few works regarding detecting whether images containing image of iris are modified or not. Scanning iris also increasing these days in various uses in modern digitisation. If these few categories can be worked and trained with appropriate models then the system would be more efficient systems to detect more various types of forgeries. There have also been very little work on detecting fake videos. And finally, although there have huge research on improving to detect forged human face but still there have been no research done to detect whether any face is forged or not wearing face mask. After Covid-19 though now most people don't wear masks still there are significant number of people who still wear masks so it is also important to detect fake human face with mask along with detecting human face with no masks. Also there have been almost no research and also there is no proper dataset available to detect forgeries in case of medical sectors such as detecting forged X-Rays from images and detecting of fake Lung cancer from images or dataset containing both real and fake lung cancers. Also a system should be made to detect and be able to tell which type of forgery occurred would be useful as the people dealing with forensic evidence can know exactly which type of forgery has happened which can be useful in situations like presenting evidence on court.

The dataset can be further improved by including more images of other type of forgeries like inpainting and morphing; covering more objects and scenarios and various light conditions that have not been used and also using more people with their consent which would help to diversify the dataset hence training of the models would be more good and improved.

To summarize, I think my thesis has produced promising outcomes in the area of picture forgery detection. I do, however, recognise that much more work has to be done in order to increase the precision and effectiveness of my model. More work is also needed to be done to improve the dataset to improve variations. I intend to make considerable progress in enhancing picture forgery detection efforts and ultimately ending the use of fabricated images and making some resources for further improved research for unfair advantage by investigating the areas of future study mentioned above.

Bibliography

[1] V. Christlein, C. Riess, J. Jordan, C. Riess, and E. Angelopoulou, "An evaluation of popular copy-move forgery detection approaches," *IEEE Transactions on information forensics and security*, vol. 7, no. 6, pp. 1841–1854, 2012.

[2] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "Imagenet classification with deep convolutional neural networks," *Communications of the ACM*, vol. 60, no. 6, pp. 84–90, 2017.

[3] Y. Wu, W. Abd-Almageed, and P. Natarajan, "Busternet: Detecting copy-move image forgery with source/target localization," in *Proceedings of the European conference on computer vision (ECCV)*, 2018, pp. 168–184.

[4] 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.

[5] A. Kuznetsov, "Digital image forgery detection using deep learning approach," in *Journal of Physics: Conference Series*, IOP Publishing, vol. 1368, 2019, p. 032 028.

[6] G. Mahfoudi, B. Tajini, F. Retraint, F. Morain-Nicolier, J. L. Dugelay, and P. Marc, "Defacto: Image and face manipulation dataset," in *2019 27th European Signal Processing Conference (EUSIPCO)*, IEEE, 2019, pp. 1–5.

[7] X. Zhang, S. Karaman, and S.-F. Chang, "Detecting and simulating artifacts in gan fake images," in *2019 IEEE international workshop on information forensics and security (WIFS)*, IEEE, 2019, pp. 1–6.

[8] G. Boato, D.-T. Dang-Nguyen, and F. G. De Natale, "Morphological filter detector for image forensics applications," *Ieee Access*, vol. 8, pp. 13 549–13 560, 2020.

[9] M. Castro, D. M. Ballesteros, and D. Renza, "A dataset of 1050-tampered color and grayscale images (cg-1050)," *Data in brief*, vol. 28, p. 104 864, 2020.

[10] L. Li, J. Bao, T. Zhang, *et al.*, "Face x-ray for more general face forgery detection," in *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, 2020, pp. 5001–5010.

[11] J. Y. Park, T. A. Kang, Y. H. Moon, and I. K. Eom, "Copy-move forgery detection using scale invariant feature and reduced local binary pattern histogram," *Symmetry*, vol. 12, no. 4, p. 492, 2020.

[12] S. Samir, E. Emary, K. El-Sayed, and H. Onsi, "Optimization of a pre-trained alexnet model for detecting and localizing image forgeries," *Information*, vol. 11, no. 5, p. 275, 2020.

[13] X. Bi, Z. Zhang, and B. Xiao, “Reality transform adversarial generators for image splicing forgery detection and localization,” in *proceedings of the IEEE/CVF international conference on computer vision*, 2021, pp. 14 294–14 303.

[14] M. Hassan, *Vgg16–convolutional network for classification and detection. neurohive*, 2021.

[15] R. Thakur, “Step by step vgg16 implementation in keras for beginners (2019),” *URl: https://towardsdatascience.com/step-by-step-vgg16-implementation-in-keras-for-beginners-a833c686ae6c*, 2021.

[16] M. AlShaikh, “A novel tamper detection watermarking approach for improving image integrity,” *Multimedia Tools and Applications*, pp. 1–22, 2022.

[17] F. Hossain, A. Gul, R. Raja, T. Dagiuklas, and C. Galkandage, *Forgery image dataset*, 2022. DOI: 10.21227/9dmj-yn86. [Online]. Available: https://dx.doi.org/10.21227/9dmj-yn86.

[18] T.-H. Nguyen, T.-N. Nguyen, and B.-V. Ngo, “A vgg-19 model with transfer learning and image segmentation for classification of tomato leaf disease,” *AgriEngineering*, vol. 4, no. 4, pp. 871–887, 2022.

[19] S. Solaiyappan and Y. Wen, “Machine learning based medical image deepfake detection: A comparative study,” *Machine Learning with Applications*, vol. 8, p. 100 298, 2022.

[20] S. Tinnathi and G. Sudhavani, “Copy-move forgery detection using superpixel clustering algorithm and enhanced gwo based alexnet model,” *Cybernetics and Information Technologies*, vol. 22, no. 4, pp. 91–110, 2022.