Digital Fake Image Detection Based on ANN and ML

A

Project Report

submitted in partial fulfillment of the requirements for the

award of the degree of

BACHELOR OF TECHNOLOGY

in

COMPUTER SCIENCE & ENGINEERING

With Specialization in

E-Commerce and Retail Automation

by

Name	**Roll No.**
Shivang Raj	R114216021
Nitish Kumar	R114217005

under the guidance of

Dr. Piyush Chauhan

Department of Informatics

School of Computer Science

University of Petroleum & Energy Studies

Bidholi, Via Prem Nagar, Dehradun, UK

December- 2020

UNIVERSITY WITH A PURPOSE

CANDIDATE'S DECLARATION

We hereby certify that the project work entitled **"Digital Fake Image Detection Based on ANN and ML"** in partial fulfillment of the requirements for the award of the Degree of BACHELOR OF TECHNOLOGY in COMPUTER SCIENCE AND ENGINEERING with specialization in **e-commerce and retail automation** and submitted to the School of Computer Science, Department of Informatics, University of Petroleum & Energy Studies, Dehradun, is an authentic record of my/ our work carried out during a period from **Aug-2020** to **Dec-2020** under the supervision of **Dr. Piyush Chauhan, Assistant Professor, Dept. of Informatics.**

The matter presented in this project has not been submitted by us for the award of any other degree of this or any other University.

Shivang Raj	**Nitish Kumar**
R114216021	R114217005

This is to certify that the above statement made by the candidate is correct to the best of my knowledge.

Date: 10/12/ 2020

Dr. Piyush Chauhan
Project Guide

Dr. T.P Singh
HOD – Informatics
School of Computer Science
University of Petroleum & Energy Studies
Dehradun – 248 001 (Uttarakhand)

ACKNOWLEDGEMENT

We wish to express our deep gratitude to our guide **Dr. Piyush Chauhan**, for all advice, encouragement and constant support he has given us throughout our project work. This work would not have been possible without his support and valuable suggestions.

We sincerely thank to our respected Program Head of the Department, **Dr. T.P. Singh**, for her great support in doing our project in ***Digital Fake Image Detection Based on ANN and ML*** at **SoCS**.

We are also grateful to **Dr. Manish Prateek,** Dean- SoCS, UPES for giving us the necessary facilities to carry out our project work successfully.

We would like to thank all our **friends** for their help and constructive criticism during our project work. Finally we have no words to express our sincere gratitude to our **parents** who have shown us this world and for every support they have given us.

Name:	**Shivang Raj**	**Nitish Kumar**
Roll no:	**R114216021**	**R114217005**

End Semester Report(2020-21)

Project Title

Digital Fake Image Detection Based on ANN and ML

Abstract

In this rising Technological world, different photograph control methods have made it simpler to play with the original image. With the extraordinary accommodation of computerized imaging and PC designs it has become more simpler to adjust the picture substance with no visual follows. Counterfeit pictures could make principal sway on the social and the affordable diaspora, transformed pictures spread quick and could make significant destruction in the general public. To scale these snags, there are different prior procedures like shading recognition, joining discovery, metadata extraction and so on The issue with these procedures is that, it can recognize just explicit altering on pictures which can be avoided effectively in this current mechanical progressed period making an escape clause. This task plans to identify counterfeit pictures utilizing ANN system what's more, various AI calculations which can recognize a wide range of altered picture. The project implements Error Level Analysis over an image which classifies the fakeness of the image using depth compression.

Keywords:: ANN (Artificial Neural Network), ELA (Error Level Analysis)

LIST OF FIGURES

1.Figure 1 - Error Level Analyzed Image

2.Figure 2 – Use Case Diagram

3.Figure 3 – Activity Diagram

4.Figure 4 – Data Flow Diagram

5.Figure 5- Flowchart

6.Figure 6- Output Authentic Image

7.Figure 7- Output tempered image

8.Figure 8- Performance Graph

9.Figure 9- Confusion Matrix

TABLE OF CONTENTS

Contents

1 Introduction **1**

2 Objective/Scope **2**

3 Problem Statement **2**

4 Background Study **3**

5 Methodology **4**

6 Design **4**

6.1 Use Case Diagram 4

6.2 Activity Diagram 5

6.3 Data Flow Diagram 5

7 Implementation **6**

8 Conclusion and future scope **7**

1. Introduction

In the todays world, enormous number of survivors of picture imitation or control are being found. With the accessibility of modest and simple picture altering programming, pictures can be modified even by a fledgling client. Continuous utilization of transformed pictures is causing some genuine weaknesses and progressively diminishing the estimation of computerized picture. These pictures can be utilized to get out phony word and substance over online media, delude the courts regarding confirmations, break any political believability and significantly more like these cases. Current legal strategies require a specialist to dissect the validity of a picture. This venture expects to actualize a framework that can decide if a picture is phony or not with the assistance of AI what's more, ANN system. The cycle of AI is like that of information mining. Both frameworks search through information to search for designs. In any case, rather than separating information for human appreciation just like the case in information mining applications AI utilizes that information to distinguish designs in information and change program activities appropriately. AI calculations are frequently sorted as being regulated or solo. Managed calculations can apply what has been realized in the past to new information. Solo calculations can draw deductions from datasets.

This Project will work with the picture criminology method known as Error Level Analysis. Of the different procedures accessible for picture fabrication recognition, this venture depends on pixel-based procedure [1]. Blunder level examination (ELA) is one of the procedures to recognize picture control by resaving the picture at a particular quality level and afterward figuring the contrast between the pressure levels. On the off chance that the pictures are not modified, 8x8 squares ought to have comparable mistake possibilities. Nonetheless, if the picture is changed, the segment of the picture that have controlled ought to have higher mistake potential than other piece of the picture.

Figure 1: Error Level Analyzed Image

2. Objective / Scope:

- To design an ANN based system which will distinguish and group the transformed pictures.

3. Problem Statement:

The existing fake image detection system detect only specific tampering methods like splicing, coloring etc. We approached the problem using machine learning and neural network to detect almost all kinds of tampering on images.

4. Review of Literature/ Feasibility Study:

Research paper based on introduction study of Tampering and Review of Tempering Detection technique (C. Raja Lakshmi, Dr. M. Germanus Alex and Dr. R. Balasubramanian) **[1]:** In this paper, the authors describe the introduction of fake image detection. Photoshop is amazing tool for altering reality and Tampering can be innocent or evil. Innocent Tampering does not change the content of the image but change images quality. Innocent tampering included various operations such as contrast brightness, adjustment, zooming, and rotation. The evil tampering aims to modify the content of the image. The evil tampering includes: - Cloning (copy/paste), Original image, Tampered image, Image splicing. Camera-based techniques focus on detecting the traces of tampering. Natural photographs are usually taken under different lighting conditions. Splicing detection techniques traces are left in the anatomy of an image when simple splicing operation is performed, Independent of image forgery is done in single or composites, all digital images are to be stored in any standard format such as JPEG one of the most interplead compression techniques .The paper mainly reviewed the different methods of image forensics.

Research paper based on Development of Photo Forensics Algorithm by Detecting Photoshop Manipulation Using Error Level Analysis(Teddy Surya Gunawan, Siti Amalina Mohammad Hanafiah, Mira Kartiwi, Nanang Ismail, Nor Farahidah Zabah, Anis Nurashikin Nordin) **[2]:** The objective of this paper is to develop a photo forensics algorithm which can detect any photo manipulation. The error level analysis (ELA) was further enhanced using vertical and horizontal histograms of ELA image to pinpoint the exact location of modification. Of the various techniques available for image forgery detection, this paper has highlighted pixel-based technique which is the most common technique. The two most common technique in pixel-based are cloning, also known as copy-move or copy paste, and slicing. Image splicing is done by copying one portion of the image and paste it to another image. Error level analysis is one of the most successful technique to detect image manipulation. In the experiment performed by the authors, the original photo from digital camera was resaved at 75 percent quality level, and then resaved again at 75 percent quality level using Adobe Photoshop. Each picture is associated with the 95 percent ELA for the image.

Research paper based on Fake Image Detection Using Machine Learning (Muhammed Afsal Villan, Johns Paul, kuncheria kuruvilla, Prof. Eldo P Elias) **[3]:** In this paper we learn how to detect a fake images, as in recent times many cases of fake images are spread- ing through digital media and other platform. The tampered images are a detected using neural network which also recognizes the regions of the image that have been manipu- lated and reveals the segments of the original image. The compression ratio of the foreign content in a fake image is different from that of the original image and is detected using Error Level Analysis. [Keywords:- Image forensics, Metadata analysis, Error level analy- sis, Multilayer perception]. Unallocated, free and slack space. Forensics Images can be created through specialized forensics software.Metadata provides information about a picture's pedigree, including the type of camera used, color space information,

and application notes. Different picture formats include different types of metadata. Some formats, like BMP, PPM, and PBM contain very little information beyond the image dimensions and color space. Metadata provides information related to how the file was generated and handled. This information can be used to identify if the metadata appears to be from a digital camera. Error Level Analysis (ELA) permits identifying areas within an image that are at different compression levels. With JPEG images, the entire picture should be at roughly the same level. If a sec- tion of the image is at a significantly different error level, then it likely indicates a digital modification. Multilayer perception network:- Multi Layer perceptron (MLP) is a feedfor- ward neural network with one or more layers between input and output layer. Feedforward means that data flows in one direction from input to output layer (forward). This type of network is trained with the backpropagation learning algorithm. MLPs are widely used for pattern classification, recognition, prediction and approximation. Multi-Layer Perceptron can solve problems which are not linearly separable. A deep neural network (DNN) is an artificial neural network (ANN) with multiple layers between the input and output layers. Each mathematical manipulation as such is considered a layer, and complex DNN have many layers. Neural network has been successfully trained using the error level analysis with 4000 fake and 4000 real images. The trained neural network was able to recognize the image as fake or real at a maximum success rate of 83 percent.

5. Methodology

1. Gathering of information about the topic from various sources.
2. Identification the dataset (what type of dataset is required).
3. Categorization of datasets as training datasets and test datasets.
4. Identification and selection of algorithm for training dataset.
5. Training and testing of dataset.
6. Confusion matrix will be used to evaluate the accuracy of predicted results.
7. At last, comparison of actual result with predicted result will be performed by visualizing both the results on various charts and maps.

6. Design

The objective of this project is to identify fake images (Fake images are the image that are digitally altered) the problem with current existing fake image detection system is that they can detect only one type of error [2]. This project takes the image as input and converts it into Error level analyzed format. One module of the project will resize this image into 100*100 pixel format which could be stored in an double dimension array. Then these 10000 pixel with RGB value (30000 inputs) is given into the layer of multilayer perceptron network. Output layer contain two neurons. One for the fake image depending upon the value of these output value of these neuron. We determine whether the image is fake or not and how much chance is there for given image to be tampered.

Diagrams

6.1 Use Case Diagram

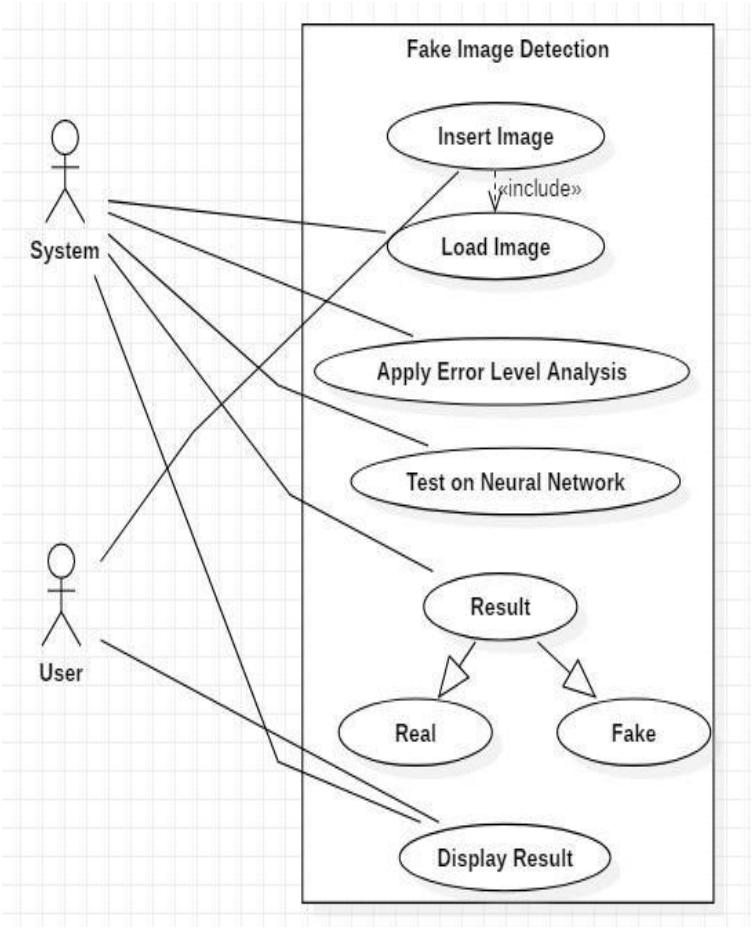

Figure 2: Use Case Diagram

6.2 Activity Diagram

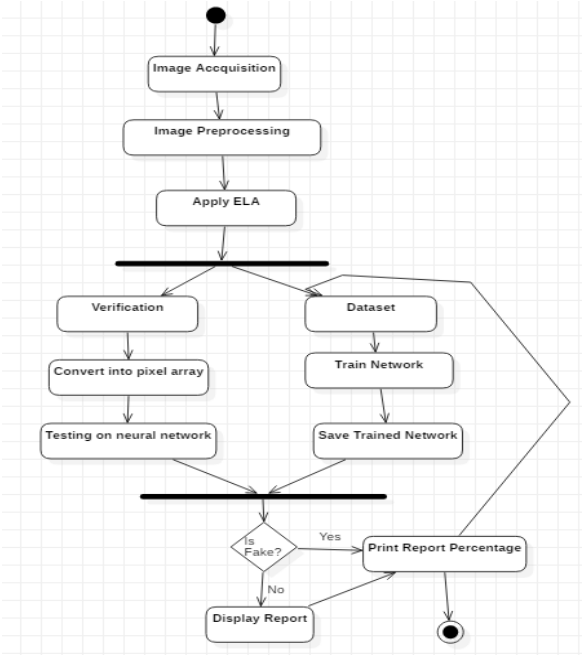

Figure 3: Activity Diagram

6.3 Data Flow Diagram

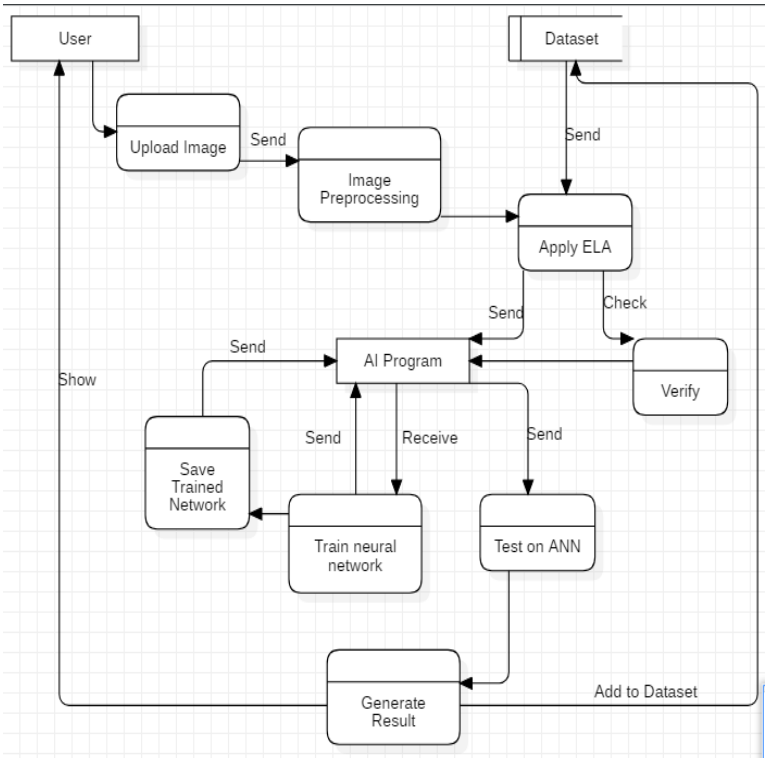

Figure 4: Data Flow Diagram

7. Implementation

7.1 Pseudo Code

int array[1000] [1000];
input image i;
read image i;
ELA (i);
arr[100][100] = ELA(i);
Test Neural(arr [i][j])
if (arr[i][j]!=parameter Database)
int i=1;
else i=0
Save pixel in form of dataset;
if (i==1)
Show fake;
else Not fake;

7.2 Flow Chart

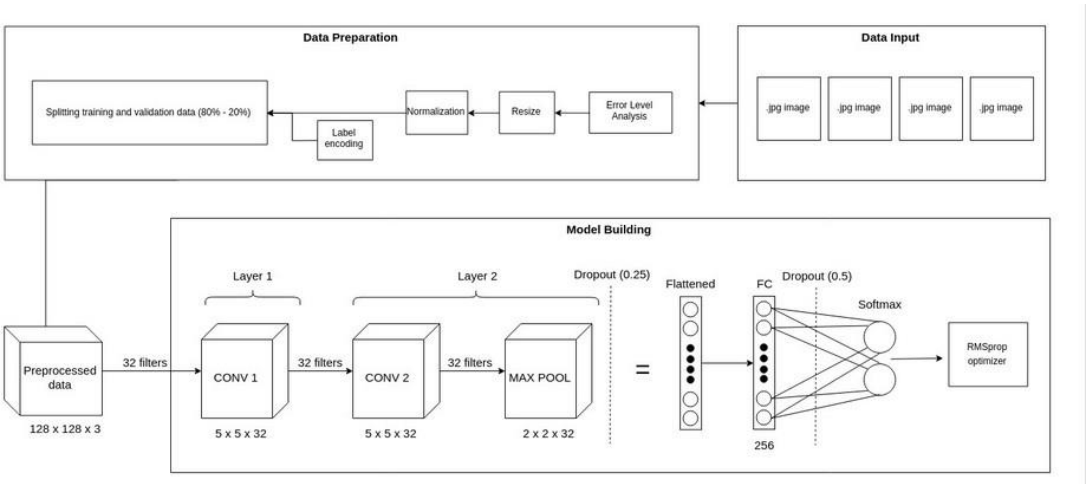

Figure 5: *Flowchart*

7.3 Algorithm For Error Level Analysis

Ela
BufferedImage Compressed = null;
try
tmpName = fname +"temp.jpg";
writeToFile =File(tmpName);
ImageOutputStream imgStream = ImageIO.createImageOutputStream(writeToFile);
Setcompressionquality(compressionlevel)

```
Getdifference(image, compressed)
height= getHeight();
width = getwidth();
if(height==compressed.getHeight() and width ==compressed.getWidth())
int[][][] original = RGBarray(img);
 int[][][] compressed= RGBarray(compressed);
diff[height][width][3];
for(i=0;i¡ hieght; i++)
for(j=0; j¡width; j++
for(band = 0; b¡3 band++)
d= original[r][c][band] - comp[r][c][band];
diff[r][c][band] = (diff[r][c][band]¿255)?255 :(diff[r][c][band]¡0)?0:diff[r][c][band];
return diff;
```

7.4 Algorithm For Convolutional Neural Network

```
Input image (i)
Img arr [i] [j] = convert Image and array (Image(i));
Sum= 0, t =0;
Convolution (int Img arr [] [] ,int fitter[][])
For (i=t; i¡n; i++)
For (j =t; j¡n; j++)
For ( k=t; k¡n; t++)
Sum = sum + img arr[i] [j] * filter [i] [j];
Feature map [j] [k] = sum;
Sum= 0
If( n ¡ Img arr.length)
n=n+1;
t=t+1;
else
Break;
Relu ( feature map [] [])
For (j= 0; j¡img.length; J++)
For (i¡0; i¡img.length; i++)
If (A feature map [i] [j]¡0)
Feature map [i] [j] =0;
Else
Break;
Pooling (feature map [] [])
t=0;
For (i=t ; i¡ img.length; i++)
For( j=t ; j¡ img.length ;j++)
Max =feature map [0] [0]
If (max¡ feature map [i][j])
Max =feature map[i][j]
Pool feature[i] [j]=max;
If(t¡ img.length)
t=t+1;
Flattering (pool.feature [] [])
```

For (i=0; i¡img.length; i++)
For(j=0;j¡img.length;j++)
Flatten[i] =pool feature[i][j]
Full Connection (flatten[])
Node[1000];
For(i=0; i¡img.length;i++)
Node[i]=flatten[i];
Apply Ann node[i];

7.5 Result Analysis

Artificial Neural Network(ANN) is composed of a large number of highly interconnected processing elements(neurons) working in union to solve specific problems. The neurons are connected by links and they interact with each other. The node can take the input data and perform simple operations on the data. Result of these operation is passed to other neurons. Modern Artificial Intelligence(AI) problems are handled by ANN. It has a high capability to learn.

For this project, we have implemented Covolutional Neural Network with 32 filters and 5 kernal size.

Inputs are:

- Authentic Image
- Tempered Image

```
In [28]: feature = feature = array(convert_to_ela_image('C:/Users/Srijan Sah/Desktop/Dataset - Copy/Single Prediction/sri
         feature = np.expand_dims(feature, axis=0)

In [29]: result = model.predict(feature)
         print (result)

         [[0. 1.]]

In [30]: if result[0][0] == 0:
             print ('Real')
         else:
             print('Fake')

         Real
```

Figure 6: *Output for Authentic Image*

```
In [73]: feature = array(convert_to_ela_image(row[0], 90).resize((64, 64)))
         feature = np.expand_dims(feature, axis=0)

In [74]: result = model.predict(feature)
         print (result)

         [[1. 0.]]

In [75]: if result[0][0] == 1:
             prediction ='Fake'
             print ('Fake')
         else:
             prediction = 'Real'
             print('Real')

         Fake
```

Figure 7: *Output for Tempered Image*

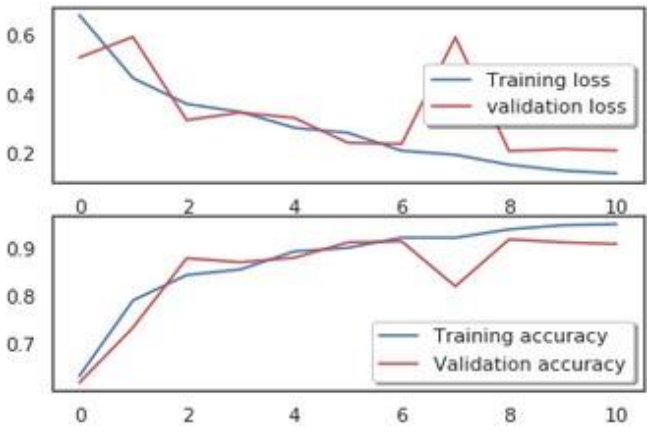

Figure 8: *Performance Graph*

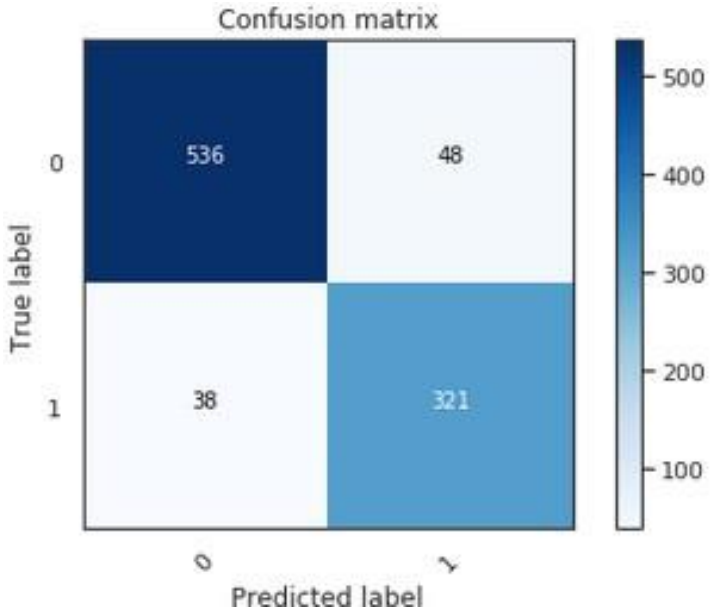

Figure 9: *Confusion Matrix*

8. Conclusion and Future Scope

In the todays technological world, huge number of victims of image forgery or manipulation are being discovered. This project is the solution to all these problems to a certain extent. The project opened many new avenues to us and taught many new things, This project has great potential and could be extended to have certain different use case. The project's efficiency and effectiveness could be improved by increasing the dataset and training the machine for longer period of time. The project could be converted into an application that would give each and every user easy to use platform and the usebility will increase. It could be extended to identify what type of error are occurring and in what part.

References

[1] M. A. Villan and J. Paul, “Fake Image Detection Using Machine Learning,” vol. 7, no. 2, pp. 19–22, 2017.

[2] T. S. Gunawan, S. Amalina, M. Hanafiah, and M. Kartiwi, “Development of photo forensics algorithm by detecting photoshop manipulation using error level analysis Development of Photo Forensics Algorithm by Detecting Photoshop Manipulation Using Error Level Analysis,” no. February 2018, pp. 131–137, 2017.

[3] J. August, I. No, C. Rajalakshmi, and M. G. Alex, “Available Online at www.ijarcs.info STUDY OF IMAGE TAMPERING AND REVIEW OF TAMPERING DETECTION TECHNIQUES,” vol. 8, no. 7, pp. 963–967, 2017.