FAKE IMAGE DETECTION USING CONVOLUTIONAL NEURAL NETWORK

BY

ZOBAER KHAN
ID: 162-15-8017

AND

MD FAHAD AHMAD
ID: 162-15-8077

This Report Presented in Partial Fulfillment of the Requirements for the Degree of Bachelor of Science in Computer Science and Engineering.

Supervised By

Sheikh Abujar
Lecturer
Department of CSE
Daffodil International University

Co-Supervised By

Mr. Abdus Sattar
Assistant Professor
Department of CSE
Daffodil International University

DAFFODIL INTERNATIONAL UNIVERSITY
DHAKA, BANGLADESH
SEPTEMBER 2019

APPROVAL

This Projecttitled "FAKE IMAGE DETECTION USING CONVOLUTIONAL NEURAL NETWORK",submitted by Md. Fahad Ahmad, ID No: 162-15-8077 and Zobaer khan, ID No: 162-15-8017 to the Department of Computer Science and Engineering, Daffodil International University has been accepted as satisfactory for the partial fulfillment of the requirements for the degree of B.Sc. in Computer Science and Engineering and approved as to its style and contents. The presentation has been held on 13September,2019.

BOARD OF EXAMINERS

Dr. Syed AkhterHossain — **Chairman**
Professor and Head
Department of Computer Science and Engineering
Faculty of Science & Information Technology
Daffodil International University

Abdus Sattar — **Internal Examiner**
Assistant Professor
Department of Computer Science and Engineering
Faculty of Science & Information Technology
Daffodil International University

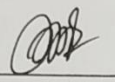
Shah Md. Tanvir Siddiquee — **Internal Examiner**
Assistant Professor
Department of Computer Science and Engineering
Faculty of Science & Information Technology
Daffodil International University

Dr. Dewan Md. Farid — **External Examiner**
Associate Professor
Department of Computer Science and Engineering
United International University

DECLARATION

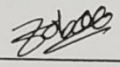
We hereby declare that this project has been done by us under the supervision of **Sheikh Abujar, Lecturer, Department of CSE** Daffodil International University. We also declare that neither this project nor any part of this project has been submitted elsewhere for award of any degree or diploma.

Supervised by:

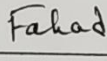


Sheikh Abujar
Lecturer
Department of CSE
Daffodil International University

Submitted by:



Zobaer Khan
ID: 162-15-8017
Department of CSE
Daffodil International University



Md Fahad Ahmad
ID: 162-15-8077
Department of CSE
Daffodil International University

ACKNOWLEDGEMENT

First we express our heartiest thanks and gratefulness to almighty God for His divine blessing makes us possible to complete the final year project/thesis successfully.

We really grateful and wish our profound our indebtedness to **Sheikh Abujar**, **Lecturer**, **Department of CSE, Daffodil International University**, Dhaka. Deep Knowledge & keen interest of our supervisor in the field of "**Fake Image Detection Using Convolutional Neural Network** " to carry out this project. His endless patience ,scholarly guidance ,continual encouragement , constant and energetic supervision, constructive criticism , valuable advice ,reading many inferior draft and correcting them at all stage have made it possible to complete this project.

We would like to express our heartiest gratitude to **Dr. Syed Akhter Hossain, Professor & Head, Department of CSE**, for his kind help to finish our project and also to other varsity Senior students of CSE department of Daffodil International University.

We would like to thank our entire direction mate in Daffodil International University, who took section in this discuss while finishing the direction work.

Finally, we must acknowledge with due respect the constant support and patience of our parents.

ABSTRACT

Nowadays many fake images are expanded through digital media and many newspapers. Images are often directed with the intent and purpose of benefiting one party. In fact, the images are often seen as the reveal of a fact or reality, therefore, false news or any form of printing that using images that have been manipulated or tempering [8] in such a way have the ability and potential to misinform the larger ones. To detection image falsification of a huge number of image data is required, and an architectural model that can process severally pixel in the image. In adding with project, effectiveness and adjustability in the training data is also required to support its usage in daily life. The concept of error level analysis big data and machine learning is the right solution to this type of problem. Therefore, with the Convolutional Neural Network (CNN) [9] architecture that utilizes Error Level Analysis (ELA) [6], image forgery detection can reach above 80% and convergence with only this time.

TABLE OF CONTENTS

CONTENTS **PAGE**

Board of examiners …….. ii

Declaration …….. iii

Acknowledgements…….. iv

Abstract ……... v

CHAPTER 1: Introduction **9-11**

1.1 Introduction ……... **9**

1.2 Motivation ……. **9**

1.3 Rationale of the Study…… **9**

1.4 Research Questions…… **10**

1.5 Expected Output ……... **10**

1.6 Report Layout …… **11**

CHAPTER 2 Background **12-14**

2.1 Introduction …… **12**

2.2 Related Works……. **12**

2.3 Research Summary…… **12**

2.4 Scope of the Problem ……... **13**

2.5 Challenges …… **14**

CHAPTER 3: Research Methodology **15-17**

3.1 Introduction ……... **15**

3.2 Research Subject and Instrumentation …….. **15**

3.3 Data Collection Procedure ……. **16**

3.4 Statistical Analysis …… **16**

3.5 Implementation Requirements …….. **17**

CHAPTER 4: Experimental Results and Discussion **18-21**

4.1 Introduction ……... **18**

4.2 Experimental Results ……. **18**

4.3 Descriptive Analysis ……. **20**

4.4 Summary ……... **21**

CHAPTER 5: Summary, Conclusion, Recommendation and Implication for Future Research **22-23**

5.1 Summary of the Study …….. **22**

5.2 Conclusions ……... **22**

5.3 Recommendations ……. **23**

5.4 Implication for Further Study……. **23**

APPENDIX

REFERENCES **24**

LIST OF FIGURE

FIGURES **PAGE NO**

Figure 2.3.1: Block diagram……..... 12

Figure 2.3.2: Real image and Fake image…… 13

Figure 3.3.1: A) Input image B)Convert input data to ELA …… 16

Figure3.4.1: Accuracy curve and damage curve for training data and validation data …… 17

Figure 4.2.1. Figure 4.1.Model Sequential …… 18

Figure 4.2.2. Figure 4.1.Model Sequential …… 19

Figure 4.2.3 Experiment result …… 19

Figure 4.2.4 Experiment result ……..... 20

CHAPTER 1
Introduction

1.1 Introduction:

There are many tools available in photo editing and photo manipulation [6], which will change our real picture. In Empty eyes, we don't catch what image is original which is fake. Social networks all picture is not correct images. Specifically, Improvements to smart devices like smartphones play a vital role in uploading and downloading images to those social networks. The social network is a platform where socialize, share and spread knowledge But caution is not practiced. Sometimes the images give us wrong information. when is the most manipulation by photoshop or any other editing software is a photo edit has many techniques for manipulating an image using a specific purpose. Pictures mesh can be made by fake for the various purposes it used that why we need accurate information.

1.2 Motivation:

There is much headway of the advanced image handling software and altering tools, a computerized picture can be effectively controlled. An image can be used as valid information because image manipulation requires identification. It used Investigating, and in various cases for the crime. The image detection techniques intend to confirm the credibility of computerized pictures with no correct information about the original image There are numerous routes for altering pictures. For example, coloring resampling, cutout, image retouching, and copy-move [1]. we have examined different types of image fabrication and their detection techniques; mainly we focused on pixel-based image fake detection techniques.

1.3 Rationale of the Study:

Nowadays in the world, so many crimes and harassment or blackmail consist of using fake images. Some people in the world working with image-related research projects such as image manipulating, blind image detect, colorize image detection [1], etc. and simultaneously developed their project. Also, fake image detection related research work available in the world but their project nothing to up to mark success ratio. So,

we are trying to solution perfectly and accuracy fake image detection and reduce the crime and harassment.

1.4 Research Questions:

A research question is a responsible investigation into specific concerns or issues. This is the first step in a research project. 'Getting Started' means once you have an idea of what you want to study, the research question is the first active step in a research project. When you are looking for the answers to all the questions, your research project will be very easy and complete, to begin with.

1.5 Expected Output:

First of all, look forward to expected outputs of a research project are the deliverables and research work properly maintained. When we are in work with fake image detection research project objectives and complete the project properly then we will try to usable and acceptance in the people. Since it is challenging fake image detection project so we will make sure to keep the success ratio above 90%. And this project delivers it to everyone so that they cannot commit the crime and misleading information using fake images.

1.6 Report Layout:

In chapter1: We have to introduce that at present, which digital zones are more vulnerable to use fake image and the and a fake image can do harm to people. We also mentioned about an introduction, motivation, rationale of the study, lookup questions and predicted the result of our thesis project. Later followed by way of the report layout.

In chapter2: We describe the related works, research summary, the scope of the hassle and challenges.

In chapter3: Focus on research methodology and additionally discuss lookup challenge and instrumentation, statistics collection procedure, statistical analysis, implementation requirements.

In chapter4: We describe the important points of experimental outcomes and dialogue which, consists of the following parts such as, experimental results, descriptive analysis, and summary.

In chapter5: We have concluded our comparison outcome and additionally about some different aspects that can be included in future works for the higher of my lookup work.

CHAPTER 2
Background

2.1 Introduction:

Imitation is not new to humanity but it is an exceptionally old problem. Nowadays a picture can be easily controlled and changed due to computerized picture handling software [3] and device changing device. Externally identifying people whether the image is unique or manipulated is a final problem. There is a rapid rise of digitally controlled false-cations on standard media and the Internet.

2.2 Related Works:

In this time many techniques and methods currently available in detection fake images. Currently, most acquirements, manipulation, and cooptation use the JPG standard for Image composition and the number of technology and research is increasing every day. firstly, find the reference image for the image database using a content-based image save system. We focus on media photo images and propose and develop a useful algorithm for detecting fraud zones in the most popular image formats JPEG and other digital camera supported image formats [11]. Although digital image fraud detection has a large number of applications related to the forensic science document Q&A section, it is very helpful for media, publications, law, military, medical image science applications, satellite imagery and the World Wide Web publications.

2.3 Research Summary:

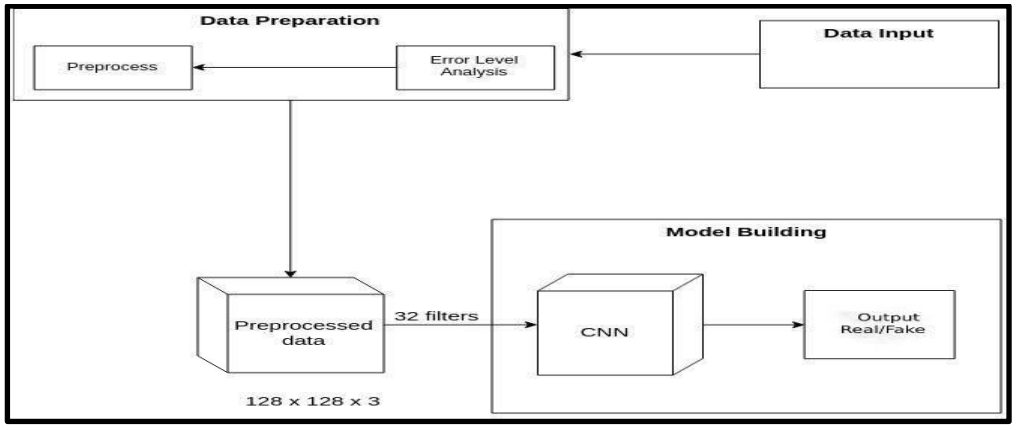


Figure 2.3.1: Block diagram

We noticed that the fake images and their associated natural images show statistical differences, which can be further used as identification

Figure 2.3.2: Real image and Fake image

2.4 Scope of the Problem:

They provide an open dataset of digital images composed of duplicate images, such as those created using different lighting conditions and image algorithms:

1. Content-Aware Completion and Patch match
2. Content-aware healing
3. Clone Stamp
4. Image reconstruction
5. Alpha Matting (for piecemeal)
6. Inpainting (image reconstruction of damaged parts - a special case of copy/paste)

2.5 Challenges:

The report will be about deep learning access to perfect the first stage of the challenge. Everything from data cleanup, preprocessing, conversational neural network architecture to instruction and amends will be stated in detail. Several limitations apply to mine this image forgery detection data [10], i.e. the raw data must be an image with loss compression (for example.jpg), not is it a computer-generated image (CGI).The

participating parties need to identify the images as fake or primitive (never intended). The second step is to identify/localize the fake regions in their duplicate images

There were some challenging tasks to implemented this such as,

1. To choose the appropriate images
2. Correct error level analysis
3. Processed data
4. Neural network make was very difficult

CHAPTER 3
Research Methodology

3.1 Introduction:

Most image files do not contain just one image the image is carrying his information into the metadata. The metadata used include various types of metadata in different image formats, providing more information about the genre of an image along with the camera type, color position information, and application notes, etc. When an image manipulation with photoshop or other editing software then occurs some error. This error is significant for finding a forgery image and we use error level analysis code [6]. Then we use techniques that machine learning with deep learning it's Convolutional Neural Network (CNN) on process images.

3.2 Research Subject and Instrumentation:

We search for fake image detection solution and the following tools, such as error-level analysis (ELA) and conversational neural network (CNN) model, are a method for learning deep learning theory. And software we use python, anaconda, an use open-source library Keras and TensorFlow

3.3 Data Collection Procedure:

Generally, architectural design is divided into two main parts, namely data readiness and model residence. In the early stages, the input data element of the images in the early stage format JPEG, with details as follows: 4000 images with tampered labels and 300 images with real labels. [8] Entered into the stage of data formulation. Error level analysis transforms the data into an image. Then, the error level analysis image will be resized to a 128 x 128 size image.

Figure 3.3.1: A) Input image. B) Convert input data to ELA

Converting raw data to images from error level analysis is a method used to increase the training efficiency of the Convolutional Neural Network model. This efficiency can be achieved because the error level analysis images contain information that is not as excessive as the original image. In addition, the pixels in error level analysis images tend to have a color that is similar or even very contrasting with nearby pixels, so training the Convolutional Neural Network model becomes more efficient. The next step is to normalize by dividing each RGB value by number 255.0 to normalize so that Convolutional Neural Network converges faster (reaching the minimum global value of loss belonging to validation data) because the value of each RGB value only ranges between 0 and 1. if represents 1 than the next process to change the label on data tampered and 0represents 0 real into categorical value.

3.4 Statistical Analysis:

The maximum accuracy of the results acquired from the proposed method is above 80.00%. The accuracy curve and the loss curve can be seen in the picture below.

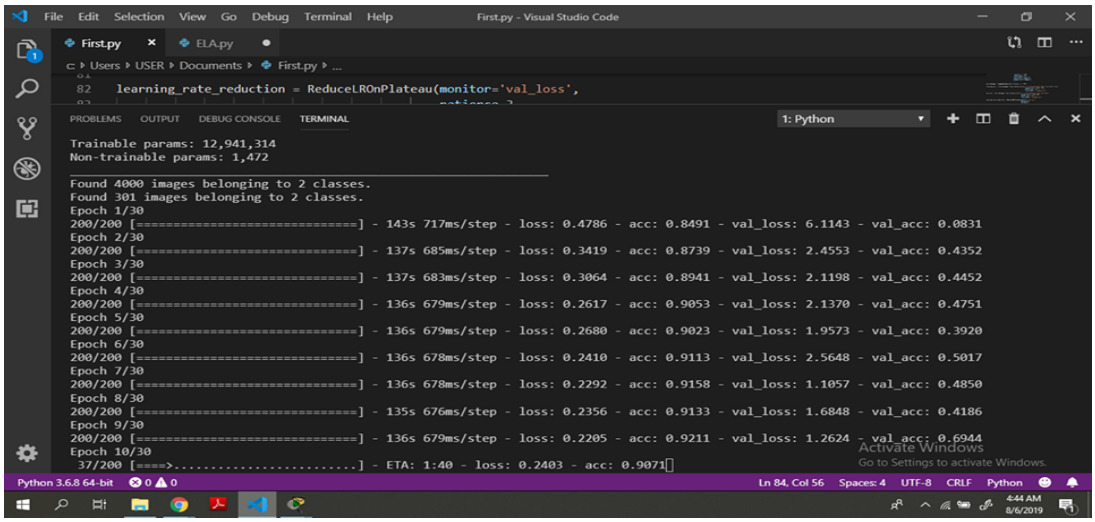


Figure3.4.1 :Accuracy curve and damage curve for training data and test data

3.5 Implementation Requirements:

Error Level Analysis is one of the techniques that use the discover image manipulation by stock images at a particular into the quality and frugal the balance between the level of compression [10]. In general, this technique is done on images that have a loss format. The type of image used in mining this data is JPEG. In JPEG images, compression is done.

independently for every 8x8 pixels in the image. If an image is not manipulated, every 8x8 pixels in the image must have the same error rate [6].

the convolutional neural network is a feedforward-based network, the flow of information is just one direction from input to output. In the convolutional neural network architectures have few layers .such as the convulsive layer and a pooling layer. Thereafter, one or more fully connected layers follow. the convolutional layer is applied as a property extractor that agency of image properties into the input on the convolutional neural networks . Usually, before the fully connected level, there are several convolutional stacks and pooling layers that work to disable more abstract representations of the features.[9]

CHAPTER 4
Experimental Results and Discussion

4.1 Introduction:

One thing we need to keep in mind before experimenting on this project is that when we input an image, the image size becomes smaller. And if the image format is JPEG it will take less time to process it. Moreover, the PNG image that is larger in size than JPEG image[1]. So we have tried to suggest every image get lossy compression size. When we project a JPEG image as input, the image will tell him error level analysis and machine learning and tell us whether the image was real or fake. This will not losses or damage the images. In this way we have been able to keep more than 80% of our accuracy when the machine is taught with a lot of image input. If you input more different types of images machine will be more learn and increase result accuracy.

4.2 Experimental Results:

Model Sequential :

```
Model: "sequential_1"
_________________________________________________________________
Layer (type)                 Output Shape              Param #
=================================================================
conv2d_1 (Conv2D)            (None, 126, 126, 32)      896
_________________________________________________________________
batch_normalization_1 (Batch (None, 126, 126, 32)      128
_________________________________________________________________
max_pooling2d_1 (MaxPooling2 (None, 63, 63, 32)        0
_________________________________________________________________
dropout_1 (Dropout)          (None, 63, 63, 32)        0
_________________________________________________________________
conv2d_2 (Conv2D)            (None, 61, 61, 64)        18496
_________________________________________________________________
batch_normalization_2 (Batch (None, 61, 61, 64)        256
_________________________________________________________________
max_pooling2d_2 (MaxPooling2 (None, 30, 30, 64)        0
_________________________________________________________________
dropout_2 (Dropout)          (None, 30, 30, 64)        0
```

Figure 4.2.1 Model Sequential

In this figure 4.2.1 we showing our Model Sequential. We use open-source library Keras Model.

```
batch_normalization_3 (Batch (None, 28, 28, 128)       512
_________________________________________________________________
max_pooling2d_3 (MaxPooling2 (None, 14, 14, 128)       0
_________________________________________________________________
dropout_3 (Dropout)          (None, 14, 14, 128)       0
_________________________________________________________________
flatten_1 (Flatten)          (None, 25088)             0
_________________________________________________________________
dense_1 (Dense)              (None, 512)               12845568
_________________________________________________________________
batch_normalization_4 (Batch (None, 512)               2048
_________________________________________________________________
dropout_4 (Dropout)          (None, 512)               0
_________________________________________________________________
dense_2 (Dense)              (None, 2)                 1026
=================================================================
Total params: 12,942,786
Trainable params: 12,941,314
Non-trainable params: 1,472
_________________________________________________________________
```

Figure 4.2.2 Model Sequential

In this figure 4.2.2 we showing our Model Sequential. We use open-source library Keras Model.

Experiment 1:

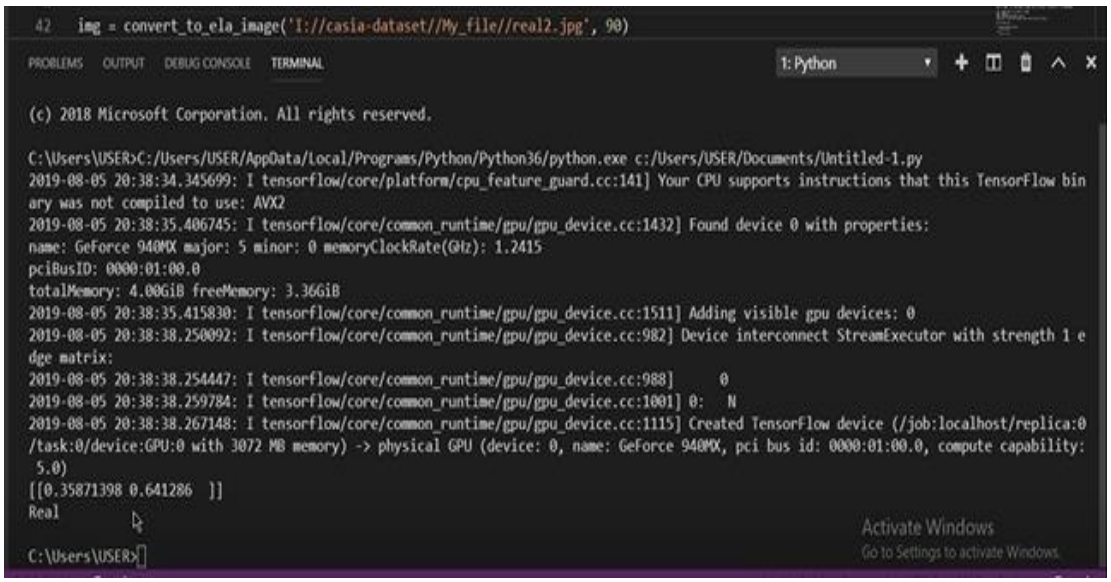


Figure 4.2.3 Experiment result

In this figure we showing our outcome. In figure 4.2.3 show real value because the input data whichever compare to training data that 35% fake value and 65 % real then showing result real.

Experiment 2:

Figure 4.2.4 Experiment result

In this figure we showing our outcome. In figure 4.2.4 show real value because the input data whichever compare to training data that 86.8% fake value and 13.1 % real then showing result real.

4.3 Descriptive Analysis:

In the picture above it can be seen that the best accuracy of the ninth period is found. After the ninth period, the standard of declining legitimacy begins to flatten and eventually increases, which is a sign of overfitting. With the help of this method, the detection method is soon to be stopped for a large number of epochs used during training, when the quality of test accuracy starts to decline or the value of the test decreases starts to increase.

The amount of training age required to achieve convergence is small because the use of the error level analysis-converted image feature makes the training model more efficient and accelerates the transformation of the convolutional neural network model according to the RGB value for each pixel.

In the case of classification administration, the results acquired by the model can be perfectly telling. This suggests that the feature is in the form of an error level analysis image.

4.4 Summary:

First of all, this research project is a fake image detection using python language. This program found a picture forgery or original and shows the result. And there are two main methods used in data error level analysis processing (ELA) [4] and machine learning techniques in the form of Convolutional Neural Network. Error Level Analysis is one of the techniques used to detect image manipulation or edited image. Then error level analysis processed data send Convolutional Neural Network architecture layer. Convolutional Neural Network has several convolutional fully attached layers. When input on Convolutional Neural Network is in from an image data then each pixel can be processed. And lastly, show the result Fake or Real an image. So training is the Convolutional Neural Network model becomes more efficient and more accuracy.

CHAPTER 5
Summary, Conclusion, Recommendation and Implication for Future Research

5.1 Summary of the Study:

Fake image detection is very significant topic in the area of computer look. In this subject we focusing on research that detection the fake images using Error Level Analysis [4] and machine learning. Machine learning neural network are most valuable and popular thesis in the globe. But before machine learning our image needs to be pre-processed and this requires Error Level Analysis. So we started to learn different types of papers such as Image Forensics [6] online for error level analysis. Then we go next step solution of fake image detection with machine learning methodology. We search online and old research paper in this related work but we catch out some research paper and related work. From all these papers and online portals we have new ideas and concepts then help consist of new recherch project.

5.2 Conclusions:

In this study, there are several things that can be concluded from the results of machine learning using Error Level analysis [6] and Convolutional Neural Network.

1. Convolutional neural network uses two convolutional layers, one MaxPooling layer, one fully connected layer, and one output layer with softmax that can achieve above 80% accuracy.
2. The use of error level analysis can increase efficiency and reduce the computational costs of the training process. This can be seen from the reduction in the number of layers from the previous method [7] and the number of epochs needed. In the proposed model, the number of epochs needed to achieve convergence is only 9.

5.3 Recommendations:

There are some limitations to this systematic review and meta-analysis. Then we create sanctions.

1. The image should first be less compressed (e.g. jpg).
2. Input requires a good quality light image.
3. Image is definitely not a computer-generated image (CGI).
4. Archaeological images cannot be used for detention.

If these are recommended think used then the detection system is work to show his current result.

5.4 Implication for Further Study:

In the future, we are Implication of the mobile application to detection fake images. It varies important to implement these actions. fake images create is human. So we have to change our mine we will implement it to day by day.

Reference:

[1] Yuanfang Guo and Xiaochun Cao and Wei Zhang and Rui Wang, "Fake Colorized Image Detection", CarXiv:1801.02768v2 [cs.MM] ,14 Jan 2018,

[2] Muhammed Afsal Villan and, "Fake Image Detection Using Machine Learning ", International Journal of Computer Science and Information Technology & Security (IJCSITS), ISSN: 2249-9555 Vol.7, No.2, Mar-April 2017

[3] Abhishek Kashyap and Rajesh Singh Parmar and Megha Agarwal and Hariom Gupta , "An Evaluation of Digital Image Forgery Detection Approaches" ,arXiv:1703.09968v2 [cs.MM] , 30 Mar 2017

[4] A. C. Popescu and H. Farid, "Exposing Digital Forgeries by Detecting Traces of Re-sampling," IEEE Trans. on Signal Processing, vol. 53,no.2, pp. 758–767, Feb. 2005

[5] T. Ng, S.F. Chang, and Q. Sun, "Blind Detection of Photomontage using Higher Order Statistics", IEEE International Symposium on Circuits and Systems, Canada, May 2004

[6] Photo Forensics: Detect Photoshop Manipulation with Error Level Analysis Diambil dari https://resources.infosecinstitute.com/error-level-analysis-detect-image-manipulation, September 2018.

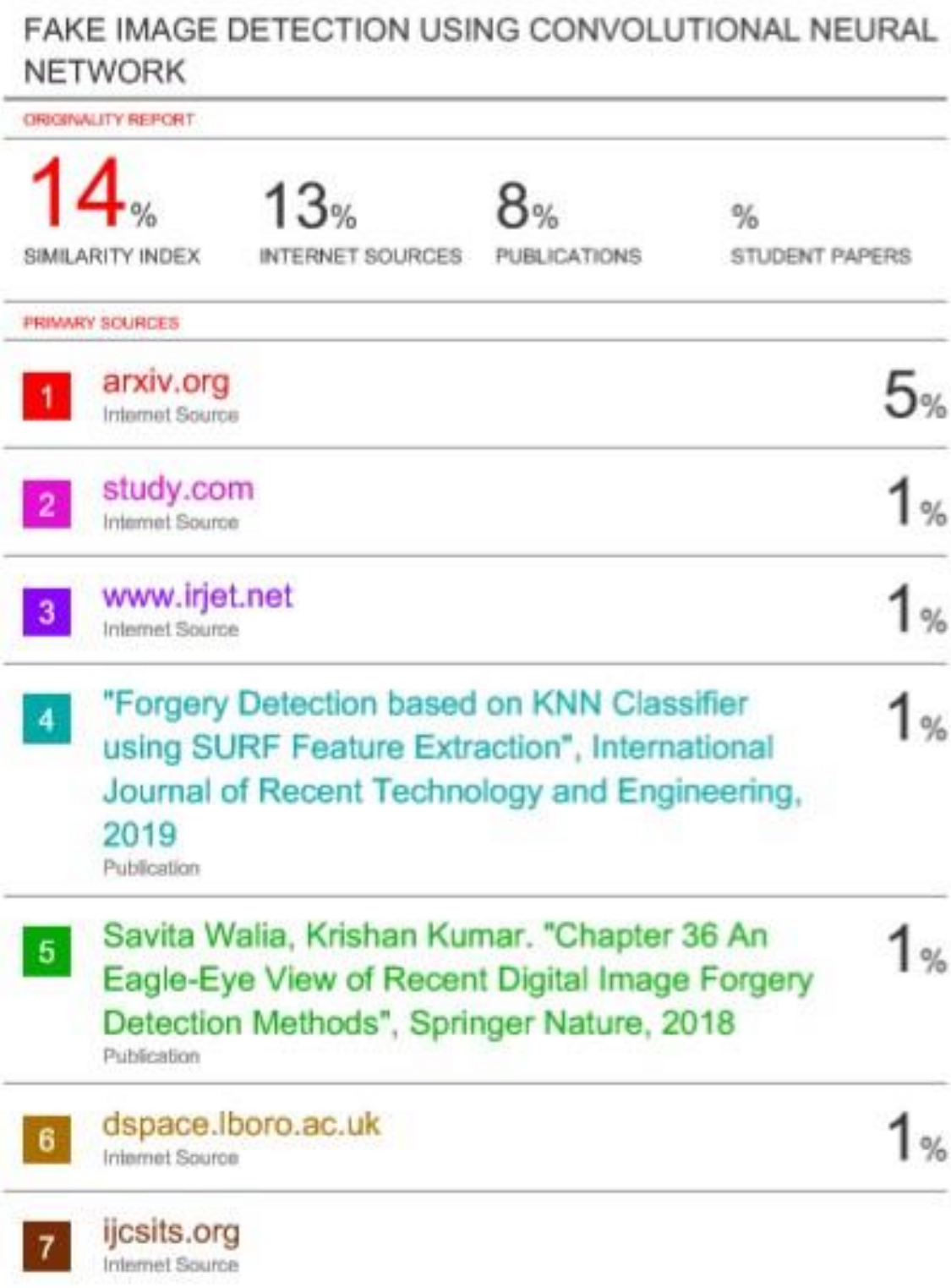
[7] Rao, Yuan, Ni, J., "A Deep Learning Approach to Detection of Splicing and CopyMove Forgeries in Images"IEEE International Workshop on Information Forensics and Security (WIFS)., 2016

[8]Jing Dong, Wei Wang and Tieniu Tan, "CASIA IMAGE TAMPERING DETECTION EVALUATION DATABASE",(CAISA TIDE) V2.0. Diambil dari ,2013[.3]

[9]Rawat, Waseem, Z. Wang, "Deep Convolutional Neural Networks for Image Classification: A Comprehensive Review", Neural Computation 29 (2017), hlm. 2352-2449. [5]

[10]N. Krawetz, "A pictures worth digital image analysis and forensics," Black Hat Briefings, hlm. 1-41, 2007.

[11]S.Murali 1, Govindraj . Chittapur , Prabhakara H. S 1 and Basavaraj S. Anami "COMPARISON AND ANALYSIS OF PHOTO IMAGE FORGERY DETECTION TECHNIQUES" International Journal on Computational Sciences & Applications (IJCSA) Vo2, No.6, December 2012



FAKE IMAGE DETECTION USING CONVOLUTIONAL NEURAL NETWORK

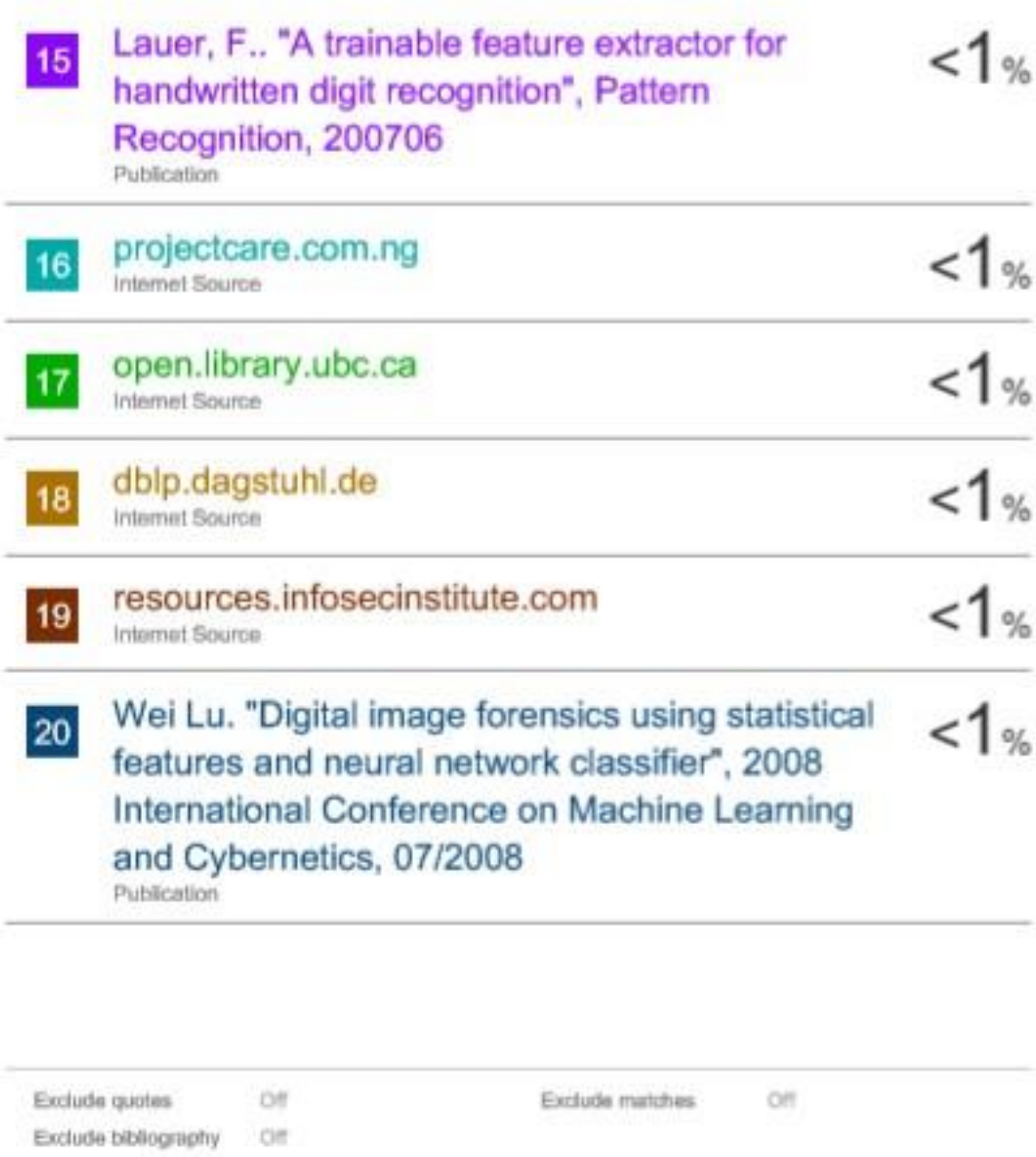
ORIGINALITY REPORT

14%	13%	8%	%
SIMILARITY INDEX	INTERNET SOURCES	PUBLICATIONS	STUDENT PAPERS

PRIMARY SOURCES

1	arxiv.org Internet Source	5%
2	study.com Internet Source	1%
3	www.irjet.net Internet Source	1%
4	"Forgery Detection based on KNN Classifier using SURF Feature Extraction", International Journal of Recent Technology and Engineering, 2019 Publication	1%
5	Savita Walia, Krishan Kumar. "Chapter 36 An Eagle-Eye View of Recent Digital Image Forgery Detection Methods", Springer Nature, 2018 Publication	1%
6	dspace.lboro.ac.uk Internet Source	1%
7	ijcsits.org Internet Source	

		1%
8	www.coursehero.com Internet Source	1%
9	Naveen Ashish, Mi-Youn Brusniak. "Automated Mouse Organ Segmentation", Proceedings of the 2018 ACM International Conference on Bioinformatics, Computational Biology, and Health Informatics - BCB '18, 2018 Publication	<1%
10	Toqeer Mahmood, Tabassam Nawaz, Rehan Ashraf, Mohsin Shah, Zakir Khan, Aun Irtaza, Zahid Mehmood. "A survey on block based copy move image forgery detection techniques", 2015 International Conference on Emerging Technologies (ICET), 2015 Publication	<1%
11	uniprojects.net Internet Source	<1%
12	www.safaribooksonline.com Internet Source	<1%
13	digital-image-forensic.de Internet Source	<1%
14	monstrochan.org Internet Source	<1%



FAKE IMAGE DETECTION USING CONVOLUTIONAL NEURAL NETWORK

GRADEMARK REPORT

FINAL GRADE

/0

GENERAL COMMENTS

Instructor

PAGE 1

PAGE 2

PAGE 3

PAGE 4

PAGE 5

PAGE 6

PAGE 7

PAGE 8

PAGE 9

PAGE 10

PAGE 11

PAGE 12

PAGE 13

PAGE 14

PAGE 15

PAGE 16

PAGE 17

PAGE 18