Web Forum And Social Media: A Model For Automatic Removal Of Fake Media Using Multilayered Neural Networks

Wilson Nwankwo, Kingsley E. Ukhurebor

Abstract: In the last decade, communication through traditional broadcast media has been severely influenced by the growing sophistication in information communication technologies. It is little wonder how mobile smart devices are rapidly emasculating the socioeconomic landscape of traditional broadcasting. In the work place, home, and leisure the trend is same. Social media has stolen the game and is quickly leading to another age, the "machine relationship" age where the usual intimate social relationships are replaced by mobile devices and social networks. Worrisome is the growing spate of fake media. Effects of such fake information propelled by miscreants and social media apologists are far-reaching as it has resulted to crisis in homes, families, relationships, organizations, and the society at large. This paper aims to contribute to the fight against fake media through the use of technology. In this paper we have proposed and developed a model that integrates multilayer neural networks to detecting and eliminating fake pictures and images uploaded to social media networks and web forums.

Index Terms: Social media, Social networks, Neural networks, Web forum, Media authentication

1. INTRODUCTION

In the last decade communication through broadcast media has been severely influenced by the growing sophistication in information communication technologies [1],[2]. It is little wonder how mobile smart devices are influencing the socioeconomic landscape. In the work place, it is no different. At home and leisure the trend is same. Social media has stolen the game and is quickly leading to another age, the "machine relationship" age where the usual intimate social relationships are replaced by mobile devices and social networks. Kai et al., [3] has noted that men and woman including teenagers and children now spend a greater amount of their lives interacting through social media platforms. What is alarming is the geometric rate at which people explore new technologies in this direction. Much attention is paid to information from social media as against the conventional mass and news media respectively. There is no gainsaying that the social media now surpasses newspapers, radio and television put together. The advantages of social media are remarkable in the dissemination of information at an alarming speed owing to their dependent on the Internet. Social media technologies have found applications in education, healthcare delivery, security, etc. [4],[5],[6]. It is instructive to note that the problem is not really the social media technologies themselves but the increasing unethical exploitation of such technologies by miscreants and social media apologists to perpetrate disaffection and other social ills in the society. Many a time, the unsuspecting recipients of such communications are often deceived and cowed to propel such further by rebroadcasting through similar social media channels. To distinguish what is fake from what is real is a herculean task in most spheres of the society especially where pictures and videos are used to authenticate such falsehood.

- *Wilson Nwankwo is an Associate Professor at the Department of Computer Science, Edo University Iyamho, Nigeria. E-mail: nwankwo.wilson@edouniversity.edu.ng*
- *Kingsley Ukhurebor is Research Fellow, WASCAL and Lecturer at the Environmental & Telecoms Unit, Department of Physics, Edo University Iyamho, Nigeria. E-mail: ukhurebor.kingsley@edouniversity.edu.ng*

As noted by [4], these technologies pose great challenges to the society. Some of these challenges are shown in Figure 1.

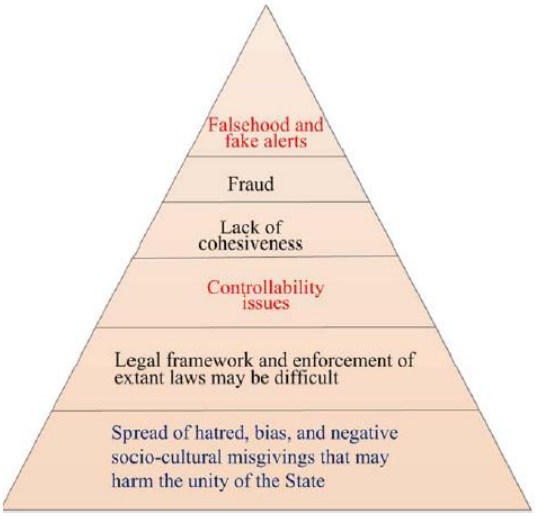

Figure 1: *Problems and challenges posed by social media technologies (Source: Nwankwo et al, 2018)*

As stated by Shao et al., [7] "the immense dispersion of forged news has been identified as a global risk and this so-called news has been alleged to impact elections and governments". Research is ongoing in the fields of sociology, communication, psychology and cognitive science, as well as computing, towards studying various channels and intentions underlying social class propensities on social media and network platforms so as to put up measures that could reverse the negative trend. To buttress the inherent problems of falsehood on social media, Allcott and Gentzkow [8] claims American democracy was repeatedly battered by modifications in media technology.

1.1 Statement of Problem

Due to the growing spate of social attachments of members of the society to social media especially through mobile devices, it has become imperative to design means through which the various information exchanged on such platforms could be checked in a manner not contrary to the

fundamental rights to freedom of information. Homes, marriages, relationships, etc. have been severely battered as a result of fake information transmitted over these media. In many cases, pieces of information are presented to look very authentic through the use of photographs. Pictorial information or infographics are known to worth more than equivalent representation in text and are appreciated faster by the brain hence are very vital in provoking emotions. For instance, in Nigeria, many of such cases abound online where photos of bad roads from a different locality are captured, photo-shopped and tagged to reflect a location in another State. Often times, the intention of the person doing it is to promote disaffection or score cheap political points. To the ordinary user of such media, whatever is relayed by way of graphics are appealing and accepted as the extant reality. Having seen the effect such misguided information has had on the minds of people, miscreants and propagandists are continuously exploiting the magnanimity of the social media especially web forums and mainstream social media networks to infuse more fake media into the Internet. Consequent upon the foregoing, the masses are often deceived to believe and further propagate fake news resulting in social contract crisis and widened distrusts in social relationships.

1.2 Aim and Objectives

The aim of this paper is to explore the use of technology in preventing and fighting crisis that may be overtly or covertly created by activities of miscreants, internet criminals, propagandists, and apologists through the propagation of fake media in social media circles such as web forums, social media apps, and social networks.

The specific objectives are:

a. To evaluate the authenticity of photographs and similar graphics media using the embedded characteristics (metadata) and neural networks.
b. To evolve a model with some intelligent capacity, that could be integrated into social network applications and web forums to automatically and remove offending media when they are uploaded to web forums, and social networks
c. To contribute to government's policy on control of fake media on the internet

2. Review of Related Literature

Social media platforms have become indispensable tools for communication in our daily lives. Different species of such technologies are continuously developed to broaden the spectrum of choice among users who are always ready to explore the functionalities of one social mobile application or the other. A visit to the google playstore, apple store and windows app store would convince anyone on the growth of such applications by the day. These apps as they are called have far-reaching effects on the population. Each may be exploited to achieve beneficial as well as used to cause havoc to others or the society at large. Popular apps include facebook, whatsapp, twitter, Instagram, skype, etc. According to Aditi et al. [9] there were ten thousand three hundred and fifty (10,350) unique tweets containing fake images transmitted through Twitter during Hurricane Sandy disaster. He had performed a characterization analysis, to comprehend the social reputation, temporal and influence patterns for the dispersion of forged images. Eighty-six percent of tweets divulging the forged images were retweets, hence very few were genuine tweets. The results, led to the conclusion that automated techniques can be used in identifying real images from fake images posted on social media. The trend is not different on other social media platforms including web forums. Satya [10] had traced the origin of image processing and advent of the use of image detection to 2001, when an algorithm for face detection was invented by Paul Viola and Michael Jones. Their demo that showed faces being detected in real time on a webcam feed was the most astonishing demonstration of computer vision and its potential at the time. Joao [11] claims the computational analysis of images is demanding as it usually involves processes such as segmentation, extraction of representative features, alignment, matching, motion analysis, tracking, 3D reconstruction, and deformation estimation. To perform each of these processes in a fully automatic, robust and efficient manner is generally demanding. In the computational vision domain, the identification of objects represented in images is typically known as segmentation. According to Lijun [12], digital image processing consists of the conversion of a physical image into a commensurate digital image and the extraction of meaningful information from the digital image by the application of various algorithms. Therefore, digital image processing constitutes image collection, image analysis, and image processing. Raturi [13], had stated that Social Networking is the main epoch of data transmission, as well as data creation in a broad scale. In his work he realized that social networking is the main platform where a great amount of data is created. Forged accounts are increasing by the day and have become a cyber threat, is among the leading contributors of the accelerated growth in the volume of data generated over the Internet. In his paper, he used machine learning to implement improved prediction on the identification of fake accounts based on their posts and status on their social networking feed. His emphasis was on Facebook and Twitter. Marra et al. [14] had stated that with the robust image editing tools available today; it is very simple to create forgeries without leaving visible traces. Partitions between host image and forgery can be hidden; brightness changed, and so on, in a naive form of counter-forensics. For this reason, a handful of modern techniques for counterfeit detection depends on the statistical distribution of micro-patterns, improved through high-level filtration. An investigation was carried out to determine the effectiveness of the proposed strategy as a function of the level of knowledge on the counterfeit detection algorithm. According to Guo [15], image forensics aims to detect the alteration of digital images. Currently, copy-move detection, splicing detection and retouching of image detection are drawing much attention from researchers. However, image editing methods improve as time goes by. One prominent image editing method is colorization, which can colorize grayscale images with realistic colors. Compared to natural images, colorized images; possess statistical diversity in the hue and saturation channels. Hacker [16], claims error level analysis (ELA) identifies areas within an image that are at different compression levels. With JPEG images, the whole picture should be at approximately the same level. If a segment of the image is at a significantly different error level, then it likely indicates a digital alteration. Regions with

uniform coloring, such as a white wall, will definitely have an ELA result (darker color) that is lower than high-contrast edges. ELA identifies and displays differences in the JPEG compression rate [17]. According to Mallick [18] the advances of the medical and biological sciences over recent decades, have made imaging a progressively essential discipline. The prevalence of digital technology has made images an essential part of a number of search areas, from nanotechnology to astronomy. ImageJ holds a special position because it not only is open source, but also runs on any platform. ImageJ can read most of the typical and important formats used in the field of biomedical imaging. According to Rueden et al., [19], ImageJ is an image analysis program extensively implemented in the biological sciences and beyond. Due to its ease of use, and extensible plug-in architecture, ImageJ relish contributions from non-programmers, and professional developers alike. ImageJ has transmogrified from a single-user, single-bench application to a versatile framework of extensible, reusable operations. The second criterion considered is metadata of the image. A parallel module is added to the program which examines the metadata to ascertain the signature of various image editing software. Since it is expensive to execute a neural network, the metadata scrutiny will substantially increase the performance by detecting modification at an early stage.

Justification/Gap Analysis

Obviously, cybercriminals and apologists who share fake images have intention of causing disaffection in the minds of the recipients and users of social media platforms. Their actions are no less than those of other cybercriminals hence a preventive measure needs be taken to curtail the growth of these cybercrimes. In addition it is also helpful to have a system that could be used by forensic analysts to check the originality of some media transmitted from one platform to another should such become an object of legal suits. In either ways, there is a need for a collective effort towards curbing this evil that is almost becoming a culture among social media users. The Nigerian government has in various times frowned against fake news and the negative influence social media platforms are having on its citizens. The efforts of the government is complemented by the idea communicated in this paper. Having surveyed the various studies made in this domain, we believe that putting it all together to show functionalities is a good leap towards realizing the fight against fake media across social media networks.

3. MATERIALS AND METHOD

3.1 Choice of Methodology

The object-oriented analysis and design methodology (OOADM) was employed. The entire system is considered a conglomerate of subsystems with each system having its component parts. The essence of adopting the OOADM is to contain any complexities that may arise. The OOADM offers immense flexibility in system decomposition and re-composition respectively.

3.2 Hardware and Software

The hardware used is a PC (HP Elitebook 820) @ 2.6Ghz Intel core-i7 processor, 16GB RAM, 1TB hard disk with Microsoft Windows 10 ultimate installed. The system is equipped with webcam for capturing images.

The software used include:

a. Netbeans IDE 8.2 with JDK 8 and JavaFX
b. Neuroph Studio Library: a lightweight Java neural network framework to create common neural network architectures. This library has impressively designed, open source Java library with basic classes and flexible Graphics user interface neural network editor to rapidly develop neural network components.
c. Metadata Extractor Library, a lightweight Java library for reading and collecting metadata from image files.
d. ImageJ Library, a Java image processing library for the display, edit, analysis, and processing of images.

3.2 Data Collection

Primary data were gathered through live camera. Five hundred live images were captured and stored. Secondary data consisting of nine thousand (9000) images were extracted from the CAMSIA dataset. Additional five hundred secondary images were gathered from the popular Nigerian web forum (Nairaland) and facebook, for analysis. The entire dataset comprises ten thousand images. 5000 images were original whereas 5000 were fake. Each category is split into training and testing sets respectively; 70% for training and 30% for testing.

3.3 Analysis of the proposed model

Analysis of the proposed model is performed by way of Use Case, Activity and a Sequence diagrams respectively as shown in Figure 1-3.

Fig. 1 *Use Case diagram of the proposed model*

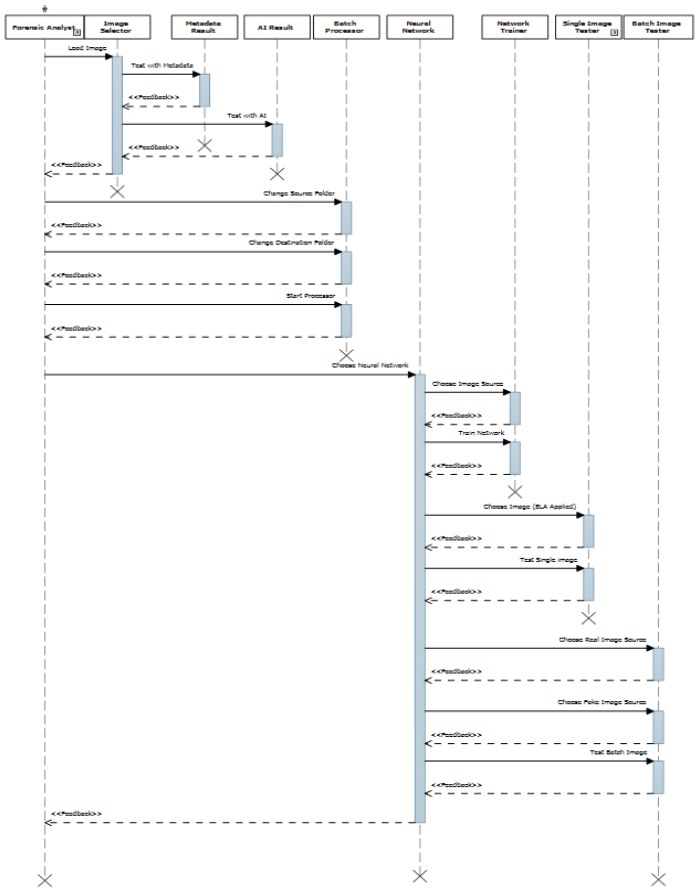

Fig. 3: *Sequence diagram of the proposed model*

3.4 Component Modelling

Component modeling is done by way of a component diagram as shown in Figure 4. The components of the model are data store, metadata analyser, error level analyser, and neural network.

Fig. 4*: Component diagram of proposed model*

3.4.1 Metadata Analysis

The image is scanned for authenticity using a tag searching algorithm that analyzes the image metadata tags for relevant details. The various properties revealed by metadata are presented in Table 1.

Table 1: *Metadata properties*

S/No	Property	Components/functions
1	EXIF Metadata Block	Model of camera
2	Image Size	Height and width of image
3	Timestamps	Time and Date when image was captured
4	Metadata type	Camera or Application-generated
5	Description	Embedded anotations
6	Missing metadata	Camera-specific information
7	Altered metadata	Processed image

The extraction of metadata of images involved the use of the metadata extractor library. Once an image is selected for processing, it is divided into 2 separate stages. The first stage is metadata analysis, while the second involves error level analysis and after extracting metadata, the metadata text is fed into metadata analysis module which implements a tag searching algorithm. For instance, if a tag like “Adobe” is found in the metadata text, then the probability of being modified is increased. Two separate variables are maintained which are called realness and fakeness of image. Each of these variables represents the weight of being a real or fake image. Once a tag is analyzed and the correlating variable is incremented by a specific predefined weight. (Table 2) the final values of realness and fakeness variables are fed into the output stage.

Table 2: *Image tags and weightings*

S/No	Reference	Classification	Preset increment
1	EXIF info	Real	2
2	Adobe	Fake	5
3	Corel	Fake	5
4	Gimp	Fake	5
5	Photoshop	Fake	5
6	Paint	Fake	5
7	Camera tag	Real	2

3.4.2 Error Level Analysis (ELA)

Error level analysis is done using the ImageJ library. ImageJ enables an image to be saved in JPEG format with a specific percentage of compression. The system first saves an image at 100% quality. Then, the same image is converted into 90% quality image using ImageJ. This image is saved as a buffered image and delivered to the neural network for further processing. The difference between these two is found out using difference method. Modifications to image affect the stability of the image, which results in the increase in their minimal error levels. Some areas in the image may show more inconsistencies. The amount of error is limited to the 8x8 cells implemented by the JPEG algorithm and after roughly 64 resaves, there is no change. However, when an image is altered, the 8x8 cells containing the alterations are no longer at the same error level as the rest of the unaltered image. During error level analysis, the image is resaved at a known error rate, such as 95%, and then the difference between the initial and resaved images are evaluated. If there is no change, then the cell has reached its local minima for error at that quality level. However, if there is a large quantity of alterations, then the pixels are not at their local minima and are hence, “original”. The amount of error introduced by each resave will modify the image such that stable areas become unstable. By analyzing the pattern, we can determine which part of the image is probably faked.

3.4.3 Data Store

The data store is an image database which holds the links to all the images stored in a directory. The images were not stored directly in the database to avoid conversion issues which may slow down the speed of the system.

3.4.4 Neural Network construction

The neural network constructed is a multilayer perceptron network which allows multiple input neurons into one input layer, three hidden layers, an activation function and two output neurons from one output layer for detecting fakeness and originality of images. The Neuroph library was used to construct a multilayer perceptron network with back propagation learning rule to minimize error function. The multilayer perceptron neural network has one input layer, 3 hidden layers and 1 output layer. Once the image is selected for evaluation, it is converted to ELA representation from Compression and Error Level Analysis phase. Once ELA is estimated, the image is preprocessed for conversion into 100x100px. After the preprocessing stage, the image is arranged into an array. The array consist of 30,000 integer values constituting 10,000 pixels. Each pixel possess red, green and blue components; hence, 10,000 pixels translated to 30,000.

Training

During training, the array is the input for the multilayer perceptron network (MLPN) with two output neurons. The first and second neuron are utilized for representing fake and real image respectively. If the selected image is a real one, then the real neuron is set to one and fake is set to zero. Otherwise, the real is set to zero and the fake is set to one. We have used momentum back propagation learning rule to adjust the neuron connection weights. This supervised learning rule tries to minimize the error function.

Testing

During testing, the image array is fed to the input neurons and values of output neurons are taken. The hidden layers were separated into three. The sigmoid activation function is used for each hidden layer. 1500 each for fake and real images were used for testing.

4. RESULTS AND DISCUSSION

Figure 5 shows the result of a metadata analysis on a test image.

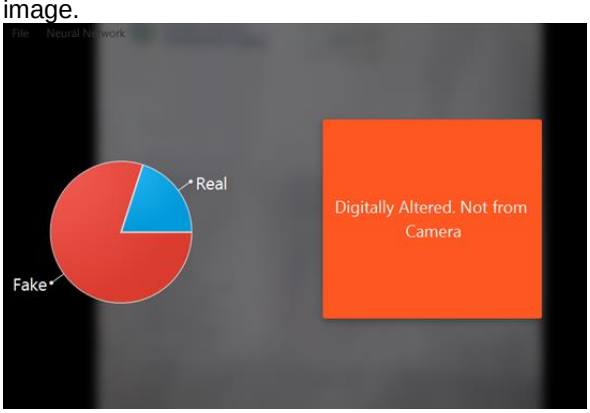

Fig. 5: *Metadata test result*

Figure 6 shows the result interface which provides a window to view the results following a test on image.

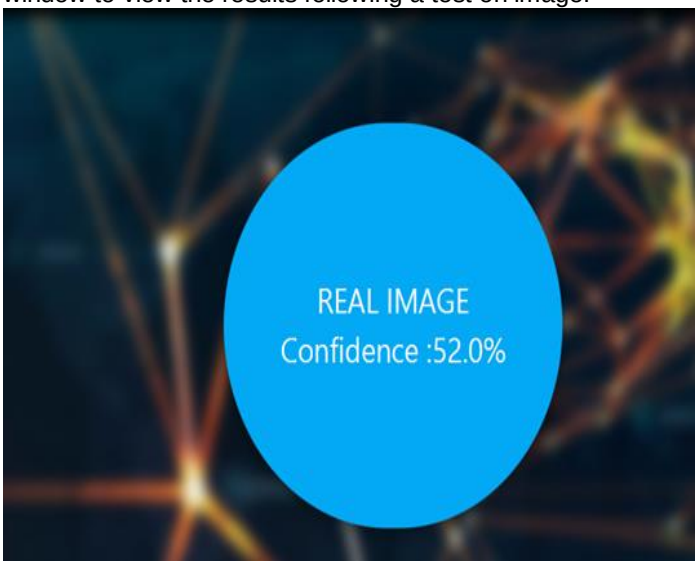

Fig. 6: *Image test result display*

The batch image processing interface enables an analyst to process images using ELA and save those images in batch, based on the directory selected, the destination folder and some other parameters. In other words, there is the

possibility of scanning through all image files in a directory to identify the likely fake and authentic images respectively.

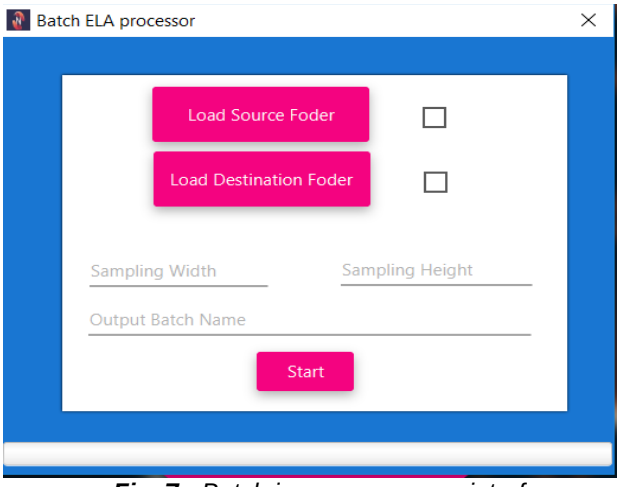

Fig. 7: *Batch image processor interface*

Neural network training interface

The neural network training interface provides options to create or train a new neural network based on various characteristics of the image file.

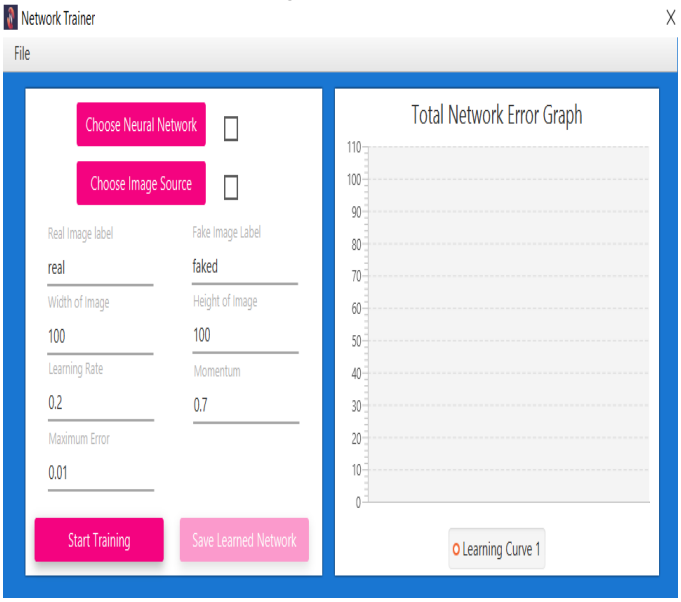

Fig. 8: *Neural network training interface*

Batch image tester interface

The batch image testing interface provides the forensic analyst with the option to perform AI tests for images in batch based on the neural network selected and the real and fake image source.

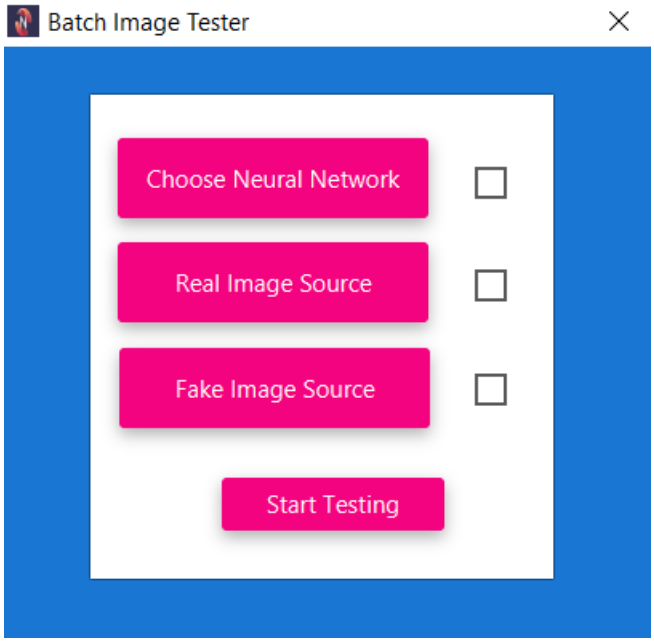

Fig. 9: *Batch image tester interface*

Training and Testing results

For the training, the chosen learning rate and momentum along with achieved efficiency is given in (Table 3). 3500 fake and 3500 real images were utilized for training. During testing, metadata analysis yielded efficient result in original images. It detected modification in all 'photo-shopped' images using small algorithmic processing. Table 3 shows the various neural network configurations and corresponding neural network efficiency. The best efficiency is achieved when momentum is set to 0.7and learning rate set to 0.2.

Table 3: MLPNN *Configuration and efficiency*

Learning rate	Momentum	Epoch	Efficiency(%)
0.01	0.5	500	60
0.05	0.5	500	62
0.1	0.5	500	68
0.2	0.5	500	66
0.1	0.4	500	69
0.1	0.3	500	68
0.1	0.6	500	75
0.1	0.7	500	76
0.2	0.7	500	82
0.2	0.7	1000	83

5. CONCLUSION

Social media has shown its undeniable impact on the socio-economic affairs of humankind as in the last decade social media networks and utilization have been on the increase. Notwithstanding the tremendous achievement recorded by this technology, it has been shown that this novel technology also has its pitfalls. It is interesting to note that many families, social groups, relationships, governance, etc. have also been marred by this technology hence a novel approach to tackle the menace of falsehood spread over the social networks by cyber criminals, cyber hobbyists and apologists With the model proposed in this paper, it is concluded web forums could be designwed in such a way to accommodate an intermediate component that would provide intelligent chacks against fake images or

photographs transmitted by participants in the forum. The implication is that technological tools are very vital in executing the war by various governments[20,21] against fake news especially the circulation of fake multimedia such as photographs and images.

REFERENCES

[1] Nwankwo, W., Ukhurebor, K.E. (2019). Investigating the Performance of Point to Multipoint Microwave Connectivity across Undulating Landscape during Rainfall. In: Journal of the Nigerian Society of Physical Sciences, 1(3), 103-115.

[2] Ukhurebor, K.E., Olayinka, S.A., Nwankwo, W., Alhasan, C. (2019). Evaluation of the Effects of some Weather Variables on UHF and VHF Receivers within Benin City, South-South Region of Nigeria. In: Journal of Physics: IOP Conference Series. 1299, 012052.

[3] Shu, K., Sliva, A., Wang, S., Tang, J., Liu, H. (2017). Fake News Detection on Social Media: A Data Mining Perspective. ACM SIGKDD Explorations Newsletter, 19(1), 22-36.

[4] Nwankwo, W., Ukhurebor, K.E. (2019). Small and Medium-Scale Software Contracts: From Initiation to Commissioning. In: International Journal of Scientific & Technology Research, 8(12), 1554-1563.

[5] Nwankwo,W., Olayinka, A.S., and Ukhurebor, K.E.(2019). The Urban Traffic Congestion Problem in Benin City and the Search for an ICT-improved Solution. International Journal of Science and Technology, 8(12).

[6] Nwankwo,W., Ukaoha,K.C.(2019). Socio-Technical perspectives on Cybersecurity: Nigeria's Cybercrime Legislation in Review; International Journal of Scientific and Technology Research, 8(9), 47-58

[7] Shao, C., Ciampaglia, G.L., Varol, O., Flammini, A., Menczer, F. (2017). The Spread of Fake News by Social Bots. ArXiv, abs/1707.07592.

[8] Allcott, H., Gentzkow, M. (2017). Social Media and Fake News in the 2016 Election.

[9] Gupta, A., Lamba, H., Kumaraguru, P., Joshi, A. (2013). Faking Sandy: Characterizing and Identifying Fake Images on Twitter during Hurricane Sandy. WWW 2013 Companion - Proceedings of the 22nd International Conference on World Wide Web. 729-736.

[10] Mallick, S. (2016). Image Recognition and Object Detection: Part 1. Retrieved from https://www.learnopencv.com.

[11] Joao, T. (2010). Image Processing and Analysis: Applications and Trends. AES-ATEMA International Conference Series - Advances and Trends in Engineering Materials and their Applications.

[12] Sun, L. (2016). Asphalt mix homogeneity. Structural Behavior of Asphalt Pavements, 821-921.

[13] Raturi, R. (2018). Machine Learning Implementation for Identifying Fake Accounts in Social Network. International Journal of Pure and Applied Mathematics, 118(20), 4785-4797.

[14] Marra, F., Poggi, G., Roli, F., Samsone, C., Verdoliva, L. (2015). Counter-Forensics in Machine Learning Based Forgery Detection. Proceedings of SPIE - The International Society for Optical Engineering, 9409.

[15] Guo, Y., Cao, X., Zhang, W., Wang, R. (2018). Fake Colorized Image Detection. 1-13.

[16] Hacker Factor (2012). Error Level Analysis. Retrieved from http://fotoforensics.com/tutorial-ela.php

[17] Hacker Factor (2012). Estimate JPEG Quality. Retrieved from http://fotoforensics.com/tutorial-estq.php

[18] Mallick, S. (2016). Image Recognition and Object Detection: Part 1. Retrieved from https://www.learnopencv.com.

[19] Rueden, C.T., Schindelin, J., Hiner, M.C., DeZonia, B.E., Walter, A.E., Arena, E.T., Eliceiri, K.W. (2017). ImageJ2: ImageJ for the next generation of scientific image data. BMC Bioinformatics, 18(1), 529.

[20] Nwankwo,W. Olayinka,A.S.(2019). Real-time Risk Management and X-ray Cargo Scanning Document Management Prototype for Trade Facilitation, International Journal of Scientific and Technology Research,8(9),93-105.Available at http://www.ijstr.org/final-print/oct2019/Implementing-A-Risk-Management-And-X-ray-Cargo-Scanning-Document-Management-Prototype.pdf

[21] Kifordu,A., Nwankwo,W.,and Ukpere, W.(2019). The Role of Public Private Partnership on the Implementation of National Cybersecurity Policies: A Case of Nigeria; Journal of Advanced Research in Dynamical and Control Systems, 11(8), 1386-1392.