International Research Journal of Engineering and Technology (IRJET) e-ISSN: 2395-0056
Volume: 09 Issue: 06 | June 2022 www.irjet.net p-ISSN: 2395-0072

Image Forgery / Tampering Detection Using Deep Learning and Cloud

Misbah Shaikh [1], Dr. Dipak Patil [2]

[1]*Student, Department of Computer Engineering, Gokhale Education Society's, R. H. Sapat College of Engineering, Management Studies and Research, Nashik, India*
[2] Professor ,*Department of Computer Engineering, Gokhale Education Society's, R. H. Sapat College of Engineering, Management Studies and Research, Nashik, India*

Abstract - *Cybercrime has become more prevalent in recent years. With modern photo editing tools as widely available as ever, it has been demonstrated that creating phony papers is incredibly simple [1]. With the help of this tool, which offers tools for doing so, documents can be scanned and forged in minutes. While photo editing software is convenient and widely available, there are also deft methods for investigating these transformed documents. This study presents a framework for investigating digitally modified documents as well as a way of distinguishing between an original document and a digitally morphing document. we created a web application to detect digitally modified photos. This method has more than 95.0 percent accuracy and has proven to be efficient and useful. Recent work on forgery detection using neural networks has proven to be very effective in detecting image forgery additionally we are using Azure Form recognizer service to read data from documents and verify it on the server, this dual approach makes the system robust and very accurate. Deep Learning methods are capable of extracting complex features in an image, resulting in increased accuracy. In contrast to traditional methods of forgery detection, a deep learning model automatically builds the required features, and as a result, it has emerged as a new area of study in image forgery.*

Key Words: morphed document, Azure form recognizer, CNN, Deep learning, neural network.

1. INTRODUCTION

With the tremendous technological improvement that has boosted the progress of every industry imaginable, one of which is security, it has also become easier to break it [2]. Legal documents can be stolen and faked, but criminal evidence, such as photographs and security footage, can also be easily tampered with. One may believe that checking IDs at the front gate is sufficient for an institution, but they do not comprehend how simple it is for a criminal to obtain false IDs. Even for inexperienced crooks, posing as someone else in public is a simple task. As previously stated, photo editing tools are not only easily available but also incredibly user-friendly. Even if you've never used picture editing software before, you can master fundamental photo editing techniques in a few hours. Photo editing is no longer particularly advanced, and counterfeiting has grown even more difficult to detect. Motives for generating forged photographs might range from financial gain to spreading rumors or making false claims in one's favor.

Deep learning is revolutionizing the field of computer vision, which is already expanding [3]. A CNN is a cutting-edge deep learning technology that learns high-level characteristics from a vast collection of labelled images. Ink analysis in document image processing allows for the determination of ink age and forgery, as well as the identification of the pen or writer. Ink spectral information in hyperspectral document pictures provides essential information about the underlying material, assisting in the identification and discrimination of inks based on their distinct spectral signatures, even though they are the same hue.

1.1 LITERATURE REVIEW

A System Based on Intrinsic Properties for Fraudulent Document Detection, the authors offer an automatic forgery detection system based on the intrinsic features of the document at the character level in this study [4]. This method is based on outlier character detection in a categorical feature space on the one hand, and strictly similar character detection on the other. As a result, a feature set is computed for each character. The character is then classed as authentic or fraudulent depending on the distance between characters of the same class. Local Binary Patterns for Detecting Document Forgery.

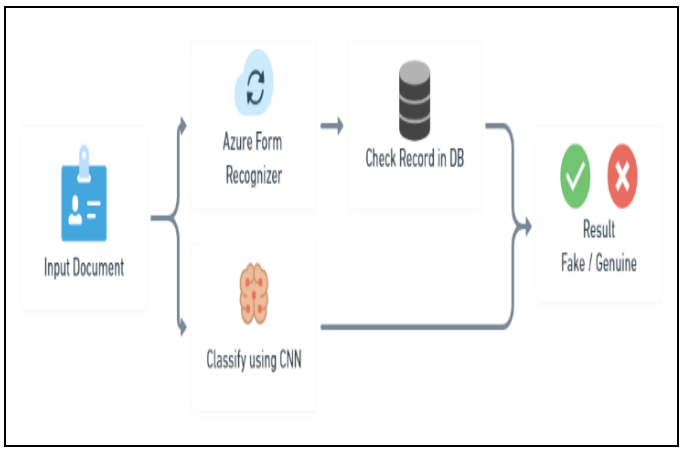
The authors of this study [5] describe a classification-based technique to forgery detection. The author employs uniform Local Binary Patterns (LBP) to capture discriminant textural characteristics seen in fabricated regions. Furthermore, the author combines numerous descriptors from nearby locations to simulate contextual information. The results of patch classification using Support Vector Machines (SVM) reveal that people can recognize numerous types of forgeries in a wide range of document categories.

This study [6] introduces a new method for detecting fake documents. The proposed method is based on network science methodologies for assessing ink spectrums of documents to determine whether they were counterfeited or fabricated. Laser-Induced Breakdown Spectroscopy is utilized to extract the spectrums of the original and

questioned documents for this purpose. The retrieved spectrums are organized to produce a dataset with nodes and edges. The dataset is then utilized to construct a network of spectrums representing both the original and disputed documents. The created network is then displayed and grouped. The detection technique is primarily based on data provided by network clusters (e.g., number of clusters). The results demonstrated that the proposed method was effective at differentiating documents.

1.2 MODELING AND ANALYSIS

Its primary purpose is to create a trustworthy and effective framework for detecting image forgeries. Figure 1 depicts the overall framework of our approach. This structure is classified into two sections.



The first section uses a cloud service (Azure form recognizer) to grab data from uploaded documents and then validate the data with the database. The second section includes all image processing operations. In this section, we also show how crucial it is to focus on the quality of the input data when trying to improve resilience. Convolutional Neural Networks are then used to explain the classification approach for recognizing bogus documents. The following section dives deep into our CNN, allowing us to better understand and develop our framework. Finally, we used a forgery detection application to put our proposed framework to the test. CNN architecture is divided into two parts: Feature-Extraction and Classification. CNN consists following layers – An input layer, convolution layer, pooling layer, and fully-connected layer. in the second phase, we are using the azure form recognizer the fetch data from the image and validate fetched data on the server. this approach makes the proposed system very accurate. As a result, the applicate is more resilient and precise than conventional solutions. Form Recognizer inspects your documents and forms, extracting text and data, mapping field relationships as key-value pairs, and producing structured JSON output. Without a lot of manual work or substantial data science knowledge, you can get accurate answers quickly that are tailored to your specific content.

2. RESULTS AND DISCUSSION

In this section, we presented the dataset we worked with and demonstrated the results and visualization. We evaluated our models using train accuracy, training loss, validation accuracy, and validation loss to derive conclusions.

a) Dataset - We used the CASIA V2.0 Image Tampering Detection Evaluation Database in our research to compare and evaluate tampering detection techniques. This dataset is publicly and easily available for research. It is a dataset for forgery classification that is divided into two classes: tampered/fake and real/original, as illustrated in fig.5. It contains almost 12,000 colored photos, with 7000+ authentic and 5000+ altered images.

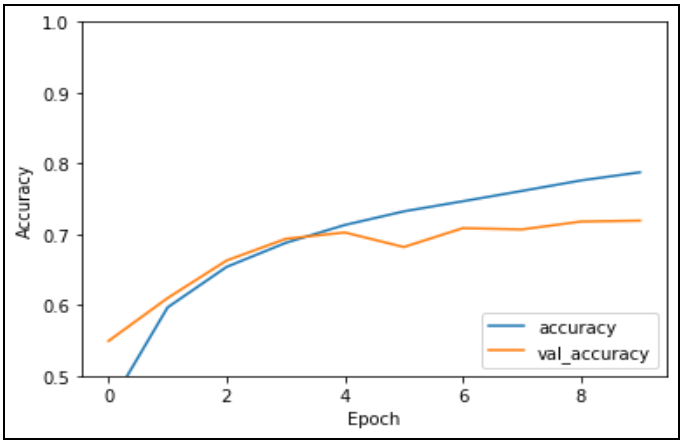


Chart -1: Accuracy of existing CNN Model

b) Result and Analysis - Following the 8 echoes, the accuracy of the present CCN architecture is above 70%, and after CCN, we used a Cloud-based service to grab the data from the document and validate it on the respective server, making our proposed system more than 90% accurate if data/record is available on the server.

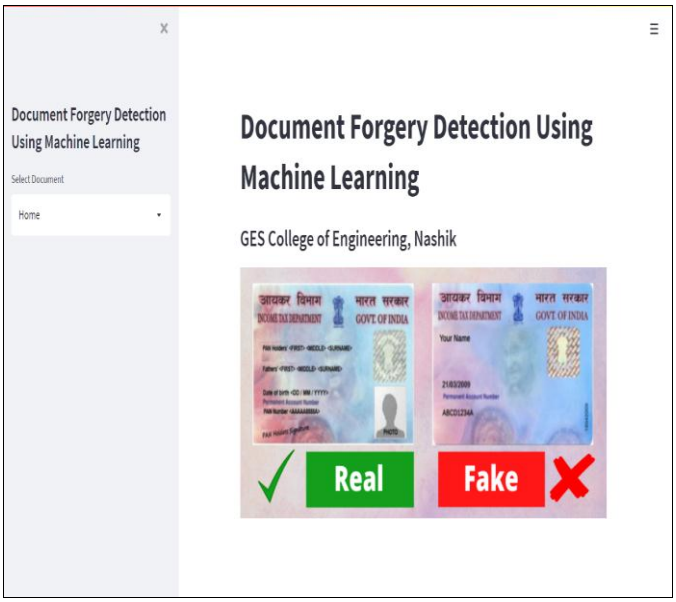


Fig -1: Landing Page of Application

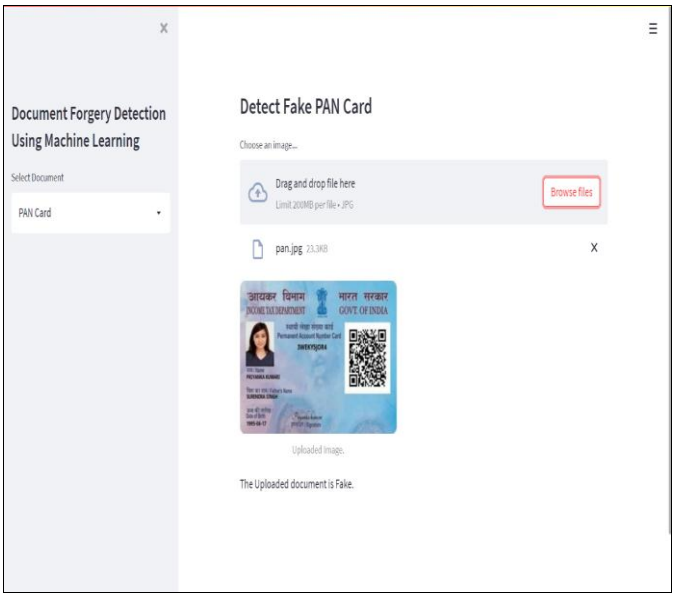


Fig -2: Landing Page of Application

3. CONCLUSION

According to the essential survey, there are two ways of detecting tampering: (i) traditional and (ii) deep learning. Traditional ways for detecting tampering include hand-crafted qualities. According to the results of the survey, traditional approaches do not operate consistently across diverse tampering methods. Instead, Deep Learning-based algorithms have been demonstrated to be capable of automatically learning abstract and complicated attributes required for the identification of tampered regions.

REFERENCES

[1] H. Benhamza, A. Djeffal and A. Cheddad, "Image forgery detection review," 2021 International Conference on Information Systems and Advanced Technologies (ICISAT), 2021, pp. 1-7, doi: 10.1109/ICISAT54145.2021.9678207.

[2] Challenges: An Opportunity to Craft Smarter Responses? CAMINO KAVANAGH, AUGUST 28, 2019 [Online], Available at: https:// carnegieendowment . org / 2019 / 08 / 28 / new -tech -new-threats - and -new - governance – challenges -opportunity -to-craft -smarter -responses-pub-79736.

[3] Khan, Zohaib & Shafait, Faisal & Mian, Ajmal. (2013). Hyperspectral Imaging for Ink Mismatch Detection. Proceedings of the International Conference on Document Analysis and Recognition, ICDAR. 877-881. 10.1109/ICDAR.2013.179.

[4] Bertrand, Romain & Gomez-Krämer, Petra & Terrades, Oriol & Franco, Patrick & Ogier, Jean-Marc. (2013). A System Based on Intrinsic Features for Fraudulent Document Detection. Proceedings of the International Conference on Document Analysis and Recognition, ICDAR. 106-110. 10.1109/ICDAR.2013.29.

[5] F. Cruz, N. Sidère, M. Coustaty, V. P. D'Andecy and J. -M. Ogier, "Local Binary Patterns for Document Forgery Detection," 2017 14th IAPR International Conference on Document Analysis and Recognition (ICDAR), 2017, pp. 1223-1228, doi: 10.1109/ICDAR.2017.202.

[6] A. Amjed, B. Mahmood and K. A. K. AlMukhtar, "Network Science as a Forgery Detection Tool in Digital Forensics," 2021 IEEE International Conference on Communication, Networks and Satellite (COMNETSAT), 2021, pp. 200-205, doi: 10.1109/COMNETSAT53002.2021.9530776.

[7] Malvika Ranjan, Manasi RajivWeginwar, Neha Joshi and Prof. A.B. Ingole Detection and Classification of Leaf Disease Using Artificial Neural Network, 2019.

[8] Jaiswal, Garima & Sharma, Arun & Yadav, Sumit. (2022). Deep feature extraction for document forgery detection with convolutional autoencoders. Computers & Electrical Engineering. 99. 107770. 10.1016/j.compeleceng.2022.107770.

[9] Bibi, Maryam & Hamid, Anmol & Moetesum, Momina & Siddiqi, Imran. (2019). Document Forgery Detection using Printer Source Identification—A Text-Independent Approach. 7-12. 10.1109/ICDARW.2019.70134.

[10] Khan, Muhammad & Yousaf, Adeel & Abbas, Asad & Khurshid, Khurram. (2018). Deep Learning for Automated Forgery Detection in Hyperspectral Document Images. Journal of Electronic Imaging. 27. 053001. 10.1117/1.JEI.27.5.053001.

[11] Ahmed, Amr & Shafait, Faisal. (2014). Forgery Detection Based on Intrinsic Document Contents. Proceedings - 11th IAPR International Workshop on Document Analysis Systems, DAS 2014. 252-256. 10.1109/DAS.2014.26.

[12] How to build a real-time live dashboard with Streamlit, By AbdulMajedRaja RS, Posted in Community, April 21 2022[Online] Available at: https://blog.streamlit.io/how-to-build-a-real-time-live-dashboard-with-streamlit/.

[13] A Beginner's Guide to matplotlib for Data Visualization and Exploration in Python, Aniruddha Bhandari — February 28, 2020 [Online] Available at: https://www.analyticsvidhya.com / blog / 2020 / 02/ beginner -guide -matplotlib -data – visualization - exploration - python/.

[14] Srivastava, Sahima & Rastogi, Vrinda & Jaiswal, Garima & Sharma, Arun. (2022). Hyperspectral Imaging in Document Forgery. 10.1007/978-981-16-6289-8_11.

[15] Khan, Muhammad & Yousaf, Adeel & Abbas, Asad & Khurshid, Khurram. (2018). Deep Learning for Automated Forgery Detection in Hyperspectral Document Images. Journal of Electronic Imaging. 27. 053001. 10.1117/1.JEI.27.5.053001.

[16] Bashir, Alsadig & Fadlalla, Yahia. (2017). Techniques in Detecting Forgery in Identity Documents.