e-ISSN:2582-7219

INTERNATIONAL JOURNAL OF MULTIDISCIPLINARY RESEARCH IN SCIENCE, ENGINEERING AND TECHNOLOGY

Volume 6, Issue 4, April 2023

INTERNATIONAL STANDARD SERIAL NUMBER INDIA

Impact Factor: 7.54

International Journal Of Multidisciplinary Research In Science, Engineering and Technology (IJMRSET)

| ISSN: 2582-7219 | www.ijmrset.com | Impact Factor: 7.54| Monthly Peer Reviewed & Referred Journal |

| Volume 6, Issue 4, April 2023 |

| DOI:10.15680/IJMRSET.2023.0604035 |

Advanced Detection of Document Tampering Using StructuralSimilarity Index and Image Analysis Techniques

Prof. Divya Pandey[1], Prof. Zeba Vishwakarma[2], Prof. Mallika Dwivedi[3], Jatin pasi[4], Shambhavi Pandey[5]

Department of Computer Science and Engineering, Baderia Global Institute of Engineering and Management, Jabalpur, MP, India[1, 2, 3,4,5]

ABSTRACT: This research focuses on the application of the Structural Similarity Index (SSIM) technique for detecting tampering in various identity documents, such as PAN cards, Aadhar cards, and voter IDs. The SSIM method is employed to assess the structural similarity between the original and provided document images. Additionally, grayscale conversion and thresholding techniques are utilized to analyze shapes and contours, further aiding in the identification of tampered areas. Experimental results show that a low SSIM score indicates potential tampering in the provided image. Visualizations, including contour overlays, difference maps, andthreshold-based comparisons, enhance the clarity of differences between original and tampereddocument images.

KEYWORDS: Document Tampering Detection, Structural Similarity Index (SSIM), Image Analysis, Grayscale Conversion, Contour Detection, Fraud Detection

I. INTRODUCTION

In today's digital age, the verification of identity documents is crucial for ensuring authenticityand preventing fraud. The integrity of these documents can be compromised through various means, including tampering and forgery. This research addresses the challenge of detecting tampered identity documents using advanced image processing techniques.

The Structural Similarity Index (SSIM) is a widely recognized method for quantifying the similarity between two images based on their structures, rather than just pixel values. By computing SSIM scores between an original document and a user-provided image, this study aims to identify alterations or discrepancies indicative of tampering.Furthermore, grayscale conversion followed by thresholding allows for a detailed analysis of shapes and contours within the images. Differences in these structural elements can provide crucial insights into the authenticity of the documents. Visual representations such as contour overlays, difference maps, and threshold-based comparisons enhance the interpretability of these findings.

The implications of this research extend to various organizations that rely on accurate identity verification processes. By implementing the proposed method, organizations can enhance their ability to detect fraudulent identity documents, thereby bolstering security and trust in verification processes.
This paper discusses the methodology, experimental results, and implications of using SSIM and image analysis techniques for document tampering detection, contributing to the advancement of fraud detection mechanisms in identity verification systems.

II. LITERATURE SURVEY

I-A. *Credit Card Fraud Detection Using Artificial Neural Networks*

Asha RB et al., in their research on credit card fraud detection using artificial neural networks, explored various techniques including data mining, machine learning, and algorithmic methods. However, their results were not satisfactory, indicating the need for more effective and efficient algorithms to detect fraud in credit card transactions.

| ISSN: 2582-7219 | www.ijmrset.com | Impact Factor: 7.54| Monthly Peer Reviewed & Referred Journal |

| Volume 6, Issue 4, April 2023 |

| DOI:10.15680/IJMRSET.2023.0604035 |

I-B. ***Credit Card Fraud Detection Using Random Forest Algorithm***

In the study conducted by M. Suresh Kumar et al., titled "Credit Card Fraud Detection Using Random Forest Algorithm," the authors highlighted the increasing prevalence of credit card fraud both online and offline. Offline fraud typically involves the physical theft of real cards, whereas online fraud can be executed with virtual cards. Fraudsters target sensitive information such as credit card numbers, bank details, and personal data to execute unauthorized transactions. This has become a major concern for banks and financial institutions, as numerous fraudulent transactions can result in significant data loss, posing a challenge for detection. The authors examined various models to detect fraudulent transactions based on transaction behavior, utilizing both supervised and unsupervised learning algorithms. Techniques like Cluster Analysis, Support Vector Machines, and Naive Bayes Classification have been employed to assess the accuracy of detecting fraudulent activities. Their study specifically employed the Random Forest Algorithm to evaluate the accuracy of identifying fraudulent transactions.

I-C. ***Fraud Detection Using Machine Learning and Deep Learning***

Pradheepan Raghavan and Neamat El Gayar, in their paper "Fraud Detection Using Machine Learning and Deep Learning," proposed that well-trained neural networks can identify distinct relationships across entire datasets. Their research involved comparing various machine learning and deep learning approaches using three datasets: European, Australian, and German. The study used an ensemble of the top three models for each dataset. Through empirical analysis, the research presented findings on the comparative performance of several machine learning and deep learning models in detecting fraud.

III. PROPOSED METHODOLOGY

II-A. ***Data Collection***

For this study, various types of identity documents, such as PAN cards, Aadhar cards, and voterIDs, were collected. Both genuine and tampered versions of these documents were included tocreate a comprehensive dataset.

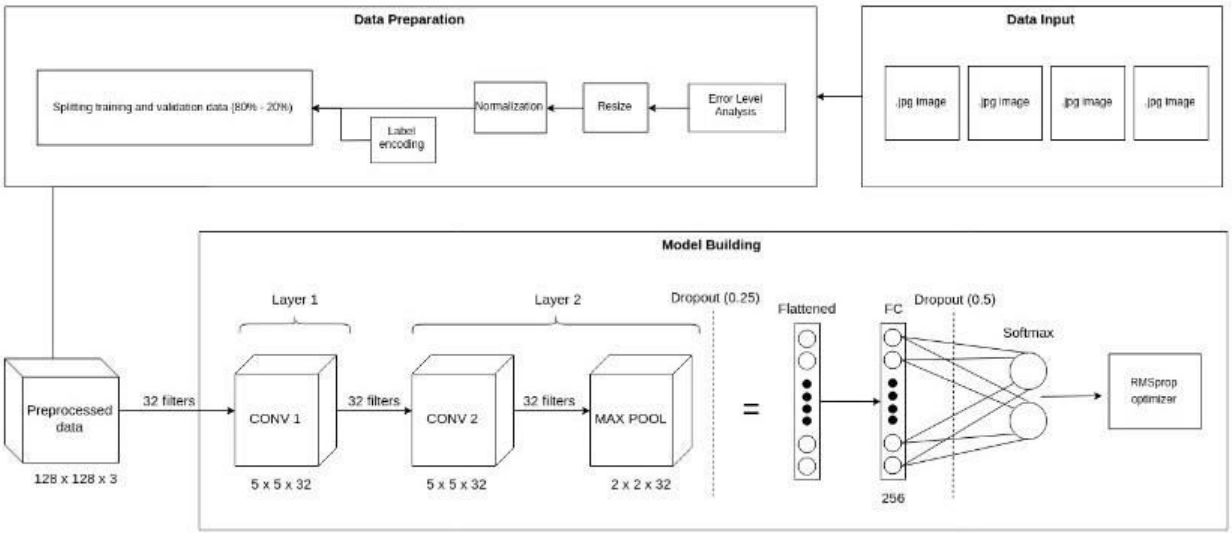

Figure 3.1 Flow diagram of proposed methodology

II-B. ***Preprocessing***

1. *Grayscale Conversion*: Each document image was converted to grayscale to simplifythe analysis and focus on structural features rather than color information.

| ISSN: 2582-7219 | www.ijmrset.com | Impact Factor: 7.54| Monthly Peer Reviewed & Referred Journal |

| Volume 6, Issue 4, April 2023 |

| DOI:10.15680/IJMRSET.2023.0604035 |

2. *Thresholding*: Grayscale images were then binarized using a global threshold value. This step helps in distinguishing the foreground (text and important features) from thebackground.

II-C. *Structural Similarity Index (SSIM) Calculation*

1. *SSIM Overview*: The SSIM index measures the similarity between two images basedon luminance, contrast, and structure. It is particularly effective in identifying subtle changes in structural information.
2. *SSIM Implementation*: The SSIM index was computed between the original document image and the provided image. An SSIM score close to 1 indicates high similarity, while a lower score suggests potential tampering.

II-D. *Shape and Contour Analysis*

1. *Contour Detection*: Contours were detected on the thresholded images to analyze shapes and structural integrity. This helps in identifying any irregularities or distortionsin the document layout.
2. *Thresholding and Contours*: Based on the thresholded images, contours were extracted and compared. Significant deviations in contour shapes and positions can indicate tampering.

II-E. *Visualization*

1. *Difference Maps*: Difference maps were generated to visually highlight discrepancies between the original and provided images.
2. *Contour Overlays:* Contours from the original and provided images were overlaid tovisually compare the structural integrity.
3. *Threshold Comparisons*: Visual comparisons of the threshold images helped inidentifying any missing or altered features.

IV. EXPERIMENTATION RESULTS

Figure 4.1 Original and Tampered Image of a PAN Card

| ISSN: 2582-7219 | www.ijmrset.com | Impact Factor: 7.54| Monthly Peer Reviewed & Referred Journal |

| Volume 6, Issue 4, April 2023 |

| DOI:10.15680/IJMRSET.2023.0604035 |

Model: "sequential_4"

Layer (type)	Output Shape	Param #
conv2d_16 (Conv2D)	(None, 98, 98, 16)	448
conv2d_17 (Conv2D)	(None, 48, 48, 32)	4640
conv2d_18 (Conv2D)	(None, 46, 46, 64)	18496
conv2d_19 (Conv2D)	(None, 22, 22, 8)	4616
dropout_4 (Dropout)	(None, 22, 22, 8)	0
flatten_4 (Flatten)	(None, 3872)	0
dense_12 (Dense)	(None, 50)	193650
dense_13 (Dense)	(None, 30)	1530
dense_14 (Dense)	(None, 2)	62

Total params: 223,442
Trainable params: 223,442
Non-trainable params: 0

Figure 4.2 Model summary of the network architecture

Figure 4.3 Original and Tampered Image with Contour

| ISSN: 2582-7219 | www.ijmrset.com | Impact Factor: 7.54| Monthly Peer Reviewed & Referred Journal |

| Volume 6, Issue 4, April 2023 |

| DOI:10.15680/IJMRSET.2023.0604035 |

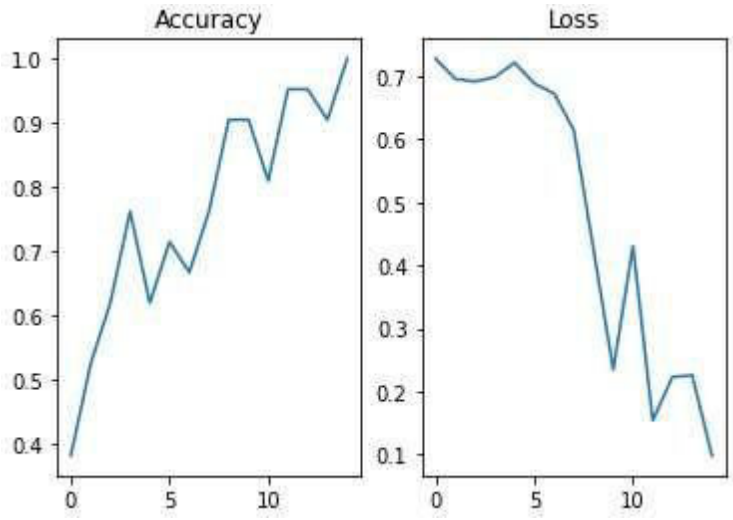

Figure 4.4 Accuracy and Loss Metrics

SSIM Scores

The experimental results showed that documents suspected of tampering exhibited significantly lower SSIM scores compared to genuine documents. For instance, an SSIM score of approximately 31.2% indicated tampering in the provided document image.

Contour and Shape Analysis

Contour analysis revealed notable differences in the shapes and positions of key features in tampered documents. The threshold-based comparisons highlighted areas where structural integrity was compromised, further confirming the presence of tampering.

V. CONCLUSION

In this study, we successfully applied the Structural Similarity Index (SSIM) and image analysis techniques to detect tampering in various identity documents. By evaluating structural similarities and performing detailed shape and contour analysis, we demonstrated an effectiveapproach to identifying fraudulent documents.

Key Findings:

1. *Effectiveness of SSIM:* The SSIM technique proved to be a reliable indicator ofdocument integrity, with low scores correlating strongly with tampered images.
2. *Contour Analysis*: Detailed contour and shape analysis provided additional evidenceof tampering, making the detection process more robust.

Practical Implications:

The methods developed in this research have significant implications for organizations that require reliable identity verification processes. By incorporating these techniques, organizations can enhance their ability to detect and prevent the use of fraudulent identity documents, thereby improving security and trust in their verification systems.

VI. FUTURE SCOPE

Further research can explore the integration of advanced machine learning techniques to enhance the accuracy and efficiency of document tampering detection. Additionally, expanding the dataset to include more types of identity documents and tampering scenarios can improve the generalizability and robustness of the proposed methods.

| ISSN: 2582-7219 | www.ijmrset.com | Impact Factor: 7.54| Monthly Peer Reviewed & Referred Journal |

| Volume 6, Issue 4, April 2023 |

| DOI:10.15680/IJMRSET.2023.0604035 |

REFERENCES

[1] Asha RB, Suresh Kumar KR," Credit card fraud detection using artificial neural network",pp. 35– 41, 2021, doi: https://doi.org/10.1016/j.gltp.2021.01.006

[2] M.Suresh Kumar, V.Soundarya, S.Kavitha, E.S. Keerthika, E.Aswini, "Credit Card Fraud Detection Using Random Forest Algorithm," 2019, doi:https://doi.org/10.1109/ICCCT2.2019.8824930

[3] Pradheepan Raghavan, Neamat El Gayar," Fraud Detection using Machine Learning and Deep Learning," December 2019, doi: https://doi.org/10.1109/ICCIKE47802.2019.9004231

[4] Badal Soni, Pradip K. Das, Dalton Meitei Thounaojam. CMFD: a detailed review of block based and key feature-based techniques in image copy-move forgery detection, 2018. IET Image Processing 12:2, pages 167-178

[5] Francisco Cruz, Nicolas Sidere, Mickael Coustaty, Vincent Poulain ¨ D'Andecy, and Jean-Marc Ogier. Local binary patterns for document forgery detection. In Document Analysis and Recognition (ICDAR), 2017 14th IAPR International Conference on, volume 1, pages 1223– 1228. IEEE, 2017

[6] Y. Sahin, E. Duman," Detecting Credit Card Fraud by ANN and Logistic Regression", 2011, doi: https://doi.org/10.1109/INISTA.2011.5946108

[7] He, Zhiwei, et al. "A new automatic extraction method of container identity codes." IEEE Transactions on intelligent transportation systems 6.1 (2005): 72-78

[8] S. Shang, N. Memon, and X. Kong, "Detecting documents forged by printing and copying,"EURASIP Journal on Advances in Signal Processing, vol. 2014, no. 1, p. 140, 2014.

[9] Ulutas, G., Muzaffer, G.: 'A new copy move forgery detection method resistant to object removal with uniform background forgery', Math. Probl. Eng., 2016, 2016, pp. 1–19

[10] Bashar, M., Noda, K., Ohnishi, N., et al.: 'Exploring duplicated regions in natural images', IEEE Trans. Image Process., 2016, 99, pp. 1–40

Impact Factor
7.54

INTERNATIONAL STANDARD SERIAL NUMBER INDIA

INTERNATIONAL JOURNAL OF MULTIDISCIPLINARY RESEARCH IN SCIENCE, ENGINEERING AND TECHNOLOGY

| Mobile No: +91-6381907438 | Whatsapp: +91-6381907438 | ijmrset@gmail.com |

www.ijmrset.com