IJIRSET
International Journal of Innovative Research in
SCIENCE | ENGINEERING | TECHNOLOGY

e-ISSN: 2319-8753 | p-ISSN: 2347-6710

INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN SCIENCE | ENGINEERING | TECHNOLOGY

Volume 12, Issue 5, May 2023

ISSN INTERNATIONAL STANDARD SERIAL NUMBER INDIA

Impact Factor: 8.423

9940 572 462 | 6381 907 438 | ijirset@gmail.com | www.ijirset.com

International Journal of Innovative Research in Science, Engineering and Technology (IJIRSET)

|e-ISSN: 2319-8753, p-ISSN: 2347-6710| www.ijirset.com | Impact Factor: 8.423| A Monthly Peer Reviewed & Referred Journal |

|| Volume 12, Issue 5, May 2023 ||

| DOI:10.15680/IJIRSET.2023.1205298 |

Deep Learning Based Forgery Attack on Document Images

Dr. Prasanna Lakshmi G S[1], Atul Adarsh[2], Shruti Dipti[3], Shruti Srijan[4] and Suman N[5]

Guide, Associate Professor, Dept. of CSE, Nagarjuna College of Engineering and Technology, Bengaluru, India[1]

U.G. Student, Dept. of CSE, Nagarjuna College of Engineering and Technology, Bengaluru, India[2,3,4,5]

ABSTRACT: With the ongoing popularization of online services, digital document images have been used in various applications. Meanwhile, there have emerged some deep learning-based text editing algorithms which alter the textual information of an image in an end-to-endfashion. In this work, we present a low-cost document forgery algorithm by the existing deep learning-based technologies to edit practical document images. To achieve this goal, the limitations of existing text editing algorithmstowards complicated characters and complex backgroundare addressed by a set of network design strategies. First, the unnecessary confusion in the supervision data is avoided by disentangling the textual and background information in the source images. Second, to capture the structure of some complicated components, the text skeleton is provided as auxiliary information and the continuity in texture is considered explicitly in the loss function. Third, the forgery traces induced by the text editing operation are mitigated by some post-processing operations which consider the distortions from the print and scan channel. Quantitative comparisons of the proposed method and the exiting approach have shown theadvantages of our design by reducing the about 2/3 reconstruction error measured in MSE, improving reconstruction quality measured in PSNR and in SSIM by4 dB and 0.21, respectively. Qualitative experiments have confirmed that the reconstruction results of the proposed method are visually better than the existing approach in both complicated characters and complex texture. More importantly, we have demonstrated the performance of theproposed document forgery algorithm under a practical scenario where an attacker is able to alter the textual information in an identity document using only one sample in the target domain. The forged-and recaptured samples created by the proposed text editing attack and recapturing operation have successfully fooled some existing document authentication systems.

I. INTRODUCTION

The rapid advancements in deep learning techniques have revolutionized various fields, including image processing and computer vision. However, with these advancements, new challenges have emerged, particularly in the area of document security and forgery detection. Document forgery has become a pressing concern due to the increasing sophistication of counterfeiters and the potential risks associated with fraudulent documents.

Deep learning, a subfield of machine learning, has shown remarkable capabilities in analyzing and understanding complex patterns within data. By leveraging the power of deep neural networks, researchers and practitioners have developed innovative approaches to detect and prevent document forgery. This introduction focuses on deep learning-based forgery attacks on document images, where adversaries exploit deep learning techniques to create convincing counterfeit documents.

Forgery attacks on document images involve malicious actors attempting to alter or fabricate the content of documents, such as identity cards, passports, financial documents, or legal contracts. Traditionally, forgery detection relied on manual examination by experts, which was time-consuming and error prone. However, deep learning-based approaches have brought automated and efficient solutions to this domain.

Paper is organized as follows. Section II describes Forgery detection using Recurrent Neural Networks. The flow diagram represents the step of the algorithm. After detection of Image, how ELA generates using an ELA technique that is given in Section III. Section IV presents experimental results showing results of images tested. Finally, Section V presents conclusion.

|e-ISSN: 2319-8753, p-ISSN: 2347-6710| www.ijirset.com | Impact Factor: 8.423| A Monthly Peer Reviewed & Referred Journal |

|| Volume 12, Issue 5, May 2023 ||

| DOI:10.15680/IJIRSET.2023.1205298 |

II. RELATED WORK

Forgery Detection in Document Images Using Recurrent Neural Networks" by Singh et al. (2019):
The authors present a forgery detection approach based on recurrent neural networks (RNNs). They leverage the sequential nature of document images to capture the temporal dependencies between different parts of the document. The RNN architecture is trained on a dataset consisting of genuine and forged document images. The proposed method achieves promising results in detecting forged regions and differentiating them from authentic content.

Adversarial Learning for Document Forgery Attack" by Li et al. (2021):
This work explores the use of generative adversarial networks (GANs) for document forgery attacks. The authors propose a GAN-based framework that consists of a generator network and a discriminator network. The generator learns to generate convincing forged document images, while the discriminator aims to distinguish between genuine and forged documents. The study highlights the vulnerability of document verification systems to GAN -based forgery attacks and suggests countermeasures to enhance security.

These related works collectively demonstrate the active research efforts in the field of deep learning -based forgery attacks on document images. They emphasize the importance of developing robust forgery detection methods to counter the evolving threats posed by advanced forgery techniques.

III. METHODOLOGY

The basic image recognition neural network called convolution neural network (CNN), has been used. Images as input is given and significant features from an image is taken out. It is a multi-layered neural network, hence at each layer it extracts certain features. As move deeper into network, it can identify even complex features. The CNN architecture is divided into 2 parts: Feature-Extraction and Classification. Dataset is split into training and testing data set with a ratio of 80% and 20% respectively is passed through the CNN, CNN_ELA combination and CNN_SHARPEN_ELA combination, separately and then compared. ELA as stated is a procedure where the tamperedimages is stored at a specific quality level and then compute the difference from the compression level. ELA is a JPEG, lossy and irreversible compression algorithm for imageforensics detection. To approximate the JPEG quality, quantization process adopted to develop this algorithm. The image is divided into 8x8 blocks and recompressed at 95%of error rate. If the image is unmodified, unaltered then every grid should have approximately same quality rate and else, if not then there will be a difference in the level of grids. Thus, existence of inconsistent in grid indicates the image modification. The manipulation in JPEG images can be identified by the error rate produced by ELA technique. To further increase the accuracy for the model to detect image forgery, examined the combination of ELA and sharpen filter. Where both the pre-processing stagesplays part as shown in results. The process of enhancement of pixel contrast between bright and dark regions to bring out features clearly is the called the sharpening. In our research, the Pillow-python image processing library has been used.

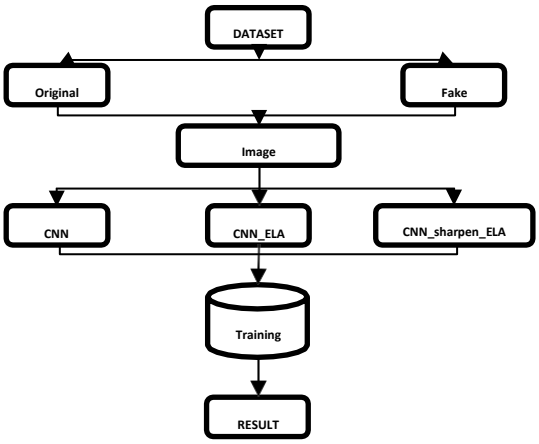


Fig. 1. Flow cha

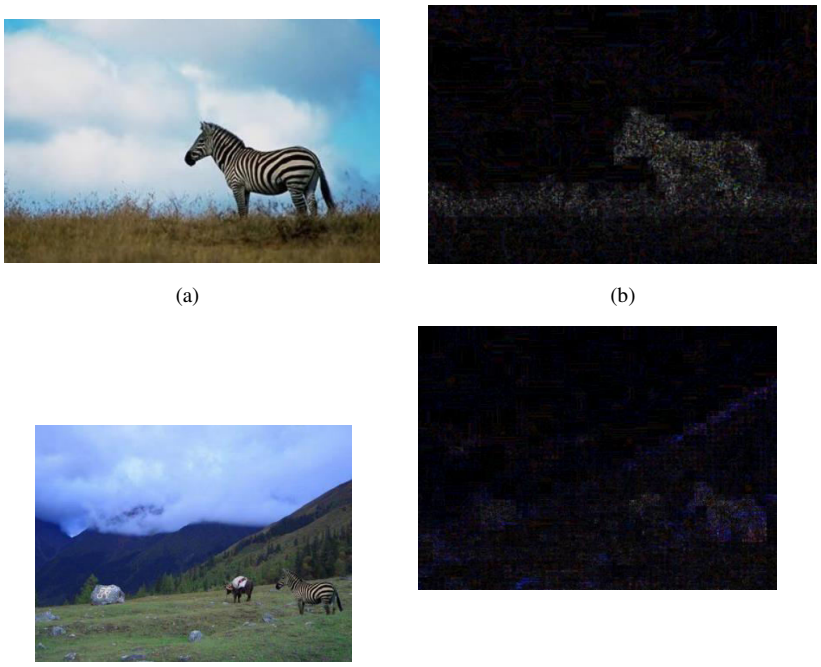
|e-ISSN: 2319-8753, p-ISSN: 2347-6710| www.ijirset.com | Impact Factor: 8.423| A Monthly Peer Reviewed & Referred Journal |

|| Volume 12, Issue 5, May 2023 ||

| DOI:10.15680/IJIRSET.2023.1205298 |

IV. EXPERIMENTAL RESULTS

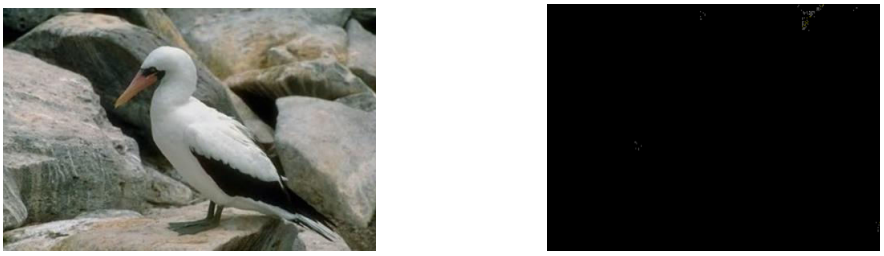
Figures shows the results of Fake Image detection from an image and inpainting by using CNN.



(a) (b)

d. (e)

Fig. 2. Fake image Detection (a) Original image (b) Image after applying ELA (c) Original image (d) Image after applying ELA



(a) (b)

Fig. 3. Fake image Detection (a) Original image (b) Image after applying ELA.

VI. CONCLUSION

In this work, the feasibility of employing deep learning- based technology to edit document images with complicated characters and complex background is studied. To achievegood editing performance, we address the limitations of existing text editing algorithms towards complicated characters and complex background by avoiding unnecessary confusions in different components of the source images, constructing texture continuity loss and providing auxiliary skeleton information (by the fine fusion and skeleton supervision. Comparisons with the existing text editing approach confirm the importance of our contributions. Moreover, we propose to mitigate the visual artifacts of text editing operation by some post-processing (color pre-compensation and inverse halftoning) considering the print-and-scan

|e-ISSN: 2319-8753, p-ISSN: 2347-6710| www.ijirset.com | Impact Factor: 8.423| A Monthly Peer Reviewed & Referred Journal |

|| Volume 12, Issue 5, May 2023 ||

| DOI:10.15680/IJIRSET.2023.1205298 |

channel. Experimental results show thatthe consistency among different regions in a document image are maintained by these post-processing. We also demonstratethe document forgery performance under a practical scenario where an attacker generates an identity document with only one sample in the target domain. Finally, the recapturing attackis employ cover the forensic traces of the text editing andpost-processing operations. The forge-and-recapture samples by the proposed attack have successfully fooled some state- of-the-art document authentication systems. From the study ofthis work, we conclude that the advancement of deep learning-based text editing techniques has already introduced significant security risk to our document images.

REFERENCES

[1] L. Wu, C. Zhang, J. Liu, J. Han, J. Liu, E. Ding, and X. Bai, "Editing Text in the Wild," in *Proceedings of the 27th ACM International Conference on Multimedia*, 2019, pp. 1500–1508.

[2] P. Roy, S. Bhattacharya, S. Ghosh, and U. Pal, "STEFANN: Scenetext editor using font adaptive neural network," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2020, pp. 13 228–13 237.

[3] Q. Yang, J. Huang, and W. Lin, "SwapText: Image based Texts Transferin Scenes," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2020, pp. 14 700–14 709.

[4] P. Korshunov and S. Marcel, "Deepfakes: A New Threat to Face Recognition? Assessment and Detection," *arXiv preprint arXiv:1812.08685*, 2018.

[5] S. Shang, X. Kong, and X. You, "Document Forgery Detection using Distortion Mutation of Geometric Parameters in Characters," *Journal of Electronic Imaging*, vol. 24, no. 2, p. 023008, 2015.

[6] S. Abramova *et al.*, "Detecting Copy–move Forgeries in Scanned Text Documents," *Electronic Imaging*, vol. 2016, no. 8, pp. 1–9, 2016.

[7] T. Thongkamwitoon, H. Muammar, and P.-L. Dragotti, "An Image Recapture Detection Algorithm based on Learning Dictionaries of Edge Profiles," *IEEE Transactions on Information Forensics and Security*, vol. 10, no. 5, pp. 953–968, 2015.

[8] S. Agarwal, W. Fan, and H. Farid, "A Diverse Large-scale Dataset for Evaluating Rebroadcast Attacks," in *2018 IEEE International Confer- ence on Acoustics, Speech and Signal Processing (ICASSP)*. IEEE, 2018, pp. 1997–2001.

INNO SPACE
SJIF Scientific Journal Impact Factor

Impact Factor: 8.423

doi cross ref

ISSN INTERNATIONAL STANDARD SERIAL NUMBER INDIA

निस्केयर NISCAIR

INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN SCIENCE | ENGINEERING | TECHNOLOGY

9940 572 462 6381 907 438 ijirset@gmail.com

www.ijirset.com

Scan to save the contact details