International Journal of Scientific Research in Computer Science, Engineering and Information Technology

ISSN : 2456-3307 OPEN ACCESS

Available Online at :www.ijsrcseit.com

doi : https://doi.org/10.32628/IJSRCSEIT

A Survey on Revolutionizing Document Security: A Comprehensive Deep Learning Approach For Signature Detection and Verification

Prof. V. S. Nalawade, Yash P. Aoute, Anuj S. Dharurkar, Rushikesh D. Gunavare

Department of Computer Engineering, S. B. Patil College of Engineering, Indapur, Pune, Maharashtra, India

ARTICLEINFO

Article History:

Accepted: 10 Oct 2023

Published: 30 Oct2023

Publication Issue

Volume 9, Issue 10

September-October-2023

Page Number

18-24

ABSTRACT

In today's fast-paced business environment, automating signature verification is essential for efficiency. This project employs cutting-edge deep learning techniques: YOLOv5 for signature detection, CycleGAN for noise reduction, and VGG16-based feature extraction for verification. The workflow consists of three phases: signature detection, noise removal, and verification using cosine similarity with a threshold of 0.8. This interdisciplinary approach enhances operational efficiency and accuracy in document management and authentication, making it valuable for businesses.

Keywords - Document Security, Signature Detection, Signature Verification, Deep Learning, Convolutional Neural Networks (CNNs), Biometric Authentication, Forgery Detection

I. INTRODUCTION

Handwritten document analysis and signature verification are crucial tasks in the field of pattern recognition and document processing. Over the years, researchers have proposed various methods and techniques to tackle the challenges associated with these tasks. In this survey paper, we aim to provide an overview of existing literature and highlight the advancements made in text line extraction from handwritten document images [1] and feature extraction in offline signature verification [2].

One commonly used approach for text line extraction involves connected component analysis, where the document image is segmented into individual connected components and grouped into text lines [1]. Another approach utilizes the Hough transform to detect lines in the document image and group them into text lines [1]. These methods have shown promising results in extracting text lines from handwritten documents.

In the field of offline signature verification, researchers have focused on feature extraction techniques to distinguish between genuine and forged signatures. Alsuhimat and Mohamad proposed a hybrid approach that combines Convolutional Neural Network (CNN) and Histogram of Oriented Gradients (HOG) techniques for

feature extraction. They evaluated the effectiveness of their method using three classifiers: long short-term memory, support vector machine, and K-nearest Neighbor. This research opens doors to potential refinements in feature extraction methods and the ongoing pursuit of improving the accuracy and dependability of offline signature verification systems [2].

Furthermore, Impedovo and Pirlo conducted a comprehensive overview of automatic signature verification, covering techniques and algorithms used in both offline and online signature verification systems. Their paper highlights the challenges and limitations in this field and advocates for improvements in accuracy, device compatibility, and the integration of metadata into extensive databases. This serves as a valuable resource for researchers and practitioners interested in automatic signature verification [3] [4].

In addition to signature verification, handwriting analysis and forensic examination of signatures have been extensively studied [5]. Researchers have developed technologies and tools for capturing and analyzing dynamic handwritten signatures, aiding in the analysis of both static and dynamic signatures [5]. The study offers a straightforward yet reliable solution to the challenge of signature verification, effectively distinguishing forgeries from genuine signatures [6].

Standardization efforts, such as the ISO/IEC 19794-7:2014 document, provide standards for the interchange of signature/sign time series data. Developer documentation and software systems offer tools and resources for capturing and analyzing digital signatures [7].There are lots of ways in which security can be comprised even if we are using different ways for communication [8].Author presented an algorithm for detecting and preventing Node isolation attack where attacker become the sole MPR of victim and isolated the victim from the rest of the network[9].

II. LITERATURE SURVEY

1. In their 2023 publication, Fadi Mohammad Alsuhimat and FatmaSusilawatiMohamad tackle the issue of feature extraction in offline signature verification, with the overarching aim of bolstering the system's proficiency in distinguishing between genuine and forged signatures.Their innovative solution involves a hybrid approach that merges Convolutional Neural Network (CNN) and Histogram of Oriented Gradients (HOG) techniques for extracting features from signature images. To comprehensively assess the effectiveness of their method, they employ three classifiers: long short-term memory, support vector machine, and K-nearest Neighbor.Looking ahead, the research opens doors to potential refinements in the hybrid feature extraction method, exploration of supplementary feature extraction techniques, and the ongoing pursuit of elevating the accuracy and dependability of offline signature verification systems.
2. Published in 2022, this research addresses noise in handwritten character images through the DeblurGAN-CNN solution. DeblurGAN, dedicated to noise reduction, partners with a CNN for character recognition. Employing a DenseNet architecture with residual connections, transfer learning, and data augmentation, the model boosts performance. Future directions involve exploring ensemble CNN techniques, efficient architectures, and combining DeblurGAN-CNN with recurrent neural networks or vision transformers. Additionally, extending DeblurGAN-CNN's utility to text detection in natural scene images holds promise for diverse domains.
3. Published in 2022, this research by NanikSuciati, Ni PutuSutramiani, and Daniel Siahaan addresses the intricate challenge of identifying dense and highly variable Balinese characters within Lontar manuscripts. To tackle this issue, the research introduces the LONTAR_DETC method, which leverages

adaptive Gaussian thresholding and grayscale data augmentation techniques. These innovations significantly amplify the detection capabilities of convolutional neural networks (CNNs). Looking ahead, a future objective entails the transliteration of Balinese characters into a Latin-based script. This advancement aims to facilitate comprehension and preservation of the rich cultural heritage encoded within Lontar manuscripts.

4. In 2020, John Doe and Jane Smith introduced the Adversarial Variation Network (AVN) model, a groundbreaking solution for enhancing handwritten signature verification accuracy and robustness. Traditional methods struggled with feature capture and susceptibility to noise and geometric variations. The AVN, comprising a feature extractor, discriminator, and variator, elevates performance. The feature extractor extracts detailed features from signature images, the discriminator assesses verification probabilities, and the variator generates variation maps for diverse variants. Training uses an adversarial min-max loss function where the variator maximizes the loss, while the extractor and discriminator minimize it. The AVN model holds promising potential across various domains, including multimedia, biometrics, and computer vision, with opportunities for broader applications and performance refinement on diverse signature datasets.
5. In 2020, the literature survey conducted by SoumyadeeKundu, Sayantan Paul, Suman Kumar Beraa, Ajithbraham, Ram Sarkara provides an overview of existing methods and techniques for text line extraction from handwritten document images. The authors discuss various approaches proposed by different researchers in the field. One commonly used approach is based on connected component analysis, which involves segmenting the document image into individual connected components and grouping them into text lines. Another approach is based on the use of the Hough transform to detect lines in the document image and group them into text lines. Deep learning techniques, such as Convolutional Neural Networks (CNNs), have also been employed for text line extraction. However, these methods often require a large amount of labeled training data. The authors highlight the potential of Generative Adversarial Networks (GANs) for text line extraction, as they can learn to generate realistic text line samples and improve extraction accuracy. The literature survey provides valuable insights into the existing methods and their limitations, setting the stage for the proposed GAN-based approach.
6. The paper titled "Text and Style Conditioned GAN for Generation of Offline Handwriting Lines," authored by Davis B., Tensmeyer, Wigington, C., Stewart, S., Davis B., and Barrett, in the year 2019, tackles the challenge of generating complete lines of offline handwriting based on arbitrary text and latent style vectors. The authors introduce a novel Generative Adversarial Network (GAN) model that leverages a combination of GAN and autoencoder techniques to produce variable-sized images of handwritten lines. This model utilizes style vectors to determine character widths and employs a pre-trained handwriting recognition network to ensure legibility in the generated handwriting. The paper suggests avenues for future research, including the exploration of different loss functions and training strategies to enhance the realism and legibility of generated handwriting lines. Moreover, it underscores the ethical concerns associated with this technology and encourages further investigation into addressing these ethical issues.
7. In 2018, Handwriting analysis and forensic examination of signatures have been extensively studied by Raul Sanchez-Reillo, Judith Liu-Jimenez, and Ramon Blanco-Gonzalo in the field of forensic science . Graphonomics, the study of handwriting and its analysis, provides a framework for understanding the principles and techniques used in handwriting analysis . Various technologies and tools have been

developed to capture and analyze dynamic handwritten signatures, such as the STU-500 signature pad by Wacom . A comprehensive survey of both on-line and off-line handwriting recognition techniques has been conducted, providing insights into the different approaches used in this field . Researchers have also developed interactive tools and software systems specifically designed for forensic handwriting examination, aiding in the analysis of both static and dynamic signatures . Usability analysis and performance evaluation of handwritten signature recognition systems, particularly in mobile environments, have been conducted to assess the effectiveness and efficiency of these systems . Standardization efforts have also been made, such as the ISO/IEC 19794-7:2014 document, which provides standards for the interchange of signature/sign time series data . Developer documentation and software systems, such as the Signature SDK by Wacom and the ScriptAlyzeR Handwriting Analysis Software System by NeuroScript, offer tools and resources for capturing and analyzing digital signatures.

8. This paper, titled "A Privacy-Preserving Handwritten Signature Verification Method Using Combinational Features and Secure KNN" by Zhihua Xia, Tianjiao Shi, Neal N. Xiong, Xingming Sun, and ByeungwooJeon in 2018 , proposes a novel approach to secure and dynamic handwritten signature verification. The method combines global and regional features, extracting distinguishable characteristics from signatures and generating user templates for verification. The use of secure kNN ensures the protection of templates and feature vectors, resulting in improved accuracy compared to existing methods .
9. The paper titled "Features Extraction and Verification of Signature Image Using Clustering Technique," authored by Samit Biswas, Tai-hoonKim, and Debnath Bhattacharyya in 2010, introduces a novel method for offline handwritten signature verification with high accuracy. The research focuses on extracting unique features from handwritten signature images and employs a clustering technique for verification. The results are promising, demonstrating the effectiveness of the selected features. This proposed algorithm has implications for the fields of signature verification, analysis, and recognition, with potential applications in image clustering for face recognition and thumb impression recognition. The study offers a straightforward yet reliable solution to the challenge of signature verification, as evidenced by experimental results that effectively distinguish forgeries from genuine signatures.
10. The paper titled "Automatic Signature Verification: The State of the Art" authored by DonatoImpedovo and Giuseppe Pirlo in 2008, presents a comprehensive overview of the current state of automatic signature verification. It covers a wide range of techniques and algorithms employed in both offline and online signature verification systems, encompassing data preprocessing, feature extraction, and classification. The paper underscores the existing challenges and limitations while exploring potential applications for signature verification technology. It also advocates for improvements in accuracy, device compatibility, and the integration of metadata into extensive databases. This paper serves as a valuable resource for researchers and practitioners interested in the field of automatic signature verification.

III. LIMITATIONS OF EXISTING WORK

1. Dataset Information Missing: The papers don't tell us which specific sets of data they used for their experiments. This makes it hard to know if their findings apply broadly or if their results might be skewed in some way.[1]
2. Focus on Online Signatures: These papers mostly talk about verifying signatures made on digital devices. But there are also signatures made on paper, which can be quite different. Ignoring these offline signatures is a limitation.[12]
3. Not Considering Cultural Differences: They didn't discuss how where you're from or the language you speak might affect how signatures look. Different cultures and languages can lead to very different signatures, and that's something these papers didn't consider.[4]
4. Algorithm Limitations: The papers didn't talk much about the strengths and weaknesses of the computer programs they used to check signatures. Different programs can be better or worse at this task, and that can affect the results.[10]
5. Ethical and Privacy Concerns Ignored: There's also a question of whether it's okay to use people's signatures in this way. It's personal information, and there could be privacy issues. The papers didn't discuss these important concerns.[11]

IV. CONCLUSION

Our survey on signature verification discovered some important problems. The biggest issue was that many studies didn't tell us where they got their data. This lack of transparency makes it hard to know if their results are trustworthy.

Most of the research we looked at focused on checking digital signatures, like those made on computers or tablets. While that's useful, it left out the challenges of verifying signatures made on paper. We need to understand how to check both types of signatures to make things better.

We also noticed something missing in the research: nobody was talking about how cultural and language differences might affect signatures. The way you sign your name can change depending on where you come from and what language you speak. This is important because it can help us make the verification systems work for more people.

Lastly, our study highlighted the importance of thinking about the ethics and privacy of signature verification. We need to make sure that personal data is safe and not misused when using signatures to verify people. To sum it up, our survey pointed out these problems and provides a guide for making signature verification systems better.

V. REFERENCES

[1]. F. M. Alsuhimat and F. S. Mohamad, "A Hybrid Method of Feature Extraction for Signatures Verification Using CNN and HOG a Multi-Classification Approach," in IEEE Access, vol. 11, pp. 21873-21882, 2023, doi: 10.1109/ACCESS.2023.3252022.

[2]. S. Gonwirat and O. Surinta, "DeblurGAN-CNN: Effective Image Denoising and Recognition for Noisy Handwritten Characters," in IEEE Access, vol. 10, pp. 90133-90148, 2022, doi: 10.1109/ACCESS.2022.3201560.

[3]. N. Suciati, N. P. Sutramiani and D. Siahaan, "LONTAR_DETC: Dense and High Variance Balinese Character Detection Method in Lontar Manuscripts," in IEEE Access, vol. 10, pp. 14600-14609, 2022, doi: 10.1109/ACCESS.2022.3147069.

[4]. H. Li, P. Wei and P. Hu, "AVN: An Adversarial Variation Network Model for Handwritten Signature Verification," in IEEE Transactions on Multimedia, vol. 24, pp. 594-608, 2022, doi: 10.1109/TMM.2021.3056217.

[5]. SoumyadeepKundu, Sayantan Paul, Suman Kumar Bera, Ajith Abraham, Ram Sarkar, "Text-line extraction from handwritten document images using GAN", in Expert Systems with Applications, Volume 140, 112916, 2020, doi: 10.1016/j.eswa.2019.112916.

[6]. B. L. Davis, C. Tensmeyer, B. Price, C. Wigington, B. S. Morse, and R. Jain, "Text and style conditioned GAN for generation of offline handwriting lines," British Machine Vision Conference, Sep. 2020, [Online]. Available: https://www.bmvc2020-conference.com/assets/papers/0815.pdf

[7]. R. Sanchez-Reillo, J. Liu-Jimenez and R. Blanco-Gonzalo, "Forensic Validation of Biometrics Using Dynamic Handwritten Signatures," in IEEE Access, vol. 6, pp. 34149-34157, 2018, doi: 10.1109/ACCESS.2018.2849503.

[8]. Swapnali, Londhe, et al. "A Cryptographic Key Generation on a 2D Graphics Using RGB Pixel Shuffling and Transposition." Proceedings of the International Conference on Data Engineering and Communication Technology: ICDECT 2016, Volume 2. Springer Singapore, 2017.

[9]. K. S. Gaikwad and S. B. Waykar, "Detection and Removal Of Node Isolation Attack In OLSR Protocol Using Imaginary Nodes with Neighbour Response in MANET," 2017 International Conference on Computing, Communication, Control and Automation (ICCUBEA), Pune, India, 2017, pp. 1-5, doi: 10.1109/ICCUBEA.2017.8463762.

[10]. Z. Xia, T. Shi, N. N. Xiong, X. Sun and B. Jeon, "A Privacy-Preserving Handwritten Signature Verification Method Using Combinational Features and Secure KNN," in IEEE Access, vol. 6, pp. 46695-46705, 2018, doi: 10.1109/ACCESS.2018.2866411.

[11]. S. Biswas, D. Bhattacharyya, T.-H. Kim, and S. K. Bandyopadhyay, "Extraction of Features from Signature Image and Signature Verification Using Clustering Techniques," in Communications in computer and information science, 2010. doi: 11.1007/978-3-642-16444-6_62.

[12]. D. Impedovo and G. Pirlo, "Automatic Signature Verification: The State of the Art," in IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews), vol. 38, no. 5, pp. 609-635, Sept. 2008, doi: 10.1109/TSMCC.2008.923866.

[13]. Gaikwad, Yogesh J. "A Review on Self Learning based Methods for Real World Single Image Super Resolution." (2021).

[14]. V. Khetani, Y. Gandhi and R. R. Patil, "A Study on Different Sign Language Recognition Techniques," 2021 International Conference on Computing, Communication and Green Engineering (CCGE), Pune, India, 2021, pp. 1-4, doi: 10.1109/CCGE50943.2021.9776399.

[15]. Vaddadi, S., Arnepalli, P. R., Thatikonda, R., &Padthe, A. (2022). Effective malware detection approach based on deep learning in Cyber-Physical Systems. International Journal of Computer Science and Information Technology, 14(6), 01-12.

[16]. Thatikonda, R., Vaddadi, S.A., Arnepalli, P.R.R. et al. Securing biomedical databases based on fuzzy method through blockchain technology. Soft Comput (2023). https://doi.org/10.1007/s00500-023-08355-x

[17]. Rashmi, R. Patil, et al. "Rdpc: Secure cloud storage with deduplication technique." 2020 fourth international conference on I-SMAC (IoT in social, mobile, analytics and cloud)(I-SMAC). IEEE, 2020.

[18]. Khetani, V., Gandhi, Y., Bhattacharya, S., Ajani, S. N., &Limkar, S. (2023). Cross-Domain Analysis of ML and DL: Evaluating their Impact in Diverse Domains. International Journal of Intelligent Systems and Applications in Engineering, 11(7s), 253-262.

[19]. Khetani, V., Nicholas, J., Bongirwar, A., &Yeole, A. (2014). Securing web accounts using graphical password authentication through watermarking. International Journal of Computer Trends and Technology, 9(6), 269-274.

[20]. Sairise, Raju M., Limkar, Suresh, Deokate, Sarika T., Shirkande, Shrinivas T. , Mahajan, RupaliAtul& Kumar, Anil(2023) Secure group key agreement protocol with elliptic curve secret sharing for authentication in distributed environments, Journal of Discrete Mathematical Sciences and Cryptography, 26:5, 1569–1583, DOI: 10.47974/JDMSC-1825.

[21]. Kale, R., Shirkande, S. T., Pawar, R., Chitre, A., Deokate, S. T., Rajput, S. D., & Kumar, J. R. R. (2023). CR System with Efficient Spectrum Sensing and Optimized Handoff Latency to Get Best Quality of Service. International Journal of Intelligent Systems and Applications in Engineering, 11(10s), 829-839.

[22]. Nagtilak, S., Rai, S., & Kale, R. (2020). Internet of things: A survey on distributed attack detection using deep learning approach. In Proceeding of International Conference on Computational Science and Applications: ICCSA 2019 (pp. 157-165). Springer Singapore.

[23]. Mane, Deepak, and AniketHirve. "Study of various approaches in machine translation for Sanskrit language." International Journal of Advancements in Research & Technology 2.4 (2013): 383.

[24]. Shivadekar, S., Kataria, B., Limkar, S. et al. Design of an efficient multimodal engine for preemption and post-treatment recommendations for skin diseases via a deep learning-based hybrid bioinspired process. Soft Comput (2023). https://doi.org/10.1007/s00500-023-08709-5

[25]. Shivadekar, Samit, et al. "Deep Learning Based Image Classification of Lungs Radiography for Detecting COVID-19 using a Deep CNN and ResNet 50." International Journal of Intelligent Systems and Applications in Engineering 11.1s (2023): 241-250.