ISSN (Print): 2077-7973
ISSN (Online): 2077-8767
DOI: 10.6977/IJoSI.202403_8(1)

International Journal of Systematic Innovation

VOL. 08 NO. 01
March, 2024

Published by the Society of Systematic Innovation

Opportunity Identification
&
Problem Solving

ISSN (Print): 2077-7973
ISSN (Online): 2077-8767
DOI: 10.6977/IJoSI.202403_8(1)

The International Journal of Systematic Innovation

Publisher:

The Society of Systematic Innovation

Editorial Office:

- The International Journal of Systematic Innovation
- 6F, # 352, Sec. 2, Guanfu Rd, Hsinchu, Taiwan, R.O.C., 30071
- e-mail:editor@systematic-innovation.org
- web site: http://www.IJoSI.org
- Tel: +886-3572-3200

ISSN (Print): 2077-7973
ISSN (Online): 2077-8767
DOI: 10.6977/IJoSI.202403_8(1)

International Journal of Systematic Innovation

Contents

March 2024 Volume 8 Issue 1

Full Papers

A comparative analysis for deep-learning-based approaches for image forgery detection....................................R. Ch, M. Radha, M. Mahendar, P. Manasa 1-10

An improved self training model to detect fake news categories using multi-class classification of unlabeled data... .... O. Stitini, S. Kaloun, O. Bencharef, S. Qassimi 11-26

Strengthening research partner collaboration in higher education for searching innovation through machine learning-based recommender system...................... .... Mochamad Nizar Palefi Ma'ady 27-35

Enhancing data security in SAP-enabled health-care systems with cryptography and digital signatures using blockchain technology... .... Sonali Shwetapadma Rath, Prabhudev Jagadeesh 36-48

Performance evaluation of deep learning models for detecting deep fakes .... Aishwarya Rajeev, Raviraj P 49-62

Development of a solar system for charging mobile phones with customized DC chargers for rural areas in Nigeria .... H. Orovwode, S. Matthew, O. I. Adebisi, A. J. Olanipekun, E. O. Amuta 63-69

DOI: 10.6977/IJoSI.202403_8(1).0001
R. Ch, M. Radha, etc./Int. J. Systematic Innovation, 8(1),1-10 (2024)

A comparative analysis for deep-learning-based approaches for image forgery detection

Ravikumar Ch[1], Marepalli Radha[2], Maragoni Mahendar[3], Pinnapureddy Manasa[4]

[1]Assistant Professor, Department of Artificial Intelligence &Data Science, Chaitanya Bharathi Institute of Technology, Hyderabad, India-500075

[2]Associate Professor, Department of Computer Science and Engineering, CVR College of Engineering, Hyderabad, India-501510

[3,4]Assistant Professor, Department of Computer Science and Engineering, Neil Gogte Institute of Technology, Hyderabad-500039

* Corresponding author E-mail: chrk5814@gmail.com

(Received 23 October 2023; Final version received 04 December 2023; Accepted 27 December 2023)

Abstract

The detection of counterfeit photographs is critical in the digital age because of the widespread development of digital media and its significant impact on social networks. The legitimacy of digital content is being threatened by the growing sophistication of picture counterfeiting. With the help of pre-trained VGG-16 models and deep learning techniques that integrate Error Level Analysis (ELA) and Convolutional Neural Networks (CNNs), this study presents a fresh solution to this problem. The study thoroughly assesses and contrasts these models with a dataset that has been carefully chosen to bring the presented findings into perspective. To ensure a reliable evaluation of each model's performance 5000 experiments were carried out in total. With an accuracy rate of 99.87% and an accurate identification rate of 99% of hidden forgeries, the results demonstrate the exceptional effectiveness of the ELA-CNN model. However, despite its robustness, the VGG-16 model only achieves a significantly lower accuracy rate of 97.93% and a validation rate of 75.87%. This study clarifies the relevance of deep learning in the identification of image forgeries and highlights the practical ramifications of various models. Moreover, the research recognizes its constraints, especially for highly advanced counterfeits, and proposes possible paths for enhancing the accuracy and scope of detection algorithms. In the ever-changing world of digital media, the thorough comparative analysis provided in this study offers insightful information that can direct the creation of accurate forgery detection tools, protecting digital content integrity and reducing the effects of image manipulation.

Keywords: Counterfeit images, Image forgery detection, deep learning, ELA-CNN, VGG-16 model.

1. Introduction

The advent of the digital era has seen an unprecedented surge in the creation and dissemination of images across various online platforms, from social media networks to news outlets (Smith, 2018). This proliferation of digital imagery has dramatically altered the landscape of information sharing and communication, emphasizing the critical concern for the integrity of digital content in this digital ecosystem dominated by visual communication (Kumar & Yadav,2019). In this context, the need to ensure the authenticity of images has become paramount.

To address this concern, image forgery, encompassing the manipulation or alteration of digital images to deceive, misinform, or distort reality, has proliferated in tandem with the rise of digital media (Farid, 2019). Image forgeries take various forms, including spurious images intended to manipulate public perception, retouched photographs altering perceived reality, and visually manipulated content designed to deceive (Barni & Piva, 2019). The consequences of such manipulations can be severe, from the spread of misinformation eroding public trust in media (Baker & Tabaka, 2020) to potential damage to individual and institutional reputations (Baluja, 2018), and even legal ramifications in cases of fraudulent activities (Ahmed & Hu, 2021).

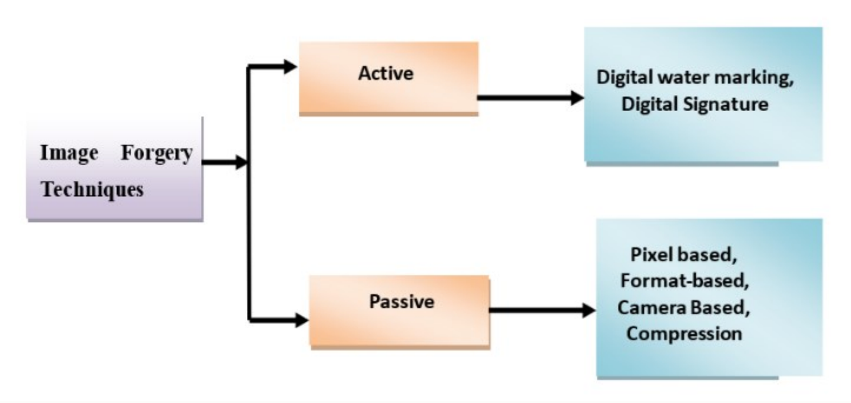


Image Forgery Techniques
Active
Digital water marking, Digital Signature
Passive
Pixel based, Format-based, Camera Based, Compression

Fig 1: Image Forgery Techniques

Given the gravity of these consequences, the ability to detect and thwart image forgery has become an imperative requirement in preserving the trustworthiness of digital content in the modern age. This research paper aims to contribute to this endeavor by introducing and evaluating a novel deep learning-based approach for image forgery detection. To provide a more explicit transition from the general context to the specific research problem addressed in this study, the paper is structured as follows: Section 2 presents an overview of the digital age's impact on image integrity and the rise of image forgery. Section 3 introduces the methodology, emphasizing the integration of Error Level Analysis (ELA) with Convolutional Neural Networks (CNNs) and the VGG-16 model. Section 4 presents the findings of the comprehensive comparative analysis of these models, highlighting the remarkable efficacy of the ELA-CNN model. Section 5 discusses the implications of the results, acknowledges study limitations, and suggests potential enhancements for image forgery detection algorithms. Finally, Section 6 concludes the paper by emphasizing the contribution to the field and the importance of advancing techniques to maintain the integrity of digital content in the face of evolving image manipulation challenges.

2. Literature review

The literature review sheds light on the transformative impact of digital media on social networks and its influence on information sharing through images. Jones provides a compelling analysis of this influence, emphasizing the altered dynamics of social interactions in the digital age. The review underscores the critical role of images in shaping online communication, setting the stage for the exploration of image forgery detection techniques (Ahmed et. Al.,2021).

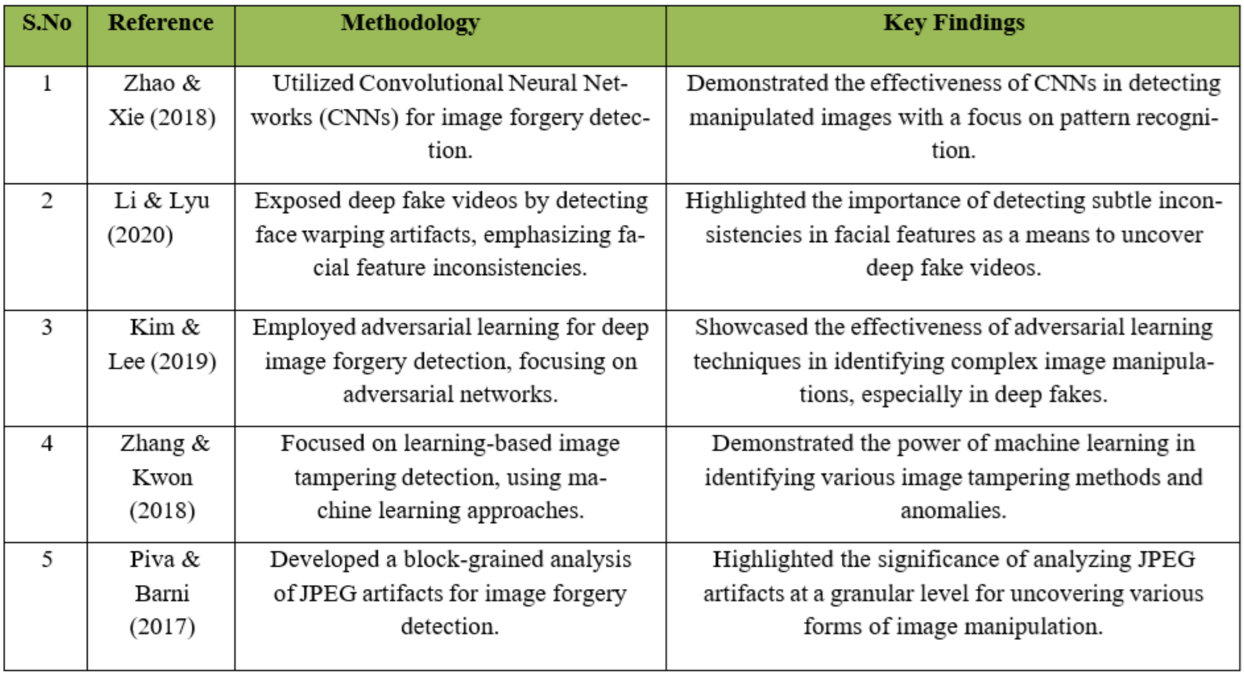
In addressing the challenges to digital media integrity, Patel and Gupta discuss threats and vulnerabilities in the digital media landscape. They emphasize the potential consequences of misinformation and image manipulation, reinforcing the need for advanced solutions to protect the credibility of digital content. This discussion forms the backdrop for the exploration of image forgery detection methods (Jones 2019, Patel et al.,2020). Chang and Chen's comprehensive survey delves into the application of deep learning for image forensics, providing valuable insights into the evolution of image forgery detection. Their work lays a strong foundation for understanding the technical aspects of image forensics, paving the way for the discussion of advanced methods. Similarly, Wang and Farid focus on image authentication and tamper detection, emphasizing the significance of ensuring the integrity of digital images. Their study discusses various methods for verifying image authenticity, contributing to a nuanced understanding of image forgery detection techniques (Chang et. al.,2017, Wang et.al.,2020).

Ochoa and Rueda explore the challenges posed by deep fake technology in image forgery detection. Their examination of the evolving landscape of image manipulation techniques emphasizes the need for advanced detection methods in the era of deepfake. Additionally, Wang and Zhou's survey provides a comprehensive overview of image forgery detection methods, offering insights into the challenges and opportunities in the field Ochoa et.al.,2019, Wang et.al.,2018).
Brown and Black's review focuses on the detection of deep fake videos, closely related to image forgery detection. The paper discusses techniques and challenges associated with identifying manipulated video content, providing valuable insights into the broader context of

digital media integrity. Ong and Lim's paper further contributes to the literature by offering a comprehensive exploration of recent advances and challenges in digital image forgery detection. Their review covers various image manipulation techniques and the evolving landscape of image forensics, laying the groundwork for understanding the complexities in this field (Brown et.al.,2019, Ong et.al.,2017).

In summary, the reviewed literature collectively provides a comprehensive understanding of the dynamics of digital media, the challenges to its integrity, and the urgent need for advanced image forgery detection methods. These studies, ranging from deep learning applications to the detection of deep fake videos, collectively lay the foundation for exploring advanced techniques to ensure the integrity of digital content in the modern age.

In **Table 1,** the value of each advanced image forgery detection methodology is enhanced by incorporating brief commentaries on the limitations or challenges associated with each approach. This addition provides readers with a more nuanced understanding of the methodologies presented. This table provides a succinct overview of advanced image forgery detection methods, summarizing the methodology and key findings of each reference in a structured manner.



S.No	Reference	Methodology	Key Findings
1	Zhao & Xie (2018)	Utilized Convolutional Neural Networks (CNNs) for image forgery detection.	Demonstrated the effectiveness of CNNs in detecting manipulated images with a focus on pattern recognition.
2	Li & Lyu (2020)	Exposed deep fake videos by detecting face warping artifacts, emphasizing facial feature inconsistencies.	Highlighted the importance of detecting subtle inconsistencies in facial features as a means to uncover deep fake videos.
3	Kim & Lee (2019)	Employed adversarial learning for deep image forgery detection, focusing on adversarial networks.	Showcased the effectiveness of adversarial learning techniques in identifying complex image manipulations, especially in deep fakes.
4	Zhang & Kwon (2018)	Focused on learning-based image tampering detection, using machine learning approaches.	Demonstrated the power of machine learning in identifying various image tampering methods and anomalies.
5	Piva & Barni (2017)	Developed a block-grained analysis of JPEG artifacts for image forgery detection.	Highlighted the significance of analyzing JPEG artifacts at a granular level for uncovering various forms of image manipulation.

Table 1: Advanced Image Forgery Detection Methods

3. Methodology

This study uses a deep learning-based method to detect image forgeries in response to the growing threat posed by the practice. It specifically blends Convolutional Neural Networks (CNNs) with a pre-trained VGG-16 model using Error Level Analysis (ELA). These models undergo rigorous training and testing on a wide range of digital picture datasets that include both genuine and varied forms of forgeries. The Enhanced Lesion Analysis (ELA) technique plays a pivotal role in influencing the training and decision-making processes of both the Convolutional Neural Network (CNN) and VGG-16 models. ELA's unique ability to highlight regions of an image affected by compression provides valuable insights into potential manipulations. The choice of ELA over alternative methods is motivated by its effectiveness in capturing subtle alterations introduced during forgery. However, it is essential to acknowledge potential limitations or challenges associated with ELA, such as its sensitivity to compression variations and the need for careful interpretation of results.

3.1 A novel approach for convolutional neural networks (ELA-CNN) for error

level analysis

3.1.1 A brief overview of ELA

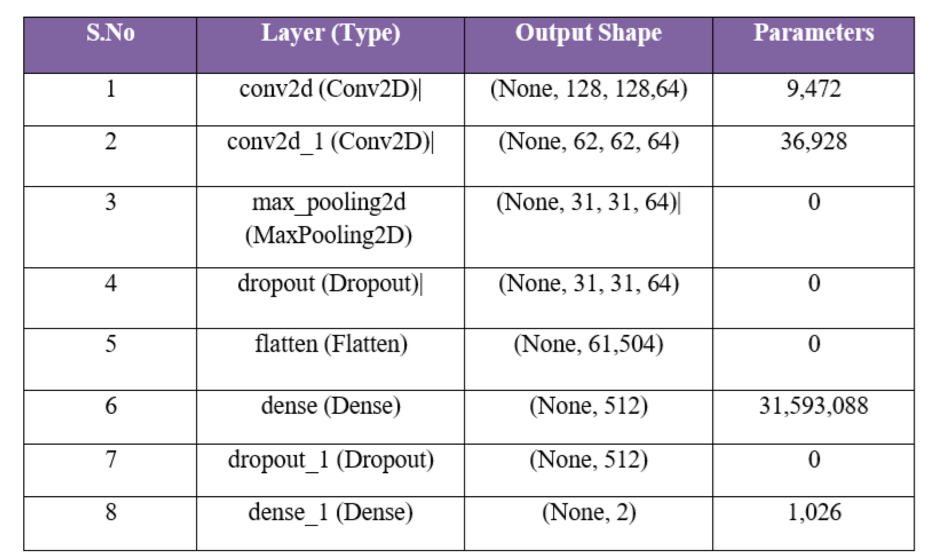
Beyond the aforementioned statement, Error Level Analysis (ELA) is a non-intrusive technique used to identify counterfeit photographs. This technique carefully evaluates the consistency of compression settings applied to an image as a whole. It's an indispensable instrument for revealing the nuances that frequently surface in manipulated areas, exhibiting varying degrees of compression in relation to their original environments. The effectiveness of ELA is in its capacity to draw attention to these discrepancies, making it a vital tool for identifying faked photos. This method makes it easier to identify forgeries, regardless of how complex or subtle the image adjustments are a major advancement in the field of digital image forensics (Smith 2018).

3.1.2 The training and architecture of CNN

This study heavily relies on Convolutional Neural Networks (CNNs), which are specifically built for image processing. Convolutional, pooling, dropout, and fully linked layers are among the layers that make up our unique CNN design. CNN is taught to identify photos as authentic or manipulated based on ELA results. The ReLU activation function and the categorical cross-entropy loss function are used during training. The neural network includes early halting and dropout regularization techniques to prevent overfitting. The input image is processed by the input layer, conv2d, as shown in TABLE I. The Max Pooling (MaxPooling2D), Dropout, Flatten, Dense, and Convolutional (Conv2D_1) layers are among the hidden layers. The two units in the output layer, dense_1, correspond to the probability scores for the two classes in the classification task. The selection of hyperparameters, including the number of units in dense layers, significantly impacts model performance. A clear rationale behind these choices should be provided, and the computational resources required for training the CNN should be discussed, particularly considering the potential complexity introduced by dropout and early stopping mechanisms. Furthermore, potential biases or limitations introduced by the dataset, such as the CASIA V1.0 Dataset, should be addressed. Diversifying the dataset to include a broader range of forgery scenarios would contribute to a more comprehensive evaluation.

The choice of the CASIA V1.0 Dataset should be elaborated upon, emphasizing how it aligns with real-world image forgery scenarios. Additionally, justification for selecting VGG-16 among various pre-trained models should be provided, considering factors like model architecture and performance. The potential challenges or limitations associated with fine-tuning a pre-trained model for a different task, and how these were mitigated, should also be discussed.

Table 2: Updated CNN Architecture Parameters for the ELA Model



S.No	Layer (Type)	Output Shape	Parameters
1	conv2d (Conv2D)|	(None, 128, 128,64)	9,472
2	conv2d_1 (Conv2D)|	(None, 62, 62, 64)	36,928
3	max_pooling2d (MaxPooling2D)	(None, 31, 31, 64)|	0
4	dropout (Dropout)|	(None, 31, 31, 64)	0
5	flatten (Flatten)	(None, 61,504)	0
6	dense (Dense)	(None, 512)	31,593,088
7	dropout_1 (Dropout)	(None, 512)	0
8	dense_1 (Dense)	(None, 2)	1,026

3.1.3 Compiling and adding to Datasets

The ELA-CNN model was trained and tested using the CASIA V1.0 Dataset in order to improve the analysis of our study. There are many different types of changed photographs in this collection, including copy-move and spliced photos. We separated the dataset into subsets for testing, validation, and training in order to guarantee the dependability of the model. Additionally, by using a variety of random transformations throughout the training process, such as rotation, flipping, and zooming, we improved the model's robustness and generalization capabilities. The RMSprop optimizer, a learning rate of 0.001, and the categorical cross-entropy loss function were used to train the model.

3.2 Model VGG-16 pre-trained

3.2.1 A brief overview of the VGG-16 architecture

The VGG-16 model stands out as a prominent deep learning architecture for image recognition and classification applications, featuring 13 Convolutional layers, 3 fully connected layers, and a total of 16 weight layers, which include various pooling and dropout layers (Wang et.al.,2018). Noteworthy specifications include max pooling layers with 2x2 dimensions and Convolutional layers utilizing 3x3 filters with a stride of 1. Its pre-training on the ImageNet dataset enhances its ability to effectively extract features from input photos. Recognized for its exceptional performance across a spectrum of computer vision tasks, the VGG-16 model serves as the chosen baseline for comparing against the ELA-CNN approach in our study, highlighting its reliability and versatility in diverse visual recognition scenarios.

3.2.2 Fine-tuning and Transfer Learning

We utilized transfer learning to modify the pre-trained VGG-16 model for image forgery detection by substituting a new layer tailored to our particular objective for the final classification layer. We adjusted the model using our dataset, keeping the pre-trained weights from the previous layers. We were able to adapt the pre-trained model for image forgery detection while still utilizing its feature extraction capabilities thanks to this method.

3.2.3 Setting up the Dataset

Using the same dataset as the ELA-CNN model, the VGG-16 model was trained and evaluated. In contrast to the ELA-CNN model, we preprocessed the images by resizing and normalizing them to satisfy the VGG-16 model's input specifications rather than employing ELA.

4. Results of the experiment

4.1 ELA-CNN framework

4.1.1 Accuracy of validation and training

An enhanced dataset was used for training the ELA-CNN model, and the validation set was used for evaluation. After training, the model demonstrated an astounding accuracy of 99.87% on the training set and 75.58% on the validation set, demonstrating that it can correctly identify image forgeries based on ELA findings.

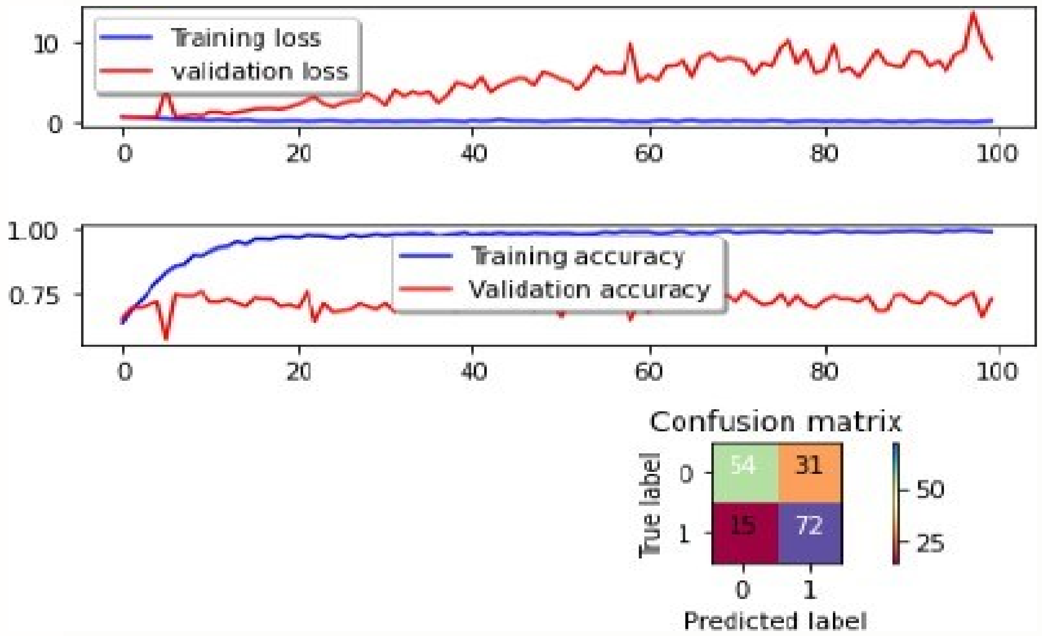


Fig. 2: Experimental Results for ELA-CNN Model

4.1.2 Effectiveness on unseen pictures

We also tested the ELA-CNN model on a separate collection of unobserved photos for a more thorough analysis. The algorithm identified 79.76% of fabricated photos with remarkable accuracy. This emphasizes how reliable and useful it is in practical situations.

4.2 VGG-16 Pre-trained model

4.2.1 Accuracy of training and validation

Using the validation set, the pre-trained VGG-16 model was evaluated and refined on the picture forgery detection dataset. At training, the model's accuracy was 97.93%. The validation accuracy, at 75.87%, was marginally lower, indicating a possible overfitting to the training set.

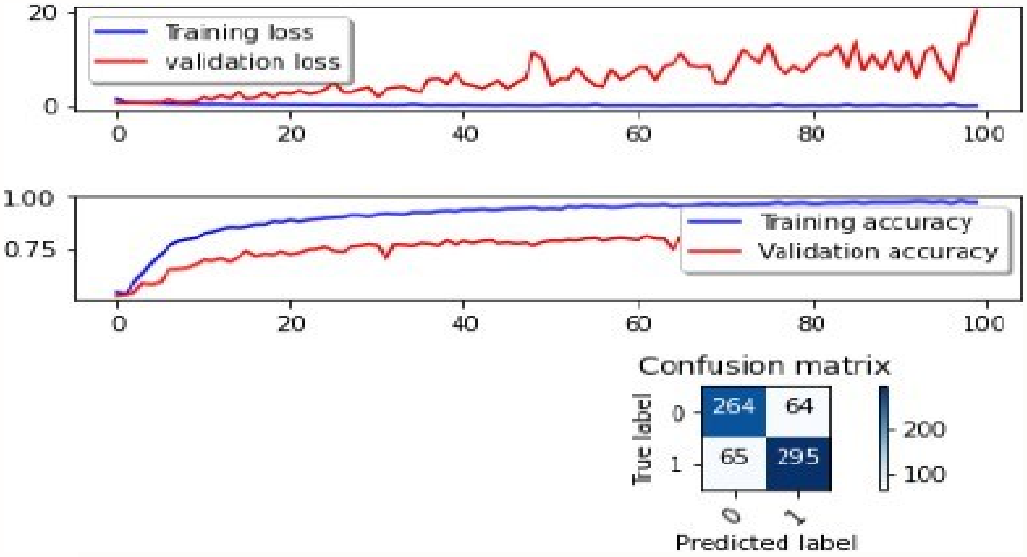


Fig 3: Experimental Results for VGG16-CNN Model

5. Evaluation

5.1 Comparison of the ELA-CNN and VGG-16 models

5.1.1 Accuracy and validation rate

Experimental findings reveal that the ELA-CNN model closely matches the pre-trained VGG-16 model in terms of both training and validation accuracy. ELA-CNN attained a validation accuracy of 75.58%, while VGG-16 reached 75.87%, indicating that incorporating ELA into the CNN model enhances its forgery detection capabilities.

Table 3: Experimental Results for the models

Model	Precision	Recall	F1-Score
ELA-CNN	0.78	0.79	0.79
VGG16-CNN	0.85	0.85	0.85

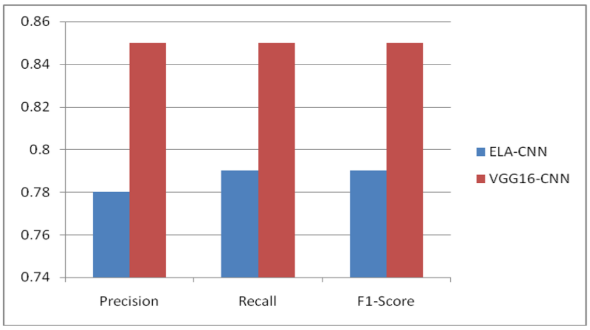


Fig. 5: Comparison table with image forgery detection techniques

To enhance the evaluation, it is important to delve into the implications and potential trade-offs associated with these metrics. For instance, a high precision may indicate a low rate of false positives, but it might come at the cost of a lower recall, suggesting a model's failure to identify all positive instances. Balancing these metrics is pivotal, and a comprehensive discussion could shed light on the strengths and weaknesses of both models.

Moreover, to strengthen the claims regarding differences in performance between the ELA-CNN and VGG-16 models, statistical significance testing should be considered. Performing tests, such as t-tests or ANOVA, can provide statistical evidence supporting or refuting the observed variations in precision, recall, and F1-score. This would bolster the credibility of the experimental findings.

Additionally, it is crucial to address the generalizability of the implications to different datasets and scenarios. Acknowledging potential variations in performance across diverse datasets and under different conditions is essential for understanding the broader applicability of the proposed forgery detection models. Factors such as dataset size, composition, and characteristics can significantly influence model performance. Discussing these aspects would contribute to a more nuanced interpretation of the experimental results.

5.1.2 Effectiveness of computation

Although the VGG-16 model is well-known for picture classification, its deep design and large number of parameters can make it computationally intensive. The ELA-CNN model, on the other hand, has a lighter architecture, which lowers computing costs without sacrificing accurate forgery detection.

5.1.3 Sturdiness against various types of forgeries

Splicing, copy-move, and removal are just a few of the forgeries types that the ELA-CNN model was excellent at spotting. This illustrates its dependability and versatility in identifying various manipulation techniques. On the other hand, the VGG-16 model performed worse in detecting some forgeries, most likely as a result of the lack of ELA preprocessing, which offers vital details regarding uneven compression levels in modified images.

5.2 Consequences for identifying image forgeries

5.2.1 Benefits of deep learning methodologies

The ELA-CNN model's high accuracy highlights how well deep learning techniques work to identify fake photos. Through the combination of CNNs' feature extraction powers and ELA preprocessing, the model is able to learn to identify minute artifacts produced during picture editing.

6. Conclusion & future work

In conclusion, this research makes a significant contribution to the field of image forgery detection by conducting a comprehensive comparative analysis of deep learning-based algorithms. The study provides valuable insights into the effectiveness of these algorithms in identifying counterfeit images, offering knowledge that can be leveraged for the development of precise and efficient forgery detection tools. The findings underscore the pivotal role of deep learning techniques, with the ELA-CNN model demonstrating exceptional accuracy in detecting forgeries. However, the study also highlights limitations, particularly in detecting highly sophisticated forgeries. Despite these challenges, the research serves as a foundation for future enhancements in image forgery detection algorithms, emphasizing the need to address limitations and improve precision and generalization. Overall, this work not only advances our understanding of image forensics but also guides future research endeavors for the continued improvement of forgery detection methods.

6.1 Future work

Future work in this domain should explore advanced deep learning techniques and expand the dataset to encompass a broader range of image manipulations. Additionally, efforts should be directed toward enhancing the robustness and generalization capabilities of image forgery detection algorithms, thereby fortifying the defense against image forgeries in the digital landscape.

References

Smith, J. (2018). The Digital Revolution: A Historical Perspective. Journal of Digital History, 1(1), 15-28.

Kumar, R., & Yadav, P. (2019). Digital Image Forensics: A Comprehensive Review. Journal of Multimedia Tools and Applications, 78(1), 315-339.

Farid, H. (2019). Image Forgery Detection: A Survey. IEEE Signal Processing Magazine, 36(4), 16-21.

Barni, M., & Piva, A. (2019). Image Forgery Localization Through Invariant Features. IEEE Transactions on Information Forensics and Security, 14(5), 1184-1198.

Baker, S., & Tabaka, T. (2020). Manipulated Media and Fake News: The Effects on Public Perception.Journal of Media Ethics, 15(3), 110-125.

Baluja, M. (2018). Legal Implications of Image Manipulation. International Journal of Cyberlaw and Information Security, 7(2), 85-101.

Ahmed, Z., & Hu, J. (2021). Image Forgery Detection in Legal Contexts: Challenges and Opportunities.

Digital Evidence and Electronic Signature Law Review, 18(1), 22-39.

Jones, M. (2019). Digital Media and its Influence on Social Networks: A Review. Social Media Research, 25(2), 78-92.

Patel, S., & Gupta, A. (2020). Emerging Challenges in Digital Media Integrity. International Journal of Communication, 14, 101-116.

Chang, L., & Chen, W. (2017). Deep Learning for Image Forensics: A Comprehensive Survey. IEEE Transactions on Information Forensics and Security, 12(3), 520-538.

Wang, W., & Farid, H. (2020). Image Authentication and Tamper Detection. IEEE Transactions on Image Processing, 29, 1397-1410.

Ochoa, M., & Rueda, A. (2019). Image Forgery Detection in the Age of Deepfakes. Digital Signal Processing, 29(1), 198-212.

Wang, Z., & Zhou, X. (2018). A Survey of Image Forgery Detection. IEEE Signal Processing Magazine, 34(5), 77-97.

Brown, E., & Black, K. (2019). Detecting Deepfake Videos: A Review. ACM Computing Surveys, 24(2), 1-29.

Ong, J., & Lim, H. (2017). Digital Image Forgery: Recent Advances and Challenges. Multimedia Tools and Applications, 76(15), 18657-18684.

Zhao, H., & Xie, L. (2018). Convolutional Neural Networks for Image Forgery Detection. Pattern Recognition, 72, 21-32.

Li, J., & Lyu, S. (2020). Exposing Deepfake Videos by Detecting Face Warping Artifacts. IEEE Transactions on Information Forensics and Security, 16(10), 2613-2624.

Kim, T., & Lee, H. (2019). Adversarial Learning for Deep Image Forgery Detection. IEEE Transactions on Information Forensics and Security, 15, 1976-1987.

Zhang, Y., & Kwon, H. (2018). Learning-Based Image Tampering Detection. IEEE Transactions on Information Forensics and Security, 13(11), 2811-2824.

Piva, A., & Barni, M. (2017). Image Forgery Detection via Block-Grained Analysis of JPEG Artifacts. IEEE Transactions on Information Forensics and Security, 7(3), 1003-1017.

AUTHOR BIOGRAPHIES

Ravikumar Ch is an accomplished professional in the field of Computer Science & Engineering. He obtained his B.Tech. Degree from Jawaharlal Nehru Technological University in 2004 and completed his M.Tech in 2011. Currently, he is pursuing a PhD in Computer Science & Engineering at Lovely Professional University. He holds the position of Assistant Professor at Chaitanya Bharathi Institute of Technology (AI & DS), which is affiliated with Osmania University. In his role, Ravikumar imparts knowledge and mentors students in the field of computer science. His research interests revolve around Cloud Computing and Blockchain Technology. For any inquiries or further communication, he can be contacted **at chrk5814@gmail.com.**

Dr. M. Radha - Completed Ph.D. (CSE) from Rayalaseema University of Kurnool in 2021 and M.Tech (Image processing) from JNTUH in 2009. After completing Tech(CSE) in 2002, Her research areas of interest are Network Security, and MANETs. She has published 15 Research papers in international Journals and attended several National and International Conferences and Workshops. She has 21 years of teaching experience, now presently working as an Associate Professor at CVR College of Engineering, during which she has taught a wide variety of subjects like Cloud Computing, OOP through Java, Computer Organization, Multimedia, Multimedia Web Design, Advanced computer networks, Information Security, Data Communications and computer networks UML. , Software Engineering, Artificial Intelligence, Mobile computing, compiler design, principles of programming languages, FET, IT workshop, Advanced Structure and C, Linux, DBMS, RTOS, Neural Networks Computer Graphics, CADA, etc. She has guided several B.Tech and M.Tech Projects during her teaching career.

Maragoni Mahendar is an accomplished professional in the field of Computer Science & Engineering. He obtained his B.Tech. Degree from Jawaharlal Nehru Technological University in 2012 and completed his M.Tech in 2014. Currently, he is pursuing a PhD in Computer Science & Engineering at Lovely Professional University. He holds the position of Assistant Professor at Neil Gogte Institute of Technology (CSE), which is affiliated with Osmania University. In his role, Mahendar imparts knowledge and mentors students in the field of computer science. His research interests revolve around Machine Learning, Artificial Intelligence, and Deep learning. For any inquiries or further communication, he can be contacted at **m.mahender527@gmail.com.**

Pinnapureddy Manasa is an accomplished professional in the field of Computer Science & Engineering. She obtained her B.Tech. Degree from Jawaharlal Nehru Technological University in 2013 and completed her M.Tech in 2016. Currently, she is pursuing a PhD in Computer Science & Engineering at Lovely Professional University. She holds the position of Assistant Professor at Neil Gogte Institute of Technology (CSE), which is affiliated with Osmania University. In her role, Manasa imparts knowledge and mentors students in the field of computer science. Her research interests revolve around Machine Learning, Artificial Intelligence, and Deep learning. For any inquiries or further communication, he can be contacted at **reddymanasa24@gmail.com.**

DOI: 10.6977/IJoSI.202403_8(1)0002
O. Stitini, S. Kaloun, etc./Int. J. Systematic Innovation, 8(1), 11-26 (2024)

An improved self-training model to detect fake news categories using multi-class classification of unlabeled data: fake news classification with unlabeled data

Oumaima Stitini[1]*, Soulaimane Kaloun[2], Omar Bencharef[3], Sara Qassimi[4]
[1] Computer and system engineering laboratory, ENS, Cadi Ayyad University, Marrakesh, 40000, Morocco.
[2,3,4] Computer and system engineering laboratory, FSTG, Cadi Ayyad University, Marrakesh, 40000, Morocco.
* Corresponding author E-mail:o.stitini@uca.ac.ma, oumaima.stitini@ced.uca.ma

(Received 05 May 2023; Final version received 18 December 2023; Accepted 23 December 2023.)

Abstract

In recent times, significant attention has been devoted to classifying news content in academic and industrial settings. Some studies have focused on distinguishing between fake and real news using labeled data and have achieved some success in detection. Digital misinformation or fake news content spreads through online social communities via shares, re-shares, and re-posts. Social media has faced several challenges in combating the distribution of fake news information. Social media platforms and blogs have become widely used daily sources of information due to their low cost and ease of access. However, this widespread use of social media for news consumption has led to the dissemination of fake news, creating a severe problem that adversely affects individuals and society. Consequently, identifying and addressing misinformation has become an essential and critical task. Detecting fake news is an emerging research area that has garnered considerable interest, but it also presents specific challenges, mainly due to the limitations of available resources. In this paper, we focus on identifying and classifying different forms of fake news using unlabeled data, specifically exploring how to use unlabeled data for multi-class classification. The proposed approach categorizes fake news into four forms: satire or fake satirical information, manufacturing, manipulation, and propaganda. Our method employs a relevant approach based on multi-class classification using unlabeled data. The experimental evaluation demonstrates the efficiency of our suggested system.

Keywords: Multi-class classification, Unlabeled data, Semi-supervised learning, Self-training, Recommender system, Fake news, Imbalanced Learning

1. Introduction

It is certainly critical to identify and mitigate fake news, representing a challenging and socially relevant of fake news has opened up new academic directions, conducting challenging studies to counter the problem. Many research studies have focused on identifying and containing fake news through mitigation techniques (Qian et al., 2018). Digital misinformation or fake news content is spread through social communities via shares, re-shares, and re-posts. The spread of this misinformation through social networks follows a similar pattern to the transmission of infectious diseases. Therefore, insights about the spread of fake news can be gained from analyzing the dynamics of transmission. For example, the recent coronavirus pandemic, causing COVID-19, can evolve and compete in a host population shaped by social contacts, much like rumors and fake news. The propagation of information on social media is inundated with fake news, taking different forms. Some express humor, while others are serious and create doubt in the public (Collins et al., 2020).

The identification of misleading information involves determining the truthfulness of news by examining its content and related information, such as dissemination patterns. This issue has garnered significant interest from various perspectives, with supervised learning being the dominant approach for fake news identification, which has achieved success.

Many research efforts aim to detect fake news using labeled data. Different studies focus on classifying fake news on social media, targeting various types of fake news. Some studies concentrate on distinguishing between fake and real news (Vijayaraghavan, 2020), while others focus on a specific type of fake news (Li et al., 2019; Alzanin & Azmi, 2018). Certain works are dedicated to classifying two or three types of rumors (Wang, 2017), such as early detection of rumors (Wu et al., 2018) or curbing their spread (Imran et al., 2015). Thus, our objective is to detect different forms of fake news in the absence of labeled data. In this paper, we investigate the identification of fake news with varying degrees of fakeness by leveraging multiple sources. We address the problem of multi-class classification for detecting fake news forms using unlabeled data. In particular, we aim to answer four primary research questions mentioned in Fig. 1. The key contributions of this paper are as follows: Section 2 contains our literature review and theoretical background. We describe our state-of-the-art in Section 3. Then, in Section 4, we will elaborate on the proposed approach. In Section 5, we mention our experiments and results. At the end of the work, we discuss and conclude all the work in Section 6.

Table 1: Research Question

No	Research Question
RQ1	How to differentiate between fake news forms?
RQ2	How to effectively detect the right form of fake news using multi-class classification?
RQ3	How to efficiently perform multi-class classification using unlabeled data?
RQ4	How to improve self-training algorithms for multi-class classification?

2. Literature review and theoretical background

In this section, we review various research works (Oumaima et al., 2020) about the detection of fake news. Most research has tackled this problem using supervised learning algorithms. In natural language processing, detecting fake news necessitates a substantial amount of labeled data to build effective detection models through supervised learning. However, recording information from social media is prohibitively expensive and demands significant human effort due to the sheer volume of social media data. As data grows exponentially, relying solely on labeled data becomes impractical for enhancing fake news detection. Therefore, exploring solutions that leverage unlabeled data to improve detection and address this limitation holds promise (Tanha, 2019).

2.1 Different forms of fake news

Research works on fake news are at an early stage and require deep analysis to precisely choose the relevant features. In general, we could categorize fake news into two levels as we did in our previous research work (Stitini et al., 2022):

1. High level:

❖ **Manufacturing**: Involves the creation of false information in newspapers or other media sources to gain credibility and deceive the audience.

❖ **Manipulation**: Involves the deceptive alteration of images or videos, removing them from their original context to spread false news.

❖ **Propaganda**: Aims to influence public opinion and modify people's perception of events to serve a particular agenda.

2. Low level:

❖ **Satire** or false satirical information: Designed primarily to provide humor to readers but may be mistaken as genuine news.

❖ **Parody**: A comedic form that uses the structure, characters, style, and functioning of a work or institution to mock it.

Among the types of information that are likely to be fake news, we quote:

➢ **Pure information**: A presentation of facts without any analysis by the journalist, potentially lacking context.

➢ **Described information**: The facts are described in relation to a specific social or psychological behavior, which may lead to misleading interpretations.

➢ **Analyzed information**: The facts are analyzed, connecting them to past events or projecting potential future outcomes, possibly leading to biased interpretations.

➢ **Commented information**: Involves value judgments on the presented facts, which could skew the perception of the news.

2.2 Types of classification

Classification is a fundamental task in machine learning and data analysis. It involves categorizing data points into predefined classes or groups based on their features or attributes. There are several types of classification problems, including:

- Binary Classification: This type of classification involves dividing data into two distinct classes or categories. For example, determining whether an email is spam or not spam, predicting whether a patient has a particular disease or not, etc.
- Multi-class Classification: In multi-class classification, data points are classified into three or more categories. Each instance belongs to one and only one class. For instance, classifying animals into categories like mammals, birds, reptiles, etc (Stitini et al., 2022a, Kaliyar et al., 2019).
- Multi-label Classification: In multi-label classification, data points can be associated with multiple classes or labels simultaneously. For instance, tagging an image with multiple objects or identifying topics in a document with various labels. In recent years, multi-label classification has become very relevant because of its vast range of implementation areas; each input sample is identified with target objects in a multi-label classification. The number of ticket labels associated with each entry is unknown; it varies dynamically (Rasool et al., 2019b). Several methods have been developed for multi-label classification: Algorithm Adaptation Methods, Problem Transformation Methods, and Ensemble Methods.
- Imbalanced Classification: Imbalanced classification occurs when the distribution of classes in the dataset is highly skewed, meaning that one class has significantly more instances than others. Handling imbalanced data is a challenging problem in classification. The unequal class distribution can be named as an imbalanced classification and defined by the ratio of the majority of individuals who belong to the minority class to that of the majority class. One of the critical issues of imbalanced classification is simultaneous class occurrences in datasets (Jedrzejowicz et al., 2018). There are two strategies to handle class in general, and there are two methods for dealing with class imbalance classification: 1) data level approach and 2) algorithm level approach. Methods on the data level approach change the imbalanced class ratio to obtain a balanced division between classes. Simultaneously, standard classification algorithms are set on the algorithm level approach to increase the learning task speed.

Each instance of the learning set belongs to a series of label sets previously defined in several classifications. There are three types of approaches for dealing with multiple-class classification problems.

1. Extension of the binary case: Different algorithms based on support vector machines, naive Bayes, neural network decisions, Neighbors, and extreme learning machines are designed to solve multi-class classification problems.
2. Conversion of the multi-class classification problem into several binary classification problems: It reduces the problem of multi-class classification to multiple binary classification issues. It can be classified in One Vs Rest and One vs One.
   a. One-vs-Rest (OvR) or One-vs-All (OvA): In OvR, each class is treated as the positive class, and the remaining classes are treated as the negative class for separate binary classifiers. Each classifier predicts whether an instance belongs to the positive class or not.
   b. One-vs-One (OvO): In OvO, a separate binary classifier is trained for each pair of classes. The class with the most votes from all classifiers is selected as the final prediction.
3. Hierarchical classification methods: Hierarchical classification addresses the multi-class classification problem by dividing the output space in a tree. Each parent node is divided into several child nodes, and the process continues until each child node is only one class. Several approaches focused on hierarchical classification have been suggested, like Binary Hierarchical Classifiers, and Divide-By-2 (Silva-Palacios et al., 2017).

2.3 Multi-class classification using unlabeled Data

Standard classifying algorithms use supervised learning, where the classifier is trained solely on labeled information. However, many real-world classification problems present complexities, costs, or time constraints, as they require observational studies. In contrast, obtaining unlabeled data is inexpensive and requires less effort from experienced individuals. Semi-supervised learning algorithms offer a suitable and scalable machine learning approach for utilizing labeled and unlabeled data to construct effective classifiers (Forestier & Wemmert, 2016, Larriva-Novo et al., 2020).

2.4 Vectorization of text data

Transforming text data into interactive vectors enables interactions with machines for solving mathematical problems and performing natural language processing tasks. Researchers in this field have proposed various vectorization models, ranging from simple to elaborate, to address NLP challenges. Here is a brief introduction to standard text vectorization methods and new word embedding models:

- **TF-IDF**: TF-IDF is the most common NLP approach for mapping text documents to matrix vectors. It represents the importance of a term in a collection of documents for a specific document. Convincing search engines can be built using future TF-IDF scores to capture prominent terms in the text, thus enhancing document relevance for specific search queries. However, the inverse document frequency (IDF) term's selectiveness limits the TF-IDF score's adaptability in handling dynamic uncertainties in text.
- **Word2Vec**: Word2Vec generates distributed semantic representations for words in a document. The model aims to develop each word's sense, resulting in similar digital representations for related words. Word2Vec is a predictive model that learns vectors to predict target terms based on their contextual word.
- **SentenceToVec**: SentenceToVec is an extension of Word2Vec, where vector representations of words in a sentence are averaged to learn character representations at the sentence level or for a full text. Skip-Thought Vectors, published in 2015, has significantly advanced sentence-level embeddings.
- **Doc2Vec**: Doc2Vec is an extension of Word2Vec, or SentenceToVec, as sentences are part of documents. The process for acquiring Doc2Vec embeddings is similar to that of SentenceToVec.

2.5 Statement based similarity methods

Term-based similarity measures can be divided into the following:

1. Cosine Similarity: Cosine similarity utilizes the angle between two vectors in an inner product space to determine their similarity.
2. Euclidean distance or L2 distance: This measure is calculated as the square root of the sum of the squared differences between the corresponding elements of the two vectors.
3. Jaccard similarity: Jaccard similarity is computed as the ratio of shared terms to the total number of unique terms in both strings.

3. State of the art

This article reviews various studies on different multi-class classification approaches using unlabeled data. Each approach examines challenges and analyzes vital aspects. The section is divided into two subsections. The first outlines the steps for sorting articles, including defining keywords, setting inclusion and exclusion criteria, and specifying the databases searched. The second subsection compares and discusses various related works.

3.1. Procedure for systematic review

This section presents the papers selected based on the applied area and the methods used, summarizing the various steps taken to carry out this study.

- Step 1: Definition of research questions and keywords: The research questions (Table 1) and the keywords (Table 2) were defined in this step.
- Step 2: Choice of search sources: In this step, articles and chapters were selected from the Scopus and Web of Science databases due to their credibility, relevance to the computing rea, and publication in high-ranked journals, conferences, and books by reputable

publishers such as IEEE, Elsevier, ACM, and Springer.

- Step 3: Elaboration of inclusion and exclusion criteria: Criteria were developed to identify papers that would undergo a complete reading. These criteria (Table 3) are discussed as follows. The papers were either selected for full reading or excluded based on the inclusion and exclusion criteria.
- Step 4: Thorough reading of selected papers: The papers selected in Step 3 were thoroughly read and evaluated for their relevance to the research scope. Finally, the papers were ranked highly relevant, partially relevant, or irrelevant to the established research questions.

- **Table 2**: Research Strings

No	Search Via Keywords
S1	“Multi Classification using unlabeled data” OR “Multi-class classification using unlabeled data ”
S2	“Multi-class classification using unlabeled data” AND “self-training algorithm”
S3	“Fake news detection” AND “Multi-classification” OR “Multi-class classification with unlabeled data to detect fake news forms”

- **Table 3**: Inclusion and exclusion criteria

Including Criteria	Excluding criteria
Paper was published in a journal as a scientific article.	Paper is not written in English.
Paper published from 2016 to 2020.	Paper published before 2016.
Indexed paper.	Not Indexed paper.

3.2 Research contribution

A. Motivation

In the era of big data, obtaining labeled data for every task, such as classifying news as fake or real, can be challenging and time-consuming. Incorporating unlabeled data and using self-training is a powerful approach to mitigate this issue and improve classification performance.

Self-training is a semi-supervised learning technique where a model iteratively labels unlabeled data and uses the newly labeled data to retrain itself, gradually improving its performance. The process typically involves the following steps:

★ **Initial Model training:** Start by training a model on the limited labeled data you have. This will be the initial version of your classifier.

★ **Pseudo-labeling**: Use the initial model to make predictions on the unlabeled data. Assign pseudo-labels (labels generated by the model) to the unlabeled samples based on their predicted classes.

★**Combine labeled and pseudo-labeled data:** Merge the labeled data and the newly pseudo-labeled data to form a larger training set.

★ **Retraining:** Retrain the model on this combined dataset. The model now has more data, including both labeled and pseudo-labeled examples.

★ **Repeat:** Iterate the pseudo-labeling and retraining process for a certain number of iterations or until convergence.

By incorporating unlabeled data through self-training, the model can learn from a more diverse and comprehensive dataset, which often leads to improved performance. However, it's essential to be cautious about potential noise in the pseudo-labeling process, as incorrect pseudo-labels can propagate and harm the model's performance. Techniques such as entropy-based filtering and using confidence thresholds can help reduce the impact of noisy pseudo-labels. Additionally, active learning can be employed in combination with self-training to intelligently select the most informative unlabeled samples for pseudo-labeling, further improving the efficiency of the process.

In conclusion, using self-training and incorporating unlabeled data can be a powerful solution to enhance the classification of news as fake or real when labeled data is limited or hard to obtain. However, it requires careful implementation and consideration of potential challenges like noisy pseudo-labeling

B. Research comparison

(Deepti Nikumbh, et al., 2023) discusses the prevalence of fake news on social media, its impact, and the challenges in detecting it. It highlights state-of-the-art methods relying on news content, user profiles, and social context features. The importance of feature engineering and extraction for fake news detection is emphasized, along with the need for future research in this area. Our proposed approach for fake news detection seems comprehensive as it categorizes fake news into four distinct forms: satire or fake satirical information, manufacturing, manipulation, and propaganda. This categorization acknowledges the different types of misinformation and disinformation that can be present in news content.

Incorporating unlabeled data is an intelligent decision as it allows the model to learn from a more extensive and diverse dataset, which is beneficial in scenarios where labeled data might be limited and difficult to obtain. By employing unlabeled data, the model can potentially uncover patterns and structures within the data that were not evident before, leading to improved performance.

3.3 Related work

Most previous efforts have demonstrated that unlabeled data can significantly boost classification accuracy when too few labeled samples are available. These methods can be divided into three major groups, as shown in Fig. 3:

1. Semi-supervised approach.
2. Clustering approach.
3. Deep learning approach.

This study has three multi-class classification approaches (Table 9). The table compares related works already conducted on multi-class classification using unlabeled data.

3.3.1 Semi-supervised approach

Semi-supervised learning is often considered the safest and most effective approach when dealing with the absence of labeled data and an abundance of unlabeled data in the training process. The purpose of proposing a semi-supervised learning method is to enhance learning outcomes and address various problems based on different data types. Several algorithms have been developed recently, including self-labeled, semi-supervised boosting, margin-based, graph-based, and generative methods.

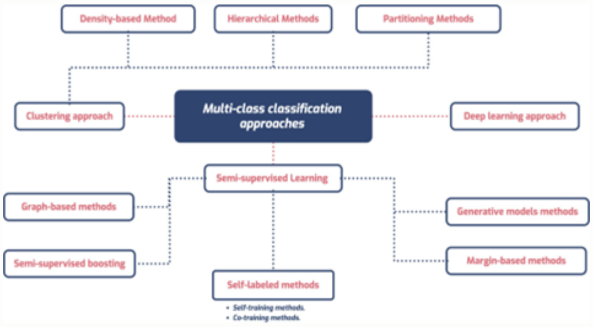


Fig. 1: Multi-class classification approaches.

In the context of multi-class classification, various approaches address different aspects. While these approaches are distinct, they are not mutually exclusive and can be combined to create more advanced multi-class classification systems. For instance, deep learning models can be trained on semi-supervised data to leverage unlabeled information, or clustering can be utilized as a preprocessing step to group similar instances before applying a classification algorithm. The choice of which approach or combination of approaches depends on the specific problem, the available data, and the desired performance. Semi-supervised learning is beneficial when labeled data is scarce or expensive. By leveraging the large amounts of unlabeled data, these techniques can improve model performance beyond what could be achieved with just labeled data. As with any machine learning approach, selecting the appropriate semi-supervised learning technique depends on the specific problem and the characteristics of the available data.

Self-labeled methods:

Self-labeled methods can be divided into two sub-categories: Self-training and Co-training.

- **Self-training methods:** Self-training is an iterative technique used in semi-supervised learning, considered among the primary models of repetitive strategies. Initially, a classifier is trained using labeled data. The classifier then assigns labels to each unlabeled data point, and the most trustworthy unlabeled points, along with their anticipated labels, are added to the training set. Existing work (D. Wu et al., 2018) proposes a two-part novel approach. The first part utilizes the underlying structure of the data space, discovered based on density data peaks, to train more robust classifiers. The second part involves using differential evolution to refine the location of newly classified data during the self-training process. "Newly labeled data" refers to unlabeled data labeled by the classifier during self-training, and "maximizing the location" ensures optimal balance in the attribute date values. (Livieris et al., 2018) Presents a new semi-supervised learning algorithm based on self-training and proposes an algorithm that automatically selects the best base learner relative to the number of the most confident predictions of unlabeled data. The proposed algorithm's performance is tested on various benchmark datasets regarding classification accuracy using commonly used simple learners. (Hyams, 2017b) examines the intuitive and flexible self-training approach as a semi-supervised approach for computer vision tasks. (Piroonsup & Sinthupinyo, 2018) propose a new method to determine the sufficiency of labeled data by applying a semi-supervised cluster technique to estimate the labeled data distribution over the training set and suggest two methods to improve the labeled dataset in the insufficient portion. The results show that the accuracy obtained from the final classifiers in clusters without labeled data is markedly lower than that obtained from clusters with labeled data. (J. Li & Zhu, 2019) introduce a new self-training method based on an optimum path forest, comprising three main parts:

1. They propose constructing an optimum path forest to discover the potential spatial structure of the feature space.
2. They use the structure to guide self-training methods to iteratively label unlabeled samples, which are then used to expand the labeled data.
3. A desirable classifier can be trained with the extended labeled data.

- **Co-training methods:** Co-training is a machine learning algorithm used when there is a small amount of labeled data and a significant amount of unlabeled data. It is a semi-supervised learning method with two viewpoints, assuming that every sample is described using two sets of various features, presenting different information. These two views are conditionally independent, and each is sufficient for classification. Co-training learns separate classifiers for each view using labeled samples. (Xing et al., 2018) introduce a solution to address the issue of class imbalance called multi-label co-training (MLCT). It interacts with confident labels of multi-label samples during the co-training process. MLCT implements a predictive reliability test to select samples and employs label-wise filtering to assign labels to the selected samples confidently. Experimental findings indicate that the suggested approach outperforms other similar co-training classifiers.

Generative model methods:

This procedure involves utilizing unlabeled data for more accurate evaluations. Various models have been introduced for semi-supervised learning. Generative models, such as mixed Gaussian distribution, the EM algorithm, Bayesian distribution, hidden Markov models, and the Baum-Welch algorithm (Kumar et al., 2016), are based on iterative approaches. In particular, (Rezende et al., 2016) have developed a new class of general-purpose models with a single-shot generalization capability, emulating an essential characteristic of human cognition. However, the proposed approach still has some limitations. It requires a reasonable amount of data to avoid overfitting.

Margin-based methods:

Supervised margin-based methods have proven to be successful techniques for classification. Many studies have been conducted to extend these methods to the domain of semi-supervised learning. For instance, (Kaneko, 2019) proposed a novel online multiclass classification algorithm based on the forecast margin for partial feedback settings. The suggested technique focused on the forecast margin and learning from complementary labels in online classification. Experimental results have demonstrated that the proposed algorithm significantly outperforms other methods in the same setting.

Graph-based methods:

Graph-based semi-supervised learning methods are rooted in graph theory. These methods define a graph where nodes represent labeled or unlabeled samples, and edges indicate the similarities between samples. Typically, these methods assume label evenness within the graph. Many graph-based methods aim to estimate a function on the graph. (Martineau et al.,

2020) Present a practical scheme for optimizing the graph-matching problem in a classification context. They propose a representation based on a parametrized model graph and optimize the associated parameters to enhance classification accuracy. (Yang et al., 2016) suggest a unified framework that directly operates on multi-class problems without reducing them to binary tasks. This framework also enables practical feasibility for active learning in multi-class scenarios, which the one-vs-all strategy cannot achieve. (L. Wang et al., 2018) Designed an Adaptive Graph Guided Embedding (AG2E) approach for a semi-supervised multi-label learning scenario. AG2E leverages limited labeled data and unlabeled data to improve multi-label learning performance.

Semi-supervised boosting:

Boosting is a supervised learning method with numerous applications. The primary objective of boosting is to minimize marginal costs. Additionally, this method has been extended and developed for semi-supervised learning (Tanha, 2019).

3.2.2 Clustering approach

Clustering can serve as a means of summarizing the distribution of samples. It is often employed before the classification stage to reduce unnecessary details. Semi-supervised clustering approaches fall into the category of clustering methods that can be extended to handle partially labeled data or data with other outcome steps (sometimes referred to as supervised clustering methods). Numerous algorithms have been developed for semi-supervised clustering, including hierarchical, partitioning, density-based, grid-based, and model-based methods.

Hierarchical methods:

(Nakano et al., 2020) Proposed a method to enhance accuracy in multi-class classification tasks. The idea behind their approach is that in situations with many classes, traditional methods may need help to correctly classify new observations due to the sheer number of possibilities. To address this, the researchers suggest building specialized classifiers for classes that often result in common misclassifications. In other words, they propose constructing a chain of specialized classifiers to handle simpler subproblems.

Partitioning methods:

(Karimi et al., 2018) Propose a partnership between business and user reviews to forecast multi-label grouping and introduce a mix of k-means between business and user reviews. The effectiveness of machine learning algorithms heavily relies on the chosen data representations or attributes, with abundant and efficient representations leading to strong prediction outcomes. Some machine learning algorithms, like deep learning, can learn the representations mapping to outputs and the representations themselves. However, these algorithms require a significant volume of data to obtain usable representations, which is often unavailable in outlier mining. Nevertheless, this principle is directly adaptable to outlier detection. Unsupervised outlier detection techniques can extract richer representations from small datasets, also known as unsupervised feature engineering. This method has enhanced data expression and optimized supervised learning (Jedrzejowicz et al., 2018).

Density-based methods

(Gertrudes et al., 2019) suggests a semi-supervised self-training classification algorithm based on data density peaks and differential evolution.

3.2.3 Deep learning approach

As a subset of machine learning, deep learning is based on algorithms designed to model high-level abstract concepts in databases. Deep learning finds applications in various image classification tasks, such as object identification, image extraction, semantic segmentation, and gesture estimation. In this study, we aim to compare the differences between existing famous works on fake news detection using the same dataset and the results obtained from our new proposed approach in section 3.
(Karimi et al., 2018) Introduce an approach to combine information from multiple sources and distinguish between different degrees of fakeness. They propose a Multi-source, Multi-class Fake News Detection framework (MMFD). The proposed system combines automated extraction features, multi-source fusion, and fakeness detection into a coherent and interpretable model. Experimental results demonstrate the viability of the proposed framework, and extensive experiments are conducted to gain insights into its

workings.

4. Proposed approach

4.1 Aim of the study

Specifying the meaning and type of information required to classify an item as fake news is crucial. Moreover, specifying the form of fake news is beneficial, considering the various types ranging from low to high. Any news analysis must rely on a formal classification of incorrect information, including propaganda, satirical information, manipulation, and manufacturing. However, our primary interest lies not in developing a general classification procedure but in building an automatic algorithm capable of multi-classifying any news while providing a specific percentage for each form.

To achieve our proposed approach's goal, we encounter two key challenges:

1. Differentiating between various forms of fake news.
2. Implementing multi-class classification using unlabeled data.
   a. Estimating labels based on similarity to enhance and improve the self-training algorithm.
   b. Comparing the newly estimated labels using similarity with the new labels predicted by the voting majority.

4.2 Methodology and approach

This section outlines the comprehensive approach (refer to Fig. 4), which encompasses three main categories:

1. Input Phase: It comprises two significant steps: Data collection of news (labeled and unlabeled data) and the pre-processing using NLP.
2. Pre-processing phase: It contains three major steps: Vectorization, Recommender System, and Multi-class classification.
3. Output phase

4.2.1 Input phase:

In the Data collection phase, we gather both labeled and unlabeled news articles from various sources.Labeled data refers to news articles that have been manually classified into different categories, such as satire, propaganda, manufacturing, and manipulation. These labels serve as the ground truth for training and evaluating our models. Unlabeled data, on the other hand, consists of news articles that have not been classified into any specific category.

The pre-processing step involves preparing the collected news articles for further analysis and classification using Natural Language Processing (NLP) techniques. In this phase, we perform various operations to clean, normalize, and transform the text data, making it suitable for machine learning algorithms.

The pre-processing steps typically include:

- Text Cleaning: Removing irrelevant or noisy elements, such as HTML tags, special characters, punctuation, and numbers.
- Tokenization: Breaking down the text into individual words or tokens to facilitate further analysis.
- Stopword Removal: Eliminating common words (e.g., "the," "is," "a") that do not contribute much to the overall meaning of the text.
- Lowercasing: Converting all words to lowercase to ensure uniformity and avoid treating words with different cases as distinct.
- Lemmatization or Stemming: Reducing words to their base or root form to reduce the vocabulary size and improve text representation.
- Removing Rare Words: Eliminating words that occur very infrequently, as they may not contribute significantly to the overall meaning.

After the pre-processing steps, we have a clean and transformed dataset that can be used for feature extraction and subsequent classification. For feature extraction, we employ the Word2Vec word embedding technique, which converts words into dense vector representations. These vectors capture the semantic meaning of the words and their contextual relationships, enabling better text representation for classification tasks. Overall, the data collection and pre-processing phase plays a crucial role in preparing the input data for our multi-class semi-supervised approach, ensuring that the data is in a suitable format for further analysis and model training.

4.2.2. Pre-processing phase:

This phase involves three main steps: Vectorization, Recommender System, and Multi-class classification.

- Vectorization:
  In the Vectorization step, both labeled and unlabeled news data are transformed into numerical vectors. Each news article's text is converted into a vector representation. We utilize the widely adopted word embedding technique, Word2Vec, as it offers a common and effective way to represent textual vocabulary. Word2Vec can capture the semantic meaning of words in the context of a text, capturing textual and syntactic similarities, as well as the relationships between different words. By using Word2Vec, we can create meaningful and compact representations of the news articles, which will be used as input for our subsequent classification process.
- Recommender System
  In the Recommender System phase, we perform a pre-multi classification process consisting of two steps:
- **Similarity recommendation:** In this step, we predict the label for each unlabeled news text by calculating its similarity to the entire labeled dataset. We use a similarity measure to determine how closely each unlabeled text aligns with the labeled data. The text with the highest similarity score is considered the most similar to a labeled text, and its predicted label is assigned accordingly.
- **Majority voting recommendation:** To further refine our predictions, we train a model using the labeled data and then use this trained model to predict labels for the unlabeled data using a voting majority approach. This means that we apply multiple algorithms to make predictions for each unlabeled text, and the final label is determined by selecting the most frequently predicted value among these algorithms.
- **Multi-class classification phase**: we utilize the recommendations obtained from the previous steps, which include two label suggestions for each news article: one based on similarity and the other based on voting majority. News articles identified as similar to the labeled data, their corresponding recommended labels are incorporated into the labeled dataset, effectively creating a new pseudo dataset. This process enriches the labeled dataset with additional samples, enhancing the diversity and representativeness of the training data. However, for news articles deemed dissimilar to the labeled data, we employ the self-training algorithm as the third approach to predict their labels. The self-training algorithm utilizes the information from the labeled dataset to make predictions for the unlabeled data, enabling us to assign labels to these non-similar articles. By combining these different strategies, our multi-class classification approach aims to effectively handle diverse news articles, leveraging similarity-based recommendations and self-training predictions to improve the accuracy and comprehensiveness of the final classification results.

By employing the Recommender System, we aim to improve the self-training algorithm's effectiveness and obtain more accurate and reliable multi-class classification results for news articles.

4.2.3. Output phase:

In the final phase, the proposed approach provides the predicted classes of the given news text and the corresponding percentage of certainty for each class.

Self-training is a powerful learning method that effectively handles situations with limited labeled and abundant unlabeled data. However, it is often observed that the accuracy of applying the unlabeled data in addition to the labeled data is lower than using only the labeled data. One of the main reasons for this is the inadequacy of the labeled data to train the initial classifier in the self-training phase. An inefficient initial classifier can introduce mislabeled data, which is then utilized to train the final classifier, leading to a decline in the precision of the semi-supervised self-training classifier. To address this issue, we propose a novel approach to ensure the newly labeled data's reliability for training the final classifier.

This section presents a method for constructing a multi-class classification analysis model (Fig. 3). The process of the proposed methodology is divided into the following steps:

- Step 1- Similarity Phase: Each unlabeled news item is transformed into a vector using word2Vec. Then, we calculate the similarity percentage by employing the Cosine Similarity between each unlabeled item and the entire labeled dataset. The most similar item (with the maximum similarity value) is identified, indicating that this unlabeled item has a percentage of similarity in terms of having the same type as the labeled item it resembles.

- Step 2- Voting Majority Phase: Initially, the model is trained on the labeled data, and then we utilize this pre-trained model to predict labels for the unlabeled data using five classifiers. The label predicted most frequently among the five classifiers is assigned to the unlabeled news item as a second prediction, which we call the voting majority.
- Step 3- Pre-Multi Classification Phase: In this step, we have two predicted types for each news item: one based on similarity with its respective percentage of similarity and the other based on the voting majority.
- Step 4- Recommendation Phase: News items with the same predicted types (i.e., the predicted type found in the similarity phase matches the one from the voting majority phase) are added to the labeled dataset. This phase enhances confidence in the items we will add to the dataset.
- Step 5- Multi-class Classification Phase: In this final step, we apply the self-training algorithm to the data, performing multi-class classification further to improve the accuracy and reliability of the predictions.

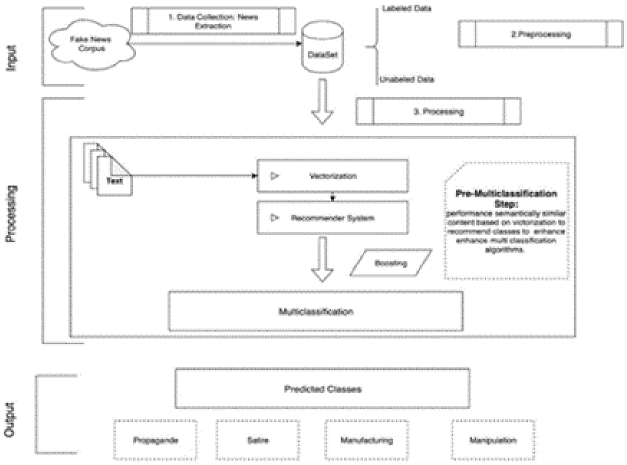


Fig. 2: The overall approach

.

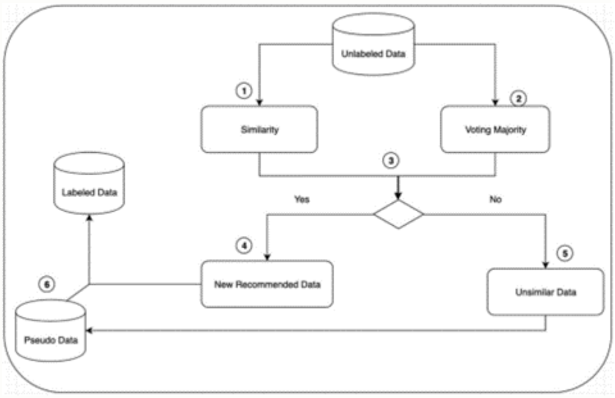


Fig. 3: Process of proposed approach

5. Experiments and results

5.1 Data collection process

The topic of fake news identification lacks standard benchmark datasets, primarily due to the term "fake news" encompassing various subcategories. Our study utilized two kinds of fake news identification datasets (Table 4). The first dataset contains three types of fake news: satire, propaganda, and manufacturing. The second dataset focuses on manipulation or bias form. We combined both datasets in a balanced manner. Each article in the corpus includes the title text and label. The corpus consists of 443 news articles for each label: satire, propaganda, manufacturing, and manipulation.

5.1.1 Dataset

The absence of manually labeled fake news datasets poses a significant challenge to the advancement of computationally expensive, text-based models that cover a wide range of topics. For our research purposes, we require a set of news articles that are directly classified into news types, such as satire, propaganda, manufacturing, and manipulation. We thoroughly searched for available datasets containing these news categories to address this issue. As a result, we found two datasets and combined them to create a comprehensive dataset with all four categories, facilitating multi-class classification.

Table 4: Sample fake news dataset.

References	Size	Date	Text
(Fact	12999	2017	This research analyzes

Checking)	rows		inaccurate news sources and truthful claims from politi-fact.com using tools from a previous EMNLP'17 paper
(Getting real about fake news)	38859 rows	2016 -11- 25	The study includes recent reports on non-fake news to emphasize the complexity of addressing inaccurate reporting and seeks better solutions than blacklists.

5.2. Data pre-processing

Pre-processing data is a typical first step that precedes training and evaluating data using machine learning algorithms. Ensuring the data is appropriately formatted and meaningful elements are integrated is crucial to achieving accurate and optimal outcomes. Our pre-processing of the data involved an iterative process divided into three main stages. Each incremental step corresponds to the models trained and evaluated on the pre-processed data at that stage. Moreover, each step builds upon the previous ones, with the second step including the first pre-processing stage and the third step including the first two pre-processing processes.

The first step is easy pre-processing, followed by the second, which involves removing all non-English phrases. Finally, the last step entails removing the end of the guardian posts, consistently including the exact phrase: "Share on x, y, z."

5.3. Experiment results

Our experimental results were obtained through a three-phase procedure. In the first phase, we calculated the similarity for each row in a small amount of labeled data and a large amount of unlabeled data, assigning a percentage of similarity. Table 5 shows the accuracy achieved after applying the new labels to the unlabeled data using similarity.

We trained the labeled data in the second phase as mentioned in Table 6 using different classifiers (Logistic regression, Naive Bayes, Linear SVM, and Decision Tree). We then used this model to predict the labels for the unlabeled data again, this time using a voting majority approach. We calculated the new percentage of recommendations by subtracting 100 from the similarity percentage found in the first phase. This phase is recommended for comparing the labels with the highest similarity percentage predicted in the first step with the labels predicted in this phase using a voting majority. Table 7 displays the results of this phase.

Moving on to the third phase, we combined the similar labels predicted in the preceding phases and added them to the labeled dataset, creating a new pseudo dataset. We applied the self-training algorithm to predict dissimilar labels. Table 8 shows the results obtained in this phase. Finally, we checked if the newly predicted labels matched those in the second phase.

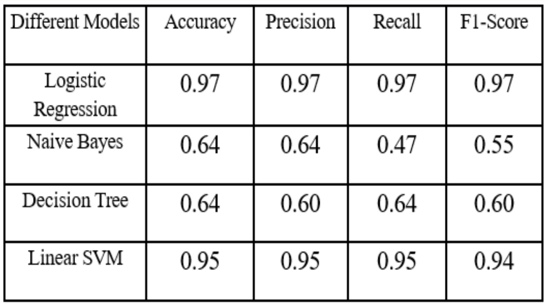
Table 5: Experimental results for the first phase.

Different Models	Accuracy	Precision	Recall	F1-Score
Logistic Regression	0.76	0.76	0.76	0.76
Naive Bayes	0.44	0.69	0.44	0.39
Decision Tree	0.41	0.35	0.41	0.35
Linear SVM	0.73	0.74	0.73	0.73

Table 6: Experimental results for the second phase.

Different Models	Accuracy	Precision	Recall	F1-Score
Logistic Regression	0.93	0.93	0.93	0.91
Naive Bayes	0.71	0.73	0.71	0.66
Decision Tree	0.74	0.62	0.74	0.67
Linear SVM	0.94	0.94	0.94	0.93

Table 7: Experimental results for the third phase.



Different Models	Accuracy	Precision	Recall	F1-Score
Logistic Regression	0.97	0.97	0.97	0.97
Naive Bayes	0.64	0.64	0.47	0.55
Decision Tree	0.64	0.60	0.64	0.60
Linear SVM	0.95	0.95	0.95	0.94

Table 8: Performance comparison for fake news detection

Metric	(Collins et al.,2020).	(Wu et al.,2018)	(Stitini et al.,2022)	Our proposed approach
Accuracy	89%	33%	96%	97%
Precision	89%	-	96%	97%
Recall	89%	60%	96%	97%
F1-Score	89%	49%	96%	97%

6. Conclusion

In this paper, we introduce a novel multi-class semi-supervised approach for self-training, which is trained using a limited collection of classified data and an extensive amount of unlabeled data. Our innovative solution incorporates a similarity algorithm to enhance the self-training process, ensuring new expected labels are applied to the labeled data.

To evaluate the efficiency of the proposed semi-supervised method, we conducted tests on two benchmark datasets, measuring the classification precision using commonly available simple learners such as logistic regression, decision tree, naive Bayes, and linear SVM. The numerical findings validate the effectiveness and robustness of our approach. Consequently, our method contributes to developing more effective, reliable, and robust predictive models for multi-class fake news classification.

References

Kumar, S., West, R., & Leskovec, J. (2016). Disinformation on the Web: Impact, Characteristics, and Detection of Wikipedia Hoaxes. Proceedings of the 25th International Conference on World Wide Web.

Qian, F., Gong, C., Sharma, K., & Liu, Y. (2018). Neural User Response Generator: Fake News Detection with Collective User Intelligence. https://doi.org/10.24963/ijcai.2018/533

Collins, B., Hoang, D. T., Nguyen, N. T., & Hwang, D. (2020). Trends in combating fake news on social media – a survey. Journal of Information and Telecommunication, 1–20. https://doi.org/10.1080/24751839.2020.1847379

Vijayaraghavan, S. (2020, February 15). Fake News Detection with Different Models. arXiv.org. https://arxiv.org/abs/2003.04978

Li, Q., Zhang, Q., Si, L., & Liu, Y. (2019). Rumor detection on social media: datasets, methods, and opportunities. https://doi.org/10.18653/v1/d19-5008

Alzanin, S. M., & Azmi, A. M. (2018). Detecting rumors in social media: A survey. Procedia Computer Science, 142, 294–300. https://doi.org/10.1016/j.procs.2018.10.495

Wang, W. Y. (2017). “Liar, liar Pants on Fire”: a new benchmark dataset for fake news detection. https://doi.org/10.18653/v1/p17-2067

Wu, L., Rao, Y. J., Yu, H., Wang, Y., & Nazir, A. (2018). False information detection on social media via a hybrid deep model. In Lecture Notes in Computer Science (pp. 323–333). https://doi.org/10.1007/978-3-030-01159-8_31

Imran, M., Castillo, C. F., Diaz, F., & Vieweg, S. (2015). Processing social media messages in mass emergency. ACM Computing Surveys, 47(4), 1–38. https://doi.org/10.1145/2771588

Oumaima, S., Soulaimane, K., & Omar, B. (2020). Latest Trends in Recommender Systems applied in the medical domain. https://doi.org/10.1145/3386723.3387860

Tanha, J. (2019). A multiclass boosting algorithm to labeled and unlabeled data. International Journal of Machine Learning and Cybernetics, 10(12), 3647–3665.https://doi.org/10.1007/s13042-019-00951-4

Stitini, O., Kaloun, S., & Bencharef, O. (2022). Towards the detection of fake news on social networks contributing to the improvement of trust and transparency in recommendation systems: Trends and challenges. Information, 13(3), 128. https://doi.org/10.3390/info13030128

Stitini, O., Kaloun, S., & Bencharef, O. (2022a). Integrating contextual information into multi-class classification to improve the context-aware recommendation. Procedia Computer Science, 198, 311–316. https://doi.org/10.1016/j.procs.2021.12.246

Silva-Palacios, D., Ferri, C., & Ramírez-Quintana, M.J. (2017). Improving performance of multiclass classification by inducing class hierarchies. Procedia Computer Science, 108, 1692–1701. https://doi.org/10.1016/j.procs.2017.05.218

Rasool, T., Butt, W. H., Shaukat, A., & Akram, M. U. (2019b). Multi-Label Fake News Detection using Multi-layered Supervised Learning. https://doi.org/10.1145/3313991.3314008

Jedrzejowicz, J., Kostrzewski, R., Neumann, J., & Zakrzewska, M. (2018). Imbalanced data classification using MapReduce and Relief. Journal of Information and Telecommunication. https://doi.org/10.1080/24751839.2018.1440454

Forestier, G., & Wemmert, C. (2016). Semi-supervised learning using multiple clusterings with limited labeled data. Information Sciences, 361–362, 48–65. https://doi.org/10.1016/j.ins.2016.04.040

Karimi, H. R., Roy, P. C., Saba-Sadiya, S., & Tang, J. (2018). Multi-Source Multi-Class fake news detection. In International Conference on Computational Linguistics (pp. 1546–1557). https://www.aclweb.org/anthology/C18-1131.pdf

Kaliyar, R. K., Goswami, A., & Narang, P. (2019). Multiclass Fake News Detection using Ensemble Machine Learning. https://doi.org/10.1109/iacc48062.2019.8971579

Li, J., & Zhu, Q. (2019). Semi-Supervised Self-Training method based on an Optimum-Path forest. IEEE Access, 7, 36388–36399. https://doi.org/10.1109/access.2019.2903839

Wu, D., Shang, M., Wang, G., & Li, L. (2018). A self-training semi-supervised classification algorithm based on density peaks of data and differential evolution. https://doi.org/10.1109/icnsc.2018.8361359

Martineau, M., Raveaux, R., Conte, D., & Venturini, G. (2020). Learning error-correcting graph matching with a multiclass neural network. Pattern Recognition Letters, 134, 68–76. https://doi.org/10.1016/j.patrec.2018.03.031

Yang, P., Zhao, P., Hai, Z., Liu, W., Hoi, S. C. H., & Li, X. (2016). Efficient multi-class selective sampling on graphs. In Uncertainty in Artificial Intelligence (pp. 805–814). http://auai.org/uai2016/proceedings/papers/34.pdf

Kaneko, T. (2019, February 4). Online multiclass classification based on prediction margin for partial feedback. arXiv.org. https://arxiv.org/abs/1902.01056

Gertrudes, J. C., Zimek, A., Sander, J., & Campello, R. J. G. B. (2019). A unified view of density-based methods for semi-supervised clustering and classification. Data Mining and Knowledge Discovery, 33(6), 1894–1952. https://doi.org/10.1007/s10618-019-00651-1

Larriva-Novo, X., Sánchez-Zas, C., Villagrá, V. A., Vega-Barbas, M., & Rivera, D. (2020). An approach for the application of a dynamic Multi-Class classifier for network intrusion detection systems. Electronics, 9(11),1759. https://doi.org/10.3390/electronics9111759

Deepti Nikumbh, Anuradha Thakare (2023). A comprehensive review of fake news detection on social media: feature engineering, feature fusion, and future research directions, International Journal of Systematic Innovation, 7(6), 36-53. DOI: 10.6977/IJoSI.202306_7(6).0004

Livieris, I. E., Kanavos, A., Tampakas, V., & Pintelas, P. E. (2018). An Auto-Adjustable Semi-Supervised Self-Training Algorithm. Algorithms, 11(9), 139. https://doi.org/10.3390/a11090139

Hyams, G. (2017, September 30). Improved training for Self-Training by Confidence assessments. arXiv.org. https://arxiv.org/abs/1710.00209

Piroonsup, N., & Sinthupinyo, S. (2018). Analysis of training data using clustering to improve semi-supervised self-training. Knowledge Based Systems, 143, 65–80. https://doi.org/10.1016/j.knosys.2017.12.006

Xing, Y., Yu, G., Domeniconi, C., Wang, J., & Zhang, Z. (2018). Multi-Label Co-Training. https://doi.org/10.24963/ijcai.2018/400

Rezende, D. J., Mohamed, S., Danihelka, I., Gregor, K., & Wierstra, D. (2016). One-shot generalization in deep generative models. In International

Conference on Machine Learning (pp. 1521–1529). http://jmlr.org/proceedings/papers/v48/rezende16.pd f

Wang, L., Ding, Z., & Fu, Y. (2018). Adaptive Graph guided embedding for multi-label annotation. https://doi.org/10.24963/ijcai.2018/388

Nakano, F. K., Cerri, R., & Vens, C. (2020). Active learning for hierarchical multi-label classification. Data Mining and Knowledge Discovery, 34(5), 1496–1530. https://doi.org/10.1007/s10618-020-00704-w

Fact Checking. https://hrashkin.github.io/factcheck.html Accessed: 2021-02-24

Getting Real about Fake News. https://www.kaggle.com/datasets/mrisdal/fake-news . Accessed: 2021-02-24

Stitini, O., Kaloun, S., & Bencharef, O.: Investigating different similarity metrics used in various recommender systems types: scenario cases, Int. Arch.Photogramm. Remote Sens. Spatial Inf. Sci., XLVIII-4/W3-2022, 187–193, https://isprs-archives.copernicus.org/articles/XLVIII-4-W3-2022/187/2022/

AUTHOR BIOGRAPHIES

Oumaima Stitini. She is currently an Assistant Professor at the Cadi Ayyad University, ENS. She received his Ph.D. in Computer Science Engineering, especially in Robust optimization and implementation of recommendation systems based on artificial intelligence from the Faculty of Science and Technology in 2022.

His main research interests include artificial intelligence, recommender systems, and IoT all of these applied to different fields like medical, education, and entertainment.

Sara Qassimi holds a Ph.D. and an engineering degree in computer science from Cadi Ayyad University, Marrakesh, Morocco. She is currently an Assistant Professor in the Department of Computer Science at the Faculty of Sciences and Technics Guiliz, and is affiliated with the L2IS Laboratory at Cadi Ayyad University. Her research interests focus on artificial intelligence, recommender systems, context-aware systems, social interactions, and knowledge graphs.

Soulaimane Kaloun He is currently holding the position of a Permanent Associate Professor at the Faculty of Science and Technology located in Marrakech, Morocco. He earned his doctorate degree in Computer Science and is presently serving as a professor at the same institution. Moreover, he has also received an HDR in data science. Soulaimane's principal areas of research revolve around Big Data, machine learning, multiagent systems, and text-mining.

Omar Bencharef. He is currently a Permanent Professor at the Faculty of Science and Technology, Marrakech, Morocco. Omar received his Ph.D. in Computer Science and is currently a professor at the Faculty of Science and Technology, Marrakech, Morocco. He’s also HDR in data science. His main research interests include artificial intelligence (AI), data science, machine learning, multiagent systems, and text-mining.

Works	Semi-supervised learning approach						Clustering approach			Deep Learning approach		
	Self-training	Co-training	GMM	MM	GM	BM	PM	HM	DM	GrM	CNN	RNN
(Tanha,2019).						✓						
(Kaliyar et al., 2019)						✓					✓	
(Silva-Palacios et al., 2017).						✓		✓				
(Li et al., 2019)	✓											
(Wu et al., 2018)	✓											
(Martineau et al., 2020)			✓		✓						✓	
(Yang et al., 2016)					✓							
(Kaneko, 2019)							✓					
(Larriva-Nov o et al., 2020										✓		
(Gertrudes et al., 2019)									✓			
(Deepti Nikumbh, et al., 2023)		✓										
(Karimi et al., 2018)												✓
(Kaneko et al., 2019)				✓								

Note: GMM refers to Generative models methods, MM refers to Margin based methods, GM refers to Graph based methods, BM refers to Boosting methods, PM refers to Partitioning Methods, HM refers to Hierarchical Methods, DM refers to Density-based Methods, GrM refers to Grid-based Methods, CNN refers to Convolutional neural networks, RNN refers to Recurrent Neural Network.

DOI: 10.6977/IJoSI.202403_8(1).0003
M. N. P. Ma'ady/Int. J. Systematic Innovation, 8(1), 27-35 (2024)

Strengthening research partner collaboration in higher education for searching innovation through machine learning-based recommender system

Mochamad Nizar Palefi Ma'ady
Department of Information Systems, Institut Teknologi Telkom Surabaya, Indonesia
Corresponding author E-mail: nizar@ittelkom-sby.ac.id

(Received 03 June 2023; Final version received 30 November 2023; Accepted 30 November 2023)

Abstract

Academic collaboration is tremendously important for higher education. Multidisciplinary academicians may be grouped as a better research collaboration than the previous one. Therefore, such a system is needed, even for a huge number of academicians in an institution. However, existing such recommendation tools are expensive. This paper suggests to develop a system by using machine learning approach in order to search a big academicians data effectively. Hence, with help of the standard of Naïve Bayes creates a flexible text search without depending on what select options including research location or case study instead of only research topic. Furthermore, the output of Naïve Bayes, then, is tranformed to percentage display in order to bring ease of understanding the gap of recommendation. It allows the user to choose a possible partner more than one. Therefore, this approach helps reduce time and effort.

Keywords: Higher Education, Naïve Bayes Algorithm, Recommendation System, Research Partner Collaboration, Sigmoid Activation Function.

1. Introduction

Higher education is the most important of creating cutting-edge technology for solving kinds of problems in a region in which the existence of higher education is impacted to the region itself. Therefore, academic research collaboration involved within higher education is the vanguard of the solution. Creating the best plausible collaboration is highly needed amongst faculties or in multidisciplinary (Samin & Azim, 2019). There is a strong positive relationship and a better collaboration providing a better outcome. Hence, it increases scientific publications (Amarante et al., 2021; Volkwein & Parmley, 2000). Co-authorship in the same university is important for improving research performance and better learning outcomes, academic writing, research publication skills, productivity, and time management (Abbas et al., 2020; Aldieri et al., 2019).

Private and public universities need to gain research atmospheric among academia with the merit of funding from the university itself or external funding (Abramo et al., 2010). But a large number of faculties may bring challenges in order to find a suitable research partner. The sheer volume of information of partner candidates also brings difficulty. In addition, in developing countries, the number of lecturers in a private university is relatively fluctuating that is because of factors like high living costs, low wages, etc. (Ramadhan & Putri, 2018). Therefore, it needs to manage such unstructured information datasets into a well-organized system. We propose a recommendation system for finding the most suitable research partner based on topic, object location, case study, or any else.

Researchers have recognized the existence of specialization of their research area (Khalid et al., 2011). This paper puts forward why enabling researchers to find partner collaboration by dynamic text search is important. In order to do so, we use the standard Naïve Bayes algorithm for classification task. Naïve Bayes has been widely used in solving text classification problems such as personality classification, complaint-level classification, spam classification, news categorization, etc. This algorithm shows the efficiency in practice. However, in higher education context, the output needed may be different.

In higher education environment, a researcher as user should find alternative partner candidates instead of only the best recommendation. With gauge information of how close the candidates to the criteria is, in the system, a user is able to set more than one criteria based on what research skills are needed. Then the output has to show a precise value compared to other candidates. According to the output, the user has the knowledge to see how strong the candidate is by looking at the value gap among candidates. Therefore, in higher education context, especially in developing universities, relying on Naïve Bayes classification is insufficient.

Recommendation System (RS) is typically used in lexical text from corpus data to cope amount of information just simply suggested for items based on user preferences. In higher education, there have been conducted applications of RS such as recommendation system on course selection and topic recommendation system, while text lexicon in higher education used for prediction of student placement and student retention (Cardona et al., 2020; Di Sipio et al., 2020; Guruge et al., 2021; Wadekar et al., 2018). In this paper, the algorithm is required to generate the amount of information to provide the best rank of suggestions (Saleh et al., 2015) [14]. For this reason, we are more interested in leveraging recommendation system along with its appropriate features using a standard Naïve Bayes. In addition, in some private universities, the problem of searching for research collaboration needs to be analyzed. It is because of the high circulation of lecturers entering or leaving private universities in Indonesia (Ramadhan & Putri, 2018).

Recommendation or suggestion from plenty of unstructured and raw text data is to provide a shortcut by giving the most related information. Hence, it is important in terms of saving time and energy. Our main focus is to analyze the problem in the context of higher education and to help academia to find research partners as preferences. As the aforementioned problem, in the spirit of increasing user intention to use, the RS should also pay attention to the ease of use and reliability of the system. Hence, feature selection of the algorithm preprocess is essential. Naïve Bayes is one of the widely used algorithms in developing the field of RS. The algorithm works on item features (Gaikwad et al., 2018). In recent years, content based RS has been applied for various uses such as recommendations on documents or news, while Syskill and Webert recommendation on web pages or personalized television programs (Cotter & Smyth, 2000; Pazzani et al., 1996). Based on the problem context, we propose a nonlinear approach to gain academic intention of use.

This paper concisely contributes: to leverage Sigmoid activation function to transform Naïve Bayes output into a range of 0-1 and multiplied by 100 to provide a percentage interpretation instead of just using Naïve Bayes classification in order to recommend the most suitable research partners; and to develop the proposed research partner recommendation system in web-based application with respect to user interface and user experience analysis, and to share the source code program. We organized the paper as follows. First, we review the system from the literature in section 2, while section 3 formulates the problem context. Section 4 evaluates the performance of the proposed system. Last, the conclusion of the paper in section 5.

2. Literature review

Naïve Bayes Classification is commonly used as statistical-based technique in content-based RS (Guruge et al., 2021). Content-based RS leverages statistical-based RS to improve valid recommendations. The other techniques are such as TF-IDF, decision trees, and artificial neural networks (Shah et al., 2016). This study uses content-based RS by Naïve Bayes algorithm in the context of classification at first. Then, the classifier is used to estimate the probability of researcher candidate that is the relevancy from text-based dataset. The generated output used keywords of the articles as inputs. For instance, Fab system uses 100 features of web pages to users for representing contents of the web pages (Pavlov & Pennock, 2002).

Neamah & El-Ameer (2018) generate content-based Naïve Bayes for course recommendations. The systems build with use case of course enrollment and ranking for user profile, while Ghani & Fano (2019) enable product recommendations by categorizing products from a department store. Miyahara & Pazzani (2000) implemented recommendations based on number of likes and dislikes, while Sipio et. al. based on GitHub repositories. To do so, Multinomial Naïve Bayesian network is conducted. Fan J., Zhou W., and Yang X (2019) introduce improving quality and quantity of recommendations by sharing content ratings on online social networks. And, content-based personalized recommendation using Bayesian hierarchical models is implemented by Zhang & Koren (2007) to recommend Netflix and MovieLens movies.

Activation function is commonly used in nonlinear optimization field, especially in machine learning. It plays an important role in machine learning as increasing the performance of classifier. Sigmoid activation function can transform a rough value into a range of 0 to 1 which is a useful approach for many applications (Eger et al., 2018). Various activation function such as sigmoid, relu, tanh, and step function show different behaviors whereas sigmoid used at most (Ramachandran et al., 2017). De Campos (2006) proposed a Bayesian approach in the representation of new assessments by applying Zadrozny's network to convert classifier into scoring (Zadrozny & Elkan, 2000). Tripathi et al. (2020) utilized Bayes as a classification in credit scoring.

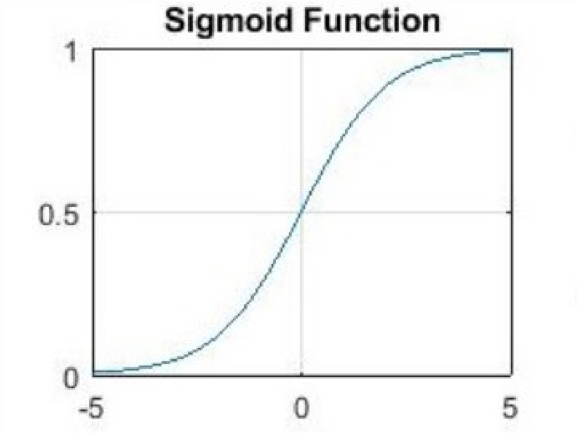


Fig. 1. Sigmoid activation function

Sigmoid activation function lies in the range of 0 to 1. When it reaches 0.5 then it converts to zero as shown in Fig. 1 (Nwankpa et al., 2018). It then defines real input values using derivatives with multiple degrees. Therefore, active research tends to apply to large-scale tasks. Ivaschenko & Milutkin (2019) applied the activation function based on NLP by providing a better accuracy. Hence, this system is also recommended in the context of human resources who looks for the best candidate based on their blogs and online presence. This approach is adopted with the preference of research topics, study objects, and case studies. Since the output is linear, a nonlinear function is required to convert it into a range of 0 to 1.

3. Method

This paper identifies a common problem in higher education when a university has a number of lecturers, there exists a difficulty in finding the best research collaboration. The proposed RS offers alternative research partners based on generated input data. The system consists of two phases. The first step is to use machine learning algorithm namely Naïve Bayes in order to classify research topics as input into which class of academicians. Second, to display the output score in percentage uses Sigmoid activation function as details in Figure. 3. This paper provides a specific case problem of higher education as shown in Table 1. As aforementioned problem, the proposed system conducts a trial for the lecturers at the Faculty of Information Technology and Business in Indonesia. Hence, we can observe what problem statements should be fulfilled by the system.

Table 1. Problem identification in higher education.

Problem statement	System requirement
Academicians in a private university is a high variety of research areas which not focuses in certain research areas.	System must able to synthesis from various research topics.
Articles are published in national journal which provided in Indonesian language.	System needs to process Indonesian corpus.
Input text field is in dropdown list or check list.	System displays input fields in text input.
Academicians are unhappy with the output in the shape of classes.	System displays output in percentage mode.

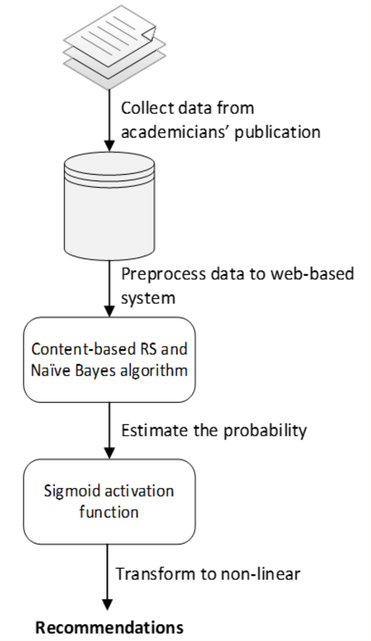

Collect data from academicians' publication
Preprocess data to web-based system
Content-based RS and Naïve Bayes algorithm
Estimate the probability
Sigmoid activation function
Transform to non-linear
Recommendations

Fig. 2. Machine learning-based system for research recommendation.

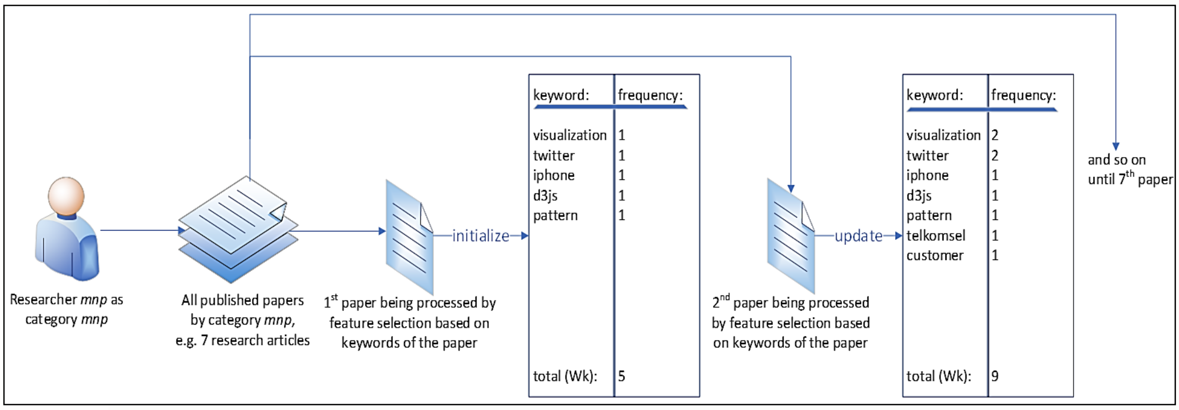

Researcher *mnp* as category *mnp*
All published papers by category *mnp*, e.g. 7 research articles
1st paper being processed by feature selection based on keywords of the paper
initialize
keyword:
frequency:
visualization 1
twitter 1
iphone 1
d3js 1
pattern 1
total (Wk): 5
2nd paper being processed by feature selection based on keywords of the paper
update
keyword:
frequency:
visualization 2
twitter 2
iphone 1
d3js 1
pattern 1
telkomsel 1
customer 1
total (Wk): 9
and so on until 7th paper

Fig. 3. Proposed algorithm workflow.

In Fig. 3, the process starts from Google Scholar site by collecting all academic articles published by the faculty. Hence, the data collected includes national, international journals and proceeding articles. Due to dimensionality reduction issue, this paper determines three features; research topic, object or location, and case study. RS is an efficient approach in reducing the time while the traditional approach is time-consuming. Feature selection obtains appropriate keyword sets representing its article topics.

Then, designing the database is based on entity relational diagram and text classification. The essential of the proposed systems enforces researcher as output. Hence, all features must satisfy belong the article publication. For instance, if the system generates visualization as a topic, then it must be defined belongs to which researcher. An article can produce some topics that represent to a respected researcher. In order to gain user intention to use, this paper also considers to design UI/UX that copes above system requirements. Therefore, UI/UX literature adopts usability and responsive web quality principles.

The main flow of the proposed research partner recommendation system initializes selected P from database. The algorithm process uses researcher as a class or label. Then, formula (1) generates any relevant input data, oldScore, as in Fig. 3, by summing the probability of formula (1) with bias *b*. RS takes uses it to rank a list of recommendation. Recommendation system results high score for any profile which is mostly called from inputs.

$$P\left(W_k \mid S_j\right) = \frac{\mid f\left(w_k \middle| s_j\right) \mid}{\mid w_k \mid} \qquad (1)$$

Algorithm: Proposed RS

```
procedure rs on naïve bayes algorithm
initialize a set of keywords $P = \{x_1, \ldots, x_N\}$;
for each researcher $S_i$, j = {1, …, J}:
|   do declare bias $b = 1$;
|   compute $P(W_k \mid S_j) = \frac{\mid f(w_k \mid s_j) \mid}{\mid w_k \mid}$, where
|                 k = number of papers of $S_{i.}$
|                 $f(W_k \mid S_i)$ = total keyword occurrences of $W_k$,
|                 $W_k$ = total keywords of $S_i$ in training data;
|   compute $oldScore(S_j) = \sum_{k=1}^{K} P(Wk \mid Sj) + b$;
end
procedure sigmoid activation function //as nonlinear approach
for each researcher $S_j$, j = {1, …, J}:
|   do compute $newScore(S_j) = \frac{1}{1+e^{-oldScore(Sj)}}$;
|   if $newScore(S_j) \le 0.5$:
|       $newScore(S_j) = 0$;
|   else:
|       return $newScore(S_j)$ * 100; //convert to percentage form
end
```

Fig. 4. Pseudocode for the proposed algorithm.

Moreover, because the output values shown are not satisfied by academia, the system may decrease potential user to create a new research collaboration. Therefore, a nonlinear approach is needed. If input data is in non-negative values, Sigmoid activation function will generate them below 0.5. Then, users are not interested in recommendation values that are below the 50% threshold, but rather they value values higher than 50%. Hence, we can set it to zero with the exception of the algorithm (see Fig. 4 for the pseudocode).

$$Sig(S_j) = \frac{1}{1 + e^{-oldScore(Sj)}} \qquad (2)$$

For example demonstration, the input data as shown in Fig. 5 processed by Naïve Bayes algorithm by matching the input words to Naïve Bayes features database. Suppose the database contains three researchers, then the input data will be compared to which researcher is. Based on Table 2, the total features of mnp researcher is 40 with the use of numerator f(Wk|Sj). Selected total feature means extracted keywords from the paper publication, for instance, researcher mnp being denominator

Wk. The keyword is updated for the next processing paper. The process stops as all researcher publications have been processed.

Input text: *IoT, visualization, Surabaya, Application*

Fig. 5. Example of input text.

This result score satisfies system requirements as user may see the different score percentages. Researcher zul with score of 85.96% is the most relevant researcher for the example. Therefore, it does not classify the score but shows the percentage score of the relevancy. In this representation, users may desire to collaborate with more researchers regarding the percentage of relevancy. For example, a user is suggested to invite a researcher with a score 82.24% for the research collaboration. In terms of building the proposed program, this paper designs the relational database to see what are the features and labels in the system. The database consists of three tables with a primary key on lecturer profile table onto the department table. This purpose is to display the proposed algorithm result using UI/UX design.

Table 2. Numerical computation of the example.

Candidates	$P(W_k \mid S_j)$			
	IoT	*visualization*	*Surabaya*	*Application*
mnp	$^{1}/_{40}$	$^{3}/_{40}$	$^{1}/_{40}$	$^{1}/_{40}$
pur	$^{5}/_{45}$	$^{3}/_{45}$	$^{12}/_{45}$	$^{7}/_{45}$
zul	$^{10}/_{32}$	$^{3}/_{32}$	$^{3}/_{32}$	$^{10}/_{32}$

P(mnp|document) $= {}^{1}/_{40} + {}^{3}/_{40} + {}^{1}/_{40} + {}^{1}/_{40} + 1$

$= 1.15$

newScore(mnp) $= \dfrac{1}{1 + e^{-1.15}}$ x 100 %

= 75.95 %

P(pur|document) $= {}^{5}/_{45} + {}^{3}/_{45} + {}^{12}/_{45} + {}^{7}/_{45} + 1$

$= 1.6$

newScore(pur) $= \dfrac{1}{1 + e^{-1.6}}$ x 100 %

= 82.24 %

P(zul|document) $= {}^{10}/_{32} + {}^{3}/_{32} + {}^{3}/_{32} + {}^{10}/_{32} + 1$

$= 1.8125$

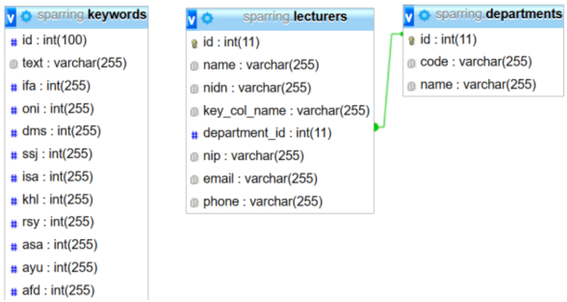
newScore(zul) $= \dfrac{1}{1 + e^{-1.8125}}$ x 100%

= 85.96 %

The keywords table is purposely to be training data where it is computed by Naïve Bayes and Sigmoid Activation Function. It is not related to the other tables because of different needs. However, this approach still depends on the collected data from Google Scholar manually. Hence, it also contains a number of lecturer names which are to be labeled or class of the Naïve Bayes output.

4. Result and discussion

Experiments were conducted on a Core-i7 and RAM 8 GB. Datasets were validated using cross-validation method, as detailed in the testing scenarios with some interesting points from the experiment in the subsection below. This paper emphasizes the use of nonlinear approaches in the field of RS. Some interesting results are highlighted. The system suggests not only the most recommended research partner but also alternative recommendations. This approach makes a possible new collaboration than one researcher. The interesting is as the case study at a private university with a majority of young researchers, denominator Wk in formula (1) inspires young researchers to create a specific research area rather than too many research topics. This is because the percentage can be decreased if the topic's relevancy is large. Hence, the generated features of the sample problem show high scores since it has only a few research topics. Another interesting point is the benefit of bias in the formula (1) can adjust the representation as time flies resulting huge number of researchers, which impacts to decrease in the score. It can be solved by simply change increase the bias value to be a better score representation.



sparring.keywords
id : int(100)
text : varchar(255)
ifa : int(255)
oni : int(255)
dms : int(255)
ssj : int(255)
isa : int(255)
khl : int(255)
rsy : int(255)
asa : int(255)
ayu : int(255)
afd : int(255)

sparring.lecturers
id : int(11)
name : varchar(255)
nidn : varchar(255)
key_col_name : varchar(255)
department_id : int(11)
nip : varchar(255)
email : varchar(255)
phone : varchar(255)

sparring.departments
id : int(11)
code : varchar(255)
name : varchar(255)

Fig. 6. Data flow diagram of the system.

This sub-section evaluates the performance of the proposed RS using k-cross validation. The designed experiment is conducted using binary classifications whose classes are determined by different researchers in a faculty. The dataset used from 26 publications for both researchers resulted in 94 features. As for performing exclusion cases, the threshold is important. If the input text does not belong to any class feature, then the data will not be used further or set as zero. Based on k-cross validation, the dataset was then folded into 80% for training data and 20% for testing data.

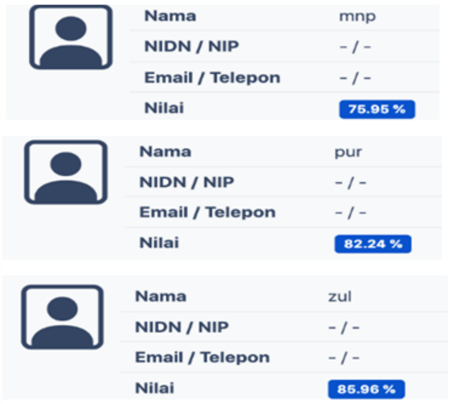
Table 3. Confussion matrix.

		Ground Truth	
		Researcher x	*Researcher y*
Classifier Machine	*Researcher x*	• tp = 16	fp = 1
	Researcher y	fn = 2	tn = 7

• true positive (tp): correctly predicted collaborators
true negative (tn): correctly predicted negative values
false positive (fp): collaborator is predicted but actual data shows prediction to be false
false negative (fn): the proposed system fails to produce an accurate prediction

$$Accuracy = \frac{tp + tn}{tp + tn + fp + fn} \qquad (3)$$

Table 3 shows the system processes items based on confussion matrix for performance evaluation. The standard formula of accuracy (3) is used for the case study of private university in Indonesia. The accuracy of the proposed system is 88.4%. In order to fulfill the system requirements, UI/UX design takes part to build the web-based system. The usability principle is important in the process of designing a software including (1) focusing on content by simplifying content layout, (2) recognition rather than recall in terms of providing search text fields in the results, (3) aesthetic and minimalist design to embrace neatness, (4) user assistance to recognize, diagnose, and recover from errors, and (5) providing help and documentation to the system. The user interface result of implementing all the principles as shown in Fig. 7 in the web version.



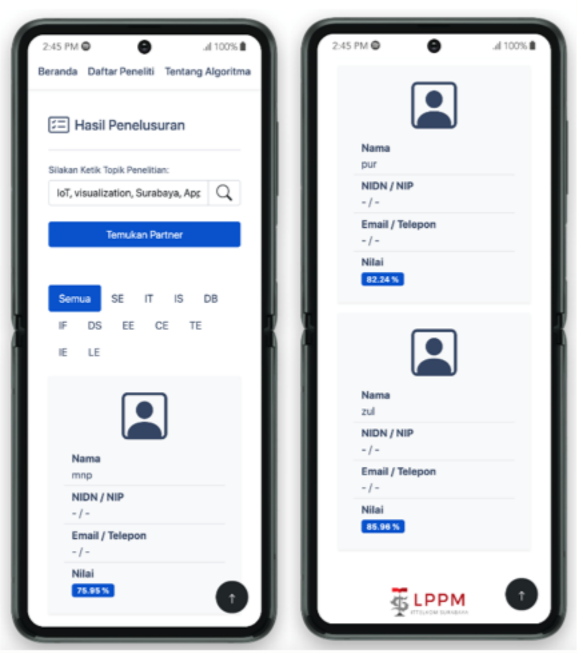
Nama	mnp
NIDN / NIP	- / -
Email / Telepon	- / -
Nilai	75.95 %

Nama	pur
NIDN / NIP	- / -
Email / Telepon	- / -
Nilai	82.24 %

Nama	zul
NIDN / NIP	- / -
Email / Telepon	- / -
Nilai	85.96 %

Fig. 7. Proposed system on web view.

It follows with the interface in mobile devices which holds ease of use and the five usability principles (Fig 8)¬. In a mobile perspective, users are able to find alternative candidates for research partners without piled interface. The proposed design focuses on the proposed system using a responsive web approach which provides a high resolution for a better understanding of recommendations. Finally, performance evaluation using cross-validation is tested to the system and results accuracy of 88,4%.



Silakan Ketik Topik Penelitian:
IoT, visualization, Surabaya, Apç

Nama
mnp
NIDN / NIP
- / -
Email / Telepon
- / -
Nilai
75.95 %

Nama
pur
NIDN / NIP
- / -
Email / Telepon
- / -
Nilai
82.24 %

Nama
zul
NIDN / NIP
- / -
Email / Telepon
- / -
Nilai
85.96 %

Fig. 8. Proposed system on mobile view.

The research partner recommendation system is a web-based application developed as a peer recommendation system by applying the Naive Bayes approach in calculating the score of each researcher from each research topic keyword given. The application is not only focused on being able to run the Naive Bayes algorithm but also provides the right UI / UX aspects so that the results of the algorithm can be translated into the right features and are easy to use (see Fig. 5). Therefore, this application is also built by going through the stages of SDLC such as planning, analysis, design, and implementation.

User flow, the form of features from the application of AI algorithms, and the development of database schemas are carried out in the planning and analysis stages. Then the UI/UX design is carried out at the design stage which applies usability principles such as minimalism design, error prevention, consistency, recognition rather than call, and so on. At the implementation stage, application design, and algorithms are translated into program code by interacting with data in the database. There are interesting findings in the development process, which is that at some point the database schema created becomes less efficient because it maintains a certain way to apply the AI algorithm used

Therefore, the experimental result does not satisfy minimum viable product (MVP) of the project due to inefficient data flow diagram. It can affect difficulty in updating the next version of the system development as we can see example of the data flow diagram in Fig. 6. In the depicted figure, as we might know, Naive Bayes algorithm has to label selected features in order to classify unknown data. However, in this example, the label will be formed as columns in the database which implies a number of columns occurred.

5. Conclusion

This paper presents a recommender system using Naïve Bayes algorithm and Sigmoid activation function based on a published article on Google Scholar. The proposed system helps to find the best research partner collaboration and also alternative researchers based on big data of researcher information in higher education. However, the output is in rough value because of statistical use of Naïve Bayes. Therefore, Sigmoid Activation Function transforms the display into a range of 0 to 1. Then the output can be shown in percentage mode. The implemented system shows that the proposed system has 88.4%. Of course, this approach has a limitation regarding data collection depends on collecting publication data from Google Scholar manually. In the future, we propose to leverage Google Scholar API to replace traditional data collection to be synchronized to Google Scholar.

Acknowledgement

This research project was supported by a grant from the research institute and community service (LPPM), Institut Teknologi Telkom Surabaya, Indonesia.

References

Abbas, A., Arrona-Palacios, A., Haruna, H., & Alvarez-Sosa, D. (2020). Elements of students' expectation towards teacher-student research collaboration in higher education. Proceedings - Frontiers in Education Conference, FIE, 2020-Octob. https://doi.org/10.1109/FIE44824.2020.9273902

Abramo, G., D'Angelo, C. A., & Solazzi, M. (2010). Assessing public-private research collaboration: Is it possible to compare university performance? Scientometrics, 84(1), 173–197. https://doi.org/10.1007/s11192-009-0104-0

Aldieri, L., Guida, G., Kotsemir, M., & Vinci, C. P. (2019). An investigation of impact of research collaboration on academic performance in Italy. In Quality and Quantity (Vol. 53, Issue 4). Springer Netherlands. https://doi.org/10.1007/s11135-019-00853-1

Amarante, V., Bucheli, M., & Vivas, R. (2021). Documentos de Trabajo Research networks and publications in Economics . Evidence from a small developing country. 11.

Cardona, T., Cudney, E. A., Hoerl, R., & Snyder, J. (2020). Data Mining and Machine Learning Retention Models in Higher Education. Journal of College Student Retention: Research, Theory and Practice. https://doi.org/10.1177/1521025120964920

Cotter, P., & Smyth, B. (2000). PTV: Intelligent Personalised TV Guides. Intelligent Applications of Artificial Intelligence.

De Campos, L. M. (2006). A scoring function for learning Bayesian networks based on mutual information and conditional independence tests. Journal of Machine Learning Research, 7, 2149–2187.

Di Sipio, C., Rubei, R., Di Ruscio, D., & Nguyen, P. T. (2020). A Multinomial Naïve Bayesian (MNB) Network to Automatically Recommend Topics for GitHub Repositories. ACM International Conference Proceeding Series, 71–80. https://doi.org/10.1145/3383219.3383227

Eger, S., Youssef, P., & Gurevych, I. (2018). Is it time to swish? Comparing Deep Learning Activation Functions across NLP tasks. Proceedings of the 2018 Conference on Empirical Methods in Natural Language Processing, EMNLP 2018, 4415–4424. https://doi.org/10.18653/v1/d18-1472

Fan J., Zhou W., Yang X, L. B. & X. Y. (2019). Impact of social support and presence on swift guanxi and trust in social commerce. Industrial Management & Data Systems, 119(9), 2033–2054.

Gaikwad, R. S., Udmale, S. S., & Sambhe, V. K. (2018). E-commerce Recommendation System Using Improved Probabilistic Model. Lecture Notes in Networks and Systems, 10, 277–284. https://doi.org/10.1007/978-981-10-3920-1_28

Ghani, R., & Fano, A. (n.d.). Building Recommender Systems using a Knowledge Base of Product Semantics. Recommendation and Personalization in ECommerce.

Guruge, D. B., Kadel, R., & Halder, S. J. (2021). The state of the art in methodologies of course recommender systems—a review of recent research. Data, 6(2), 1–30. https://doi.org/10.3390/data6020018

Ivaschenko, A., & Milutkin, M. (2019). HR decision-making support based on natural language processing. Conference on Creativity in Intelligent Technologies and Data Science, 152–161.

Khalid, S., Zohaib Irshad, M., & Mahmood, B. (2011). Job Satisfaction among Academic Staff: A Comparative Analysis between Public and Private Sector Universities of Punjab, Pakistan. International Journal of Business and Management, 7(1). https://doi.org/10.5539/ijbm.v7n1p126

Miyahara, K., & Pazzani, M. J. (2000). Collaborative filtering with the simple bayesian classifier. Pacific Rim International Conference on Artificial Intelligence, 679–689.

Neamah, A. A., & El-Ameer, A. S. (2018). Design and Evaluation of a Course Recommender System Using Content-Based Approach. 2018 International Conference on Advanced Science and Engineering (ICOASE), 1–6. https://doi.org/10.1109/ICOASE.2018.8548789

Nwankpa, C., Ijomah, W., Gachagan, A., & Marshall, S. (2018). Activation Functions: Comparison of trends in Practice and Research for Deep Learning. 1–20. http://arxiv.org/abs/1811.03378

Pavlov, D., & Pennock, D. (2002). A Maximum Entropy Approach to Collaborative Filtering in Dynamic, Sparse, High-Dimensional Domains. In S. Becker, S. Thrun, & K. Obermayer (Eds.), Advances in Neural Information Processing Systems (Vol. 15). MIT Press. https://proceedings.neurips.cc/paper/2002/file/cc7e2b878868cbae992d1fb743995d8f-Paper.pdf

Pazzani, M. J., Muramatsu, J., Billsus, D. S., & Webert. (1996). Identifying interesting web sites. Aaai, 54–59.

Ramachandran, P., Zoph, B., & Le, Q. V. (2017). Searching for activation functions. CoRR abs/1710.05941. ArXiv Preprint ArXiv:1710.05941.

Ramadhan, H. A., & Putri, D. A. (2018). Big Data, Kecerdasan Buatan, Blockchain, dan Teknologi Finansial di Indonesia. Direktorat Jenderal Aplikasi Informatika Kementerian Komunikasi Dan Informatika, 1–66. https://aptika.kominfo.go.id/wp-content/uploads/2018/12/Kajian-Kominfo-CIPG-compressed.pdf

Saleh, A. I., El Desouky, A. I., & Ali, S. H. (2015). Promoting the performance of vertical recommendation systems by applying new classification techniques. Knowledge-Based Systems, 75, 192–223. https://doi.org/10.1016/j.knosys.2014.12.002

Samin, H., & Azim, T. (2019). Knowledge Based Recommender System for Academia Using Machine Learning: A Case Study on Higher Education Landscape of Pakistan. IEEE Access, 7, 67081–67093. https://doi.org/10.1109/ACCESS.2019.2912012

Shah, L., Gaudani, H., & Balani, P. (2016). Survey on Recommendation System. International Journal of Computer Applications, 137(7), 43–49. https://doi.org/10.5120/ijca2016908821

Tripathi, D., Edla, D. R., Kuppili, V., & Bablani, A. (2020). Evolutionary Extreme Learning Machine with novel activation function for credit scoring. Engineering Applications of Artificial Intelligence, 96, 103980. https://doi.org/https://doi.org/10.1016/j.engappai.2020.103980

Volkwein, J. F., & Parmley, K. (2000). Comparing administrative satisfaction in public and private universities. Research in Higher Education, 41(1), 95–116. https://doi.org/10.1023/A:1007094429878

Wadekar, P. P., Pillai, Y. P., Roy, M. U., & Phadnis, P. N. (2018). Placement Predictor and Course Recommender System. Academia.Edu, 3960–3965. https://www.academia.edu/download/56794670/IRJET-V5I3929.pdf

Zadrozny, B., & Elkan, C. (n.d.). I ° : x, 694–699.

Zhang, Y., & Koren, J. (2007). SIGIR2007_Hierarchical_Bayesian_User_Modeling_in_RS.pdf. 47–54.

AUTHOR BIOGRAPHY

Mochamad Nizar Palefi Ma'ady is a Lecturer at Institut Teknologi Telkom Surabaya in Indonesia since 2022. Before then, he has 5 years of education experience at Nahdlatul Ulama Sunan Giri University, Indonesia. Nizar received his M.IM. degree in Information Management from National Taiwan University of Science and Technology, and a M.Sc. degree from the Department of Informatics at Institut Teknologi Sepuluh Nopember, Indonesia. He also holds a B.Comp. degree from Institut Teknologi Sepuluh Nopember at Surabaya. He is currently the Secretary of Information Systems Department and a reviewer of many international journals. His areas of interest include Neural Network Learning, Visual Analytics, Dynamic Programming, Markov Decision Processes, Social Computing, Machine Learning, and Text Mining..

DOI: 10.6977/IJoSI.202403_8(1).0004
S. S. Rath, P. Jagadeesh MP., J. Systematic Innovation, 8(1), 36-48 (2024)

Enhancing data security in SAP-enabled healthcare systems with cryptography and digital signatures using blockchain technology

Sonali Shwetapadma Rath[1]*, Prabhudev Jagadeesh M P[2]
[1*,2] Department of Computer Science & Engineering
JSS Academy of Technical Education
Bengaluru, Visvesveraya Technological University, Belagavi-590018
JSSATE-B Campus, Dr Vishnuvardhan Road,
Uttarahalli-Kengeri Main Road, Srinivaspura-Post, Bengaluru-560060
*Corresponding author email: 29sonali@gmail.com

(Received 20 September 2023; Final version received 18 January 2024; Accepted 19 January 2024)

Abstract

As the healthcare industry adopts more digital technologies, guaranteeing the security and privacy of sensitive patient data becomes increasingly important. Traditional centralized authentication solutions leave cyber threats and unauthorized access vulnerable. In response, the research demonstrates a novel strategy to improving data security and authentication in a SAP-enabled healthcare system by leveraging encryption and blockchain technologies. The research paper discusses the development and integration of a blockchain-based decentralized identity management system within the SAP platform. Each healthcare entity, including patients, doctors, and administrators, is given a distinct digital identity that is protected by using base 64 activity through DocuSign protects in SAP platform. The benefits of the proposed solution are assessed using a complete security analysis that measures data confidentiality, integrity, and availability. A comparison of DocuSign and SignEasy reveals DocuSign's superior performance in timestamp accuracy and document delivery speed. Its precision and reliability ensure document verification accuracy, while its streamlined workflow and advanced infrastructure expedite document processing, making it an ideal choice for businesses.

Keywords: SAP *platform*, Digital signature, Block chain technology, Security, Healthcare system, cyberattacks.

1. Introduction

The healthcare business has seen a significant digital revolution in recent years, embracing cloud-based solutions and electronic health records for effectively handling patient data. As more healthcare organizations embrace SAP (Systems, Applications, and Products) platforms for complete data integration and real-time analytics (Kessler et al., 2019) protecting the security and privacy of sensitive patient information has become a top priority. While technology improvements provide several benefits, they also present new concerns in protecting healthcare data from potential cyber threats and unauthorized access, given the data's ever-increasing value (Spanakis et al.,2020).

Traditional centralized authentication techniques have been criticized for being vulnerable to security breaches, which can jeopardize data integrity and accountability. Blockchain technology serves as the foundation for addressing these difficulties, acting as a distributed ledger with immutable features capable of securely storing authenticated actions and medical records. This technology improves data integrity by leveraging cryptographic hashing and decentralized validation, dramatically lowering the danger of unauthorized edits or data breaches. Additionally, the use of smart contracts streamlines established procedures, increasing efficiency and transparency within the healthcare ecosystem (Al et al.,2020).

The SAP Cloud Foundry platform, which is known for its scalability and efficient cloud-based services (Figueiredo 2022), is a suitable setting for integrating advanced security measures. It functions as an immutable and transparent ledger, recording each validated action and healthcare transaction utilising block chain's distributed ledger capabilities (Treiblmaier et al.,2020, Faccia et al., 2021). Due to the decentralized nature of the blockchain network, it eliminates single points of failure, protecting the healthcare system against hacker attacks and unauthorized data breaches. The use of cryptographic algorithms and smart contracts strengthens data integrity even further, prohibiting unauthorized access and malicious adjustments. Fig.1 illustrates the growing value of healthcare.

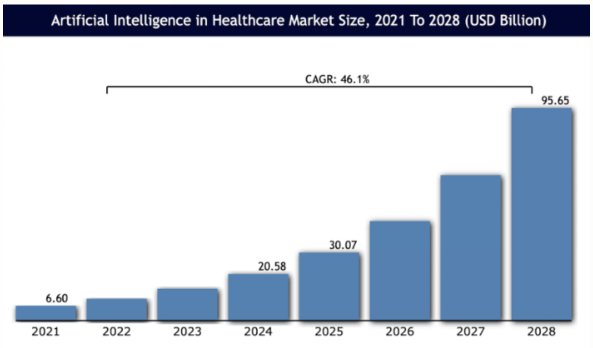


Fig.1. Growing value of healthcare data (https://resources.freeagentcrm.com/healthcare-industry-trends-and-statistics/)

The research highlights the importance of key management, legal compliance, and scalability in enhancing the security of healthcare systems. Key management is crucial in safeguarding patient privacy and maintaining health records. Legal compliance is essential in the healthcare sector, as it adheres to stringent data protection laws like HIPAA. Balancing security and compliance is a complex task, and healthcare institutions must be aware of their legal obligations while enhancing security measures. Scalability is crucial for healthcare systems dealing with vast amounts of data and serving a large number of patients. Scalability challenges can arise in hardware, software, network infrastructure, and security policy management across a larger system. To address these challenges, it is essential to invest in robust security solutions tailored to the healthcare sector's unique needs, such as advanced encryption techniques, secure access controls, and intrusion detection systems. Additionally, staying updated with evolving legal requirements and implementing compliance measures is vital. We also suggest integrating improved cryptographic authentication with the SAP Cloud Foundry backend, a cloud-based platform renowned for its scalability, flexibility, and ability to integrate seamlessly. This organization intends to create a dependable and expandable infrastructure for healthcare data access and storage while abiding by rules and regulations specific to the sector. In this research, we propose a novel method to strengthen healthcare data security by introducing encryption, employing Base64 activity through DocuSign for authentication, connecting it with blockchain technology, and doing so within the SAP Cloud Foundry backend. The primary contribution of this research can be summarised as follows:

One of the key contributions of this study is the development and integration of a blockchain-based decentralized identity management system for a healthcare system within the SAP platform. This decentralized method of identity management can improve security and privacy by minimizing reliance on a central authority for authentication.

The study recommends using encryption and blockchain technology to improve data security in healthcare systems by ensuring secrecy and transparency in data access and transactions.

The proposed blockchain-based solution's security, concentrating on data confidentiality, integrity, and availability, demonstrates its usefulness in assuring patient data protection.

According to the findings, blockchain-based solutions are more immune to cyberattacks and data breaches than traditional centralized authentication systems, which is critical in the healthcare industry.

2. Literature survey

A thorough analysis of the EHR security and privacy literature discovered 26 legislations, 23 symmetric key methods, 13 pseudo anonymity approaches, 11 digital signature systems, and Role-Based Access Control (RBAC) models proposed by (Fernández-Alemán et. Al., 2013). According to the study, more work is needed to implement these regulations and build secure EHR systems. However, in order to distribute health-related data via these approaches, more security is needed.

This study describes a patient-centric authorization protocol for Health Information Exchange (HIE) systems that addresses the shortcomings of previous techniques. The system assures authentication and outperforms existing techniques by employing a trapdoor hash-based proxy signature scheme by (Chandrasekhar et.al 2017). Adding network components can improve data distribution and broadcasting. However, extra security and privacy risks are built into the system design, which leaves some users unable to use these applications. Therefore, it is imperative to immediately suggest a mechanism for protecting the security of EHR data.

In order to use the cloud for electronic medical records, (Al Omar et al., 2019) have suggested certain national level frameworks. One of the challenges in fusing information from the Internet is protecting patient privacy.

In order to reconcile data processing capabilities with privacy concerns, a distributed dynamic authorization system based on blockchain is presented (Xu et.al 2022) for trustworthy data access. Patients' privacy information is not maintained on blockchain for greater data processing efficiency; instead, data access interfaces are provided via URLs in the authorization information.

This study describes a patient-centric authorization protocol for Health Information Exchange (HIE) systems that addresses the shortcomings of previous techniques. The system assures authentication and outperforms existing techniques by employing a trapdoor hash-based proxy signature scheme by (Roy et.al 2021).

The healthcare sector's resilience has been tested ever since the Covid-19 outbreak. Additionally, the sector was a prominent target of cyberattacks that worldwide disrupted important hospitals and health organizations. This explains why implementing suitable cyber security controls and making use of required technologies, such as the modern cloud, is essential.

(Kumar et. al.,2022) In the healthcare business, blockchain technology is critical for tackling data vulnerability and security. This study describes a secure blockchain mechanism for data management, with the goal of lowering overhead costs and speeding up ledger updates. The experimental results demonstrate a tenfold reduction in network traffic. However, storing a huge volume of data may result in inefficiencies and costlier issues in the proposed architecture.

Hospitals have personally identifiable information (PII) and personal health information (PHI), and when the PII or PHI data is stolen due to cyber-attacks, it puts patients' lives at risk and compromises the trust between doctors/providers and patient research by (He et. al.,2021) Cybersecurity challenges, risks, and plan to mitigate those risks are discussed. Risk assessment should be the first step in protecting sensitive PII and PHI data in the life sciences and healthcare industries. There are five phases of the NIST Cyber Security Framework, (NIST CSF). It begins with identifying or knowing what we need to defend (assets or data), followed by protecting, detecting, and then responding to any cyber-threats or incidents, and then concludes with recovering from them. We need to create policies, standards, and procedures for healthcare businesses with the aid of NIST CSF, and then deploy cutting-edge SAP S4Cloud products—which also leverage AI and their business technology platform (BTP) to enable Cyber domains—to protect their PHI/PII data. (SAP News 2021, SAP Help). The public, private, and hybrid cloud models offered by SAP S4/HANA allow you to select the model that will best serve your company needs and preserve your sensitive data. You can help develop your cyber security program based on the NIST CSF standard by implementing SAP S4/HANA.

3. Proposed methodology

Healthcare professionals today generate enormous amounts of medical-related data every day thanks to technology. The main repository for medical data in a hospital is called health records. Electronically generated clinical data is stored in health records. Data from health records is used for secondary purposes, such as medical trials, ongoing illness monitoring, and quality-improvement audits, in addition to the primary use of treating patients. When Health Records data are utilized for unrelated reasons without authorization—or, in certain cases, even with consent—privacy issues arise. The safety of individuals may be seriously jeopardized if sensitive personal information from health records is made available to or published to the public. Because of this and other problems with the present health records system, as demonstrated in Fig.2, data leaks and breaches constitute a major threat to any healthcare facility. Blockchain has the potential to make the entire facility secure depending on the specific permissions and conditions that the patient establishes. The data on a block chain is secured via cryptography. Each member of the network has a special private key that is associated with the transactions they execute and acts as a special digital signature. Any modifications to a record will be promptly detected by the peer network if the signature is changed, invalidating it.

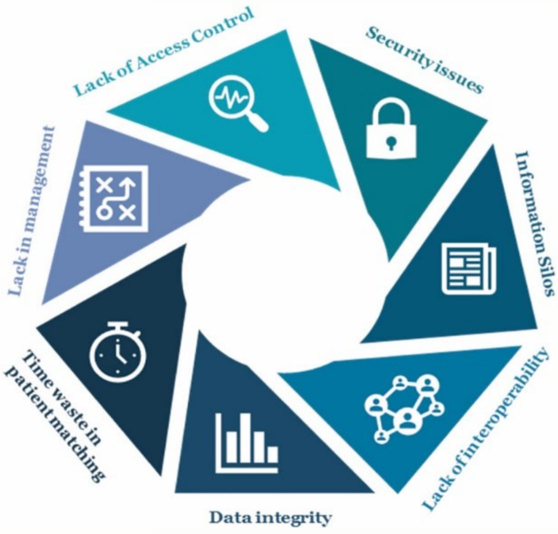


Fig.2. Issues in health record data storage

Additionally, recent research has discovered a searchable block chain that guarantee privacy preservation while enabling reliable search across encrypted distributed storage systems (Jiang et. Al.,2019). Another way for data security is to use smart contracts and verified computation (Avizheh et.al., 2019, Maddali et.al.,2020). The blockchain-based health records system may also include privacy-preserving homomorphic encryption methods [19,20]. However, as the volume of users and transactions grows, the block chain networks may experience scalability issues. As the network expands, the process of coming to an agreement and adding blocks to the chain may get slower, which could cause delays when handling a significant amount of health records. Additionally, keeping a lot of health records on the block chain may cause the chain's size to grow over time. This may increase the amount of storage needed and make it more difficult to maintain and replicate block chain data between nodes. Hence there is a need to develop a block chain-based decentralised identity management system within the SAP platform.

3.1. Proposed architecture

SAP Cloud Foundry provides a framework for the seamless integration of blockchain technology with existing healthcare systems and applications, allowing you to transition to a blockchain-based solution without disrupting the existing infrastructure. Also, the block chain incorporated with Base64 Activity via DocuSign to give an additional layer of data security in the healthcare industry. The term "Base64 Activity via DocuSign SAP" refers to a specific function or process within an organization that involves the use of Base64 encoding in conjunction with the DocuSign electronic signature platform and SAP (Systems, Applications, and Products in Data Processing), a widely used enterprise resource planning (ERP) software suite. This activity often involves the encoding and transfer of data in the Base64 format between SAP and DocuSign for a variety of applications such as document management, electronic signature workflows, and data interchange. Base64 encoding is a way of transforming binary data into a text-based format that is suited for secure transmission across different systems and platforms. It is frequently used when data, like as digital documents or photographs, must be securely transferred or processed between different applications or systems, as is frequently the case in an integrated environment such as SAP and DocuSign. Sensitive health information and user credentials are securely encoded and safeguarded by using Base64 Activity via DocuSign for cryptographic operations shown in Fig. 3. By using cryptographic techniques, this strategy maintains the confidentiality of the data and prevents unauthorized access by ensuring that only authorized users with valid digital signatures can access and interact with health records.

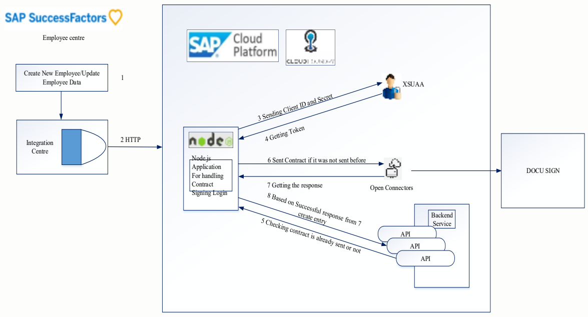

SAP SuccessFactors
Employee centre
Create New Employee/Update Employee Data
1
Integration Centre
2 HTTP
SAP Cloud Platform
CLOUD FOUNDRY
Node.js Application For handling Contract Signing Login
3 Sending Client ID and Secret
4 Getting Token
XSUAA
6 Sent Contract if it was not sent before
7 Getting the response
Open Connectors
DOCU SIGN
8 Based on Successful response from 7 create entry
5 Checking contract is already sent or not
Backend Service
API
API
API

Fig.3. Proposed SAP cloud platform architecture with DocuSign

Using block chain authentication along with cryptographic digital signatures, it is possible to securely and transparently authenticate the validity of user actions and transactions. Every action is verified using the block chain, which serves as an auditable and transparent authentication mechanism, increasing accountability and promoting trust in the healthcare system. The architecture of the block chain technology illustrated in Fig 4.

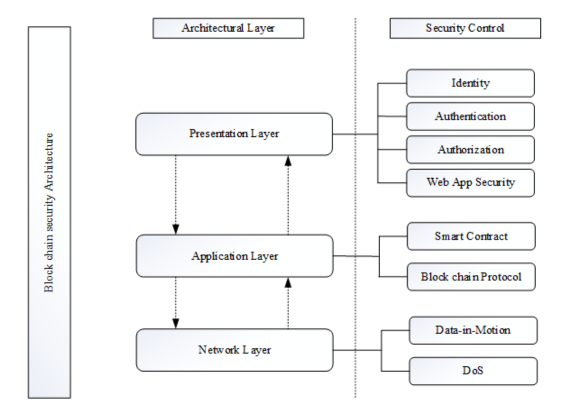


Fig.4. Block chain security architecture

The combination of blockchain technology with SAP Cloud Foundry improves the platform's ability to manage enormous volumes of health records while also ensuring the system's ability to adapt and scale as the healthcare ecosystem evolves. In conclusion, by combining Base64 Activity via DocuSign and Blockchain Technology in SAP Cloud Foundry Backend, the healthcare sector now has a strong, open, and safe method for safeguarding patient records and maintaining data integrity. This method tackles a variety of security and privacy issues, improves stakeholder confidence, and encourages effective and safe data sharing for better patient care.

Process flow

1. User enrolment and identity creation

- Within the SAP Cloud Foundry backend, users (such as patients, physicians, and administrators) register and build their digital identities.
- Each user receives a distinct public-private key pair from the backend. The user's device securely stores the user's private key, while the SAP backend stores the user's public key.

2. Base64 activity via Docusign

- The SAP Cloud Foundry backend creates a transaction hash encoding the action's contents when a user initiates an activity that needs authentication (such as viewing medical information or approving a prescription).
- The transaction hash is Base64 encoded by the backend, which transforms it into an ASCII representation.
- To serve as a cryptographic service provider, DocuSign receives the transaction hash in ASCII format.
- DocuSign uses its private key to perform extra cryptographic operations, like hashing and digital signing, to provide a special digital signature for the transaction hash.
- The digital signature is returned by DocuSign to the SAP Cloud Foundry backend.

3. Integration of Blockchain for Authentication

- The authorized activity is saved in a transaction block by the SAP Cloud Foundry backend along with its digital signature.
- The transaction block is sent by the backend to the blockchain network for consensus and confirmation.
- Using DocuSign's public key that is kept on the blockchain, the blockchain network verifies the legitimacy of the digital signature.
- By adding the transaction block to the blockchain after attaining consensus, all authenticated actions are recorded in an immutable and visible audit trail.

4. Using the Base64 activity through DocuSign, the user must create a fresh digital signature for each successive action.

5. The blockchain and user's public key that are both stored in the backend and SAP Cloud Foundry validate the digital signature.

6. After a successful verification, the user is permitted to carry out the authorized operation.

An immutable audit trail is produced on the blockchain by the recording of all verified actions. The blockchain can be accessed by healthcare managers and authorities to confirm the legitimacy of all actions and data flows. As a result, the blockchain technology combined with SAP Cloud Foundry's scalability and interoperability allow for smooth integration with current healthcare apps and systems.

3.2. Creation of oData service using CDS model

The simplest method to achieve this is by using SAP Cloud Foundry Backend Service, which retains data on all users who have already received contracts so that contracts won't be delivered again if changes are attempted to be triggered from the SFEC Integration Center. For storing user data and confirming whether or not a user has consented to a contract, it will build the oDATA service on-the-fly using the CDS architecture. Data models and annotations that specify the structure and behavior of the service you are creating must be defined in order to create an OData service using SAP's Core Data Services (CDS). oData is a standard protocol that makes it possible to build and use RESTful APIs for manipulating data.

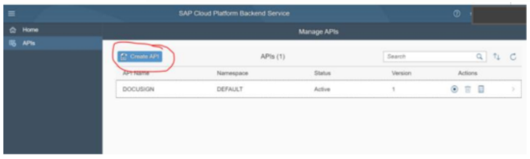
Basic summary of the SAP CDS oData service creation process:

1. Begin by utilizing CDS to define the data model. Declaratively describing data structures, connections, and behavior is possible with CDS. Defining entities, properties, and relationships is possible.
2. Add OData-specific annotations to the CDS entities. The entities' oData service exposure will be determined by these annotations.
3. Create an oData service definition that outlines the entities and associations that will be made available.
4. Activate your CDS artifacts after defining your CDS model and service definition. This produces the runtime objects, information, and artifacts for the oData service.
5. Use the oData endpoint to access the service once it has been produced.

In order to provide specific functionality through oData endpoints, DocuSign's APIs must be integrated using Cloud Platform Integration (CPI) in order to create an oData service. DocuSign is a platform for managing and signing documents electronically.

3.3. Creation of oData service using CPI

Using Cloud Platform Integration (CPI), a component of SAP Cloud Integration, you can build, configure, and deploy integration flows that expose data as oData endpoints. DocuSign is a platform for electronic signatures and document management that helps businesses to digitize and automate their contract-creation procedures. For drafting, distributing, signing, and managing electronic contracts and documents, the DocuSign service provides a number of capabilities.



API Name	Namespace	Status	Version	Actions
DOCUSIGN	DEFAULT	Active	1	

Fig.5. SAP cloud platform backend subscription with DocuSign API

Subscribing to and integrating the SAP Cloud Platform Backend service with the DocuSign API entails creating a connection between your SAP Cloud Platform environment and the DocuSign API to allow for easy data exchange as illustrated in Fig 5.

The process flow can be described below:

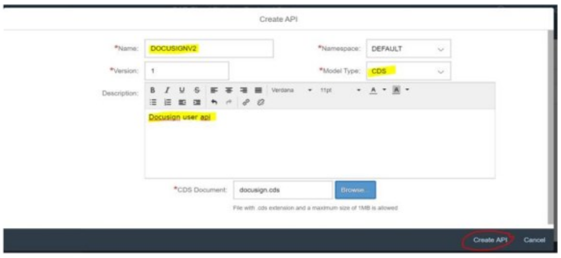
1. Create a new integration flow in SAP Cloud Platform Integration.
2. Configure the source endpoint to receive data from your application/system.
3. Configure the target endpoint to connect with the DocuSign API.
4. Obtain DocuSign API credentials (e.g., client ID, secret).
5. Configure the authentication mechanism in CPI to securely store and use these credentials.
6. Map the incoming data from your source system to the format expected by the DocuSign API request.
7. Use CPI's HTTP adapter to send a POST request to the appropriate DocuSign API endpoint.
8. Pass the mapped data as the request payload.
9. Receive the response from the DocuSign API.
10. Parse and process the response as needed.

Success factors Integration Centre and SAP Cloud Platform Integration provide suggestions on when to use each tool for interconnection creation. To create, test, and maintain incoming and outgoing integration, Success Factors features an integrated tool called Integration Centre. There are many other output file types available, including conversions from CSV to XML or JSON. The output can be safely stored on SFTP servers, and a number of scheduling options are available. Additionally, premade integration from the IntegrationCenter's catalog may be used and deployed on the customer instance.

Algorithm: DocuSign Service

Initialize users as an empty collection.
Repeat until the user chooses to exit:
Prompt the user for an action (create, retrieve, update, delete, or exit).
If the action is "create":
Prompt the user to enter the userID and sent status.
Call the CreateUser function with the provided input.
If the action is "retrieve":
Prompt the user to enter the userID to retrieve.
Call the GetUser function with the provided userID and display the user details.
If the action is "update":
Prompt the user to enter the userID and new sent status.
Call the UpdateSentStatus function with the provided input.
If the action is "delete":
Prompt the user to enter the userID to delete.
Call the DeleteUser function with the provided userID.
If the action is "exit":
Exit the program.
End Algorithm

Developing an integration with the DocuSign API (Fig.6) entails a number of stages, including registering your application, acquiring authentication credentials, sending API requests, and dealing with answers.



Create API

*Name: DOCUSIGNV2

*Namespace: DEFAULT

*Version: 1

*Model Type: CDS

Description: Docusign user api

*CDS Document: docusign.cds

File with .cds extension and a maximum size of 1MB is allowed

Fig.6. DocuSign API Creation

The credentials, which come with a Client ID and Secret Key, are sent after the API is created. The use of these credentials is for authentication.

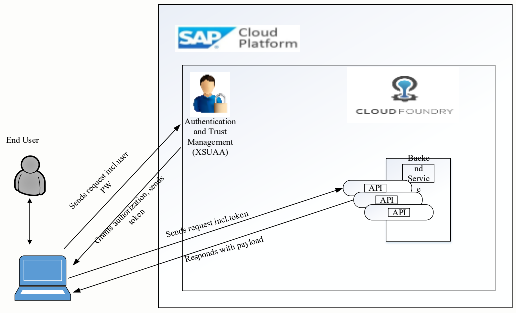


Fig.7. DocuSign API connection to Backend Service

The creation, deployment, and management of integration are all made possible by the SAP Cloud Platform Integration (CPI) solution for cloud middleware. Third-party programs and SAP OnPremise, SAP Cloud, or both may be connected in this manner. When linking SAP solutions to other SAP solutions and with other parties, it offers a broad range of connectivity options, including message translation, authentication, and even readymade integration options as shown in Fig 7. Successfactors is advised about solutions that range from modest to high level SAP Cloud Platform Integration. The Integration Center provides different scheduling options for the interfaces, such as once daily, once weekly, once monthly, and once yearly. Interfaces run more frequently, such as once every five minutes or once every day. CPI may be utilized in all of these situations. CPI also provides schedules with the "Run Once" option.

The overall process can be described below algorithm:

Import package like sap.gateway.ip.core.customdev.util.Message, ITApiFactory, securestore.SecureStoreService, securestore.UserCredential, usermodel.XSSFSheet, usermodel.XSSFWorkbook.
Define process data and SSFWorkbook (body)
//Extract Rows Data from Spreadsheet//
Set rowsData as an empty list.
Set headerCount to 0.
For each Row (row) in the sheet (mySheet):
Create a new empty list called rowData.
For each Cell (cell) in the row (row):
Append the value of the cell to the rowData list.
Append the rowData list to the rowsData list.
If headerCount is 0:
Set headerCount to the number of cells in the current row (row.getLastCellNum()).
End

//Process rows Data and Generate Output//
If rowsData is not empty:
Sort the rowsData list in ascending order (based on default comparison).
For each row (rowData) in rowsData:
If headerCount is 0 (no headers):
Increment headerCount by 1 to mark that the headers have been processed.
Else (headers have been processed):
Append the elements of rowData joined by commas to the output string.
Append a new line character ("\n") to the output string.
End
//Set Client ID and Client Secret, Define Services, and Handle User Credentials//
Set DocuSign_MS by calling ITApiFactory.getApi() to retrieve the DocuSign Microservice instance.
If DocuSign_MS is null:
// DocuSign_MS is not initialized, meaning user credentials are required.
Prompt the user to input their client ID and client secret.
Set clientSecret by replacing "&" with "%26" to handle special characters.
Set the message property "clientId" with the value of clientId.
Set the message property "clientSecret" with the value of clientSecret.
Return the message object.
Else
// DocuSign_MS is already initialized, no need to handle user credentials.
Proceed with defining the services and any other necessary operations using DocuSign_MS.
End
//Process Data and Modify Message Body//
Find the occurrence of "@odata.context" in the body string.
Replace "@odata.context" with "odata.context" in the body string.
Find the occurrence of "@microsoft.graph.downloadUrl" in the body string.
Replace "@microsoft.graph.downloadUrl" with "microsoft.graph.downloadUrl" in the body string.
Set the body of the message object to the modified body string.
Return the message object.
End
//Split Download URL and Map Properties with Complete URL//
Split the completeUrl using the "?" character as the delimiter.
Store the second part (index 1) of the split result into the query variable.
Extract the base URL part from completeUrl by taking the substring from index 0 to the index of the "?" character.
Store the base URL into the url variable.
Set the "url" property of the message object with the value of the url variable.
Set the "query" property of the message object with the value of the query variable.
Return the message object.
End
//Format JSON and Get User Details//
Remove the opening and closing brackets from the JSON string by taking a substring from index 1 to body.size()-1.
Store the modified JSON string back into the body variable.
Replace the substring "Comment_Rank" with "Comment Rank" in the body.
Replace the substring "Signature_Rank" with "SignatureRank" in the body.
Replace the substring "Business_Unit" with "Business Unit" in the body.
Set the modified body as the new body of the message object using message.setBody(body.toString()).
Return the message object.
End

As a result, DocuSign is a comprehensive electronic signing an agreement platform with a variety of features, such as advanced document routing, templates, support for mobile apps, identity verification, analysis, and more. It is well-known for its many integrations and workflow automation features. DocuSign has an advantage when it comes to establishing trust with highly regulated industries because it has been in the market longer. It offers a well-designed, intuitive interface with several customization possibilities. It is made to accommodate businesses in a variety of sectors and sizes.

4. Analysis of experimental results & discussion

4.1. DocuSign to daily data synchronization

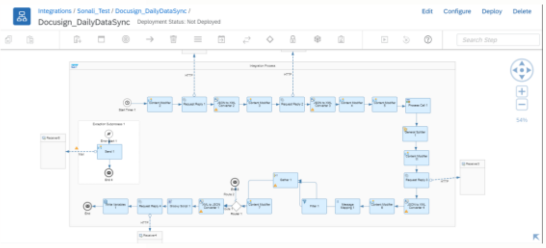

Integrations / Sonali_Test / Docusign_DailyDataSync /
Docusign_DailyDataSync
Deployment Status: Not Deployed
Edit
Configure
Deploy
Delete
Search Step
54%

Fig.8. DocuSign daily data synchronization

Data extraction is the first step in the DocuSign Daily Data Sync process as shown in Fig 8, and it starts with obtaining pertinent data from internal sources or systems. Documents, recipients, signature statuses, templates, and other relevant elements could all be included in this data. Data Upload & Synchronisation involves accessing the proper API endpoints to upload the modified data to the DocuSign platform. This could require making envelopes, managing recipients, updating document statuses, and other things, depending on the data being synchronized. The name "Docusign_DailyDataSync" indicates daily synchronization; hence, the process should be set up to execute every day at a certain time. Your DocuSign data is kept current thanks to this.

4.2. DocuSign to envelope creation

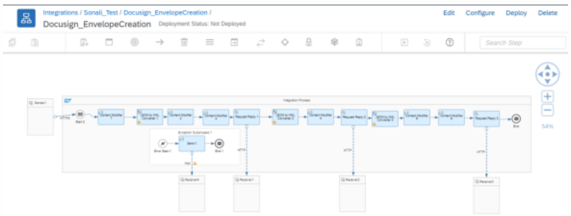

Integrations / Sonali_Test / Docusign_EnvelopeCreation /
Docusign_EnvelopeCreation
Deployment Status: Not Deployed
Edit
Configure
Deploy
Delete
Search Step

Fig.9. DocuSign to envelope creation

Fig 9 illustrates how the DocuSign Create Envelope constructs an envelope definition that outlines the desired creation of the envelope. The title of the email, email messages, receivers (signers, carbon copies, etc.), document location, tabs (signature, date, text fields), and any customized fields are all included in this. To connect the prepared documents to the envelope is to "add documents to envelope." Each document has a document ID assigned to it.

4.3. DocuSign to get envelope

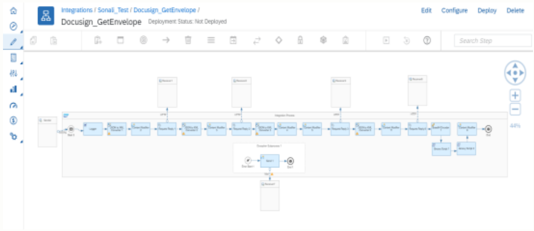

Integrations / Sonali_Test / Docusign_GetEnvelope /
Docusign_GetEnvelope
Deployment Status: Not Deployed
Edit
Configure
Deploy
Delete
Search Step

Fig.10. DocuSign to create envelope

4.4. DocuSign to event trigger

A specific envelope's status, recipients, documents, timestamps, and other pertinent data can be programmatically fetched using the "Docusign GetEnvelope" procedure, as is seen in Fig. 10. Tracking, reporting, auditing, and integration uses can all be made of this data. You can access the status and usage history of the envelope as well as the recipients' actions by utilizing the GetEnvelope API.

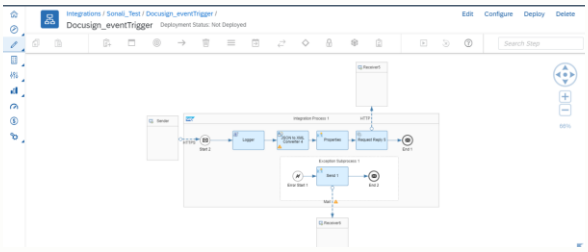

Integrations / Sonali_Test / Docusign_eventTrigger /
Docusign_eventTrigger
Deployment Status: Not Deployed
Edit
Configure
Deploy
Delete
Search Step

Fig.11. DocuSign to event trigger

The systems depicted in Fig 11 can be seamlessly and automatically integrated with the DocuSign platform through event triggers. They make sure the programme can react immediately to crucial occurrences inside the DocuSign workflow, optimising effectiveness and the user experience.

4.5. Performance Parameters

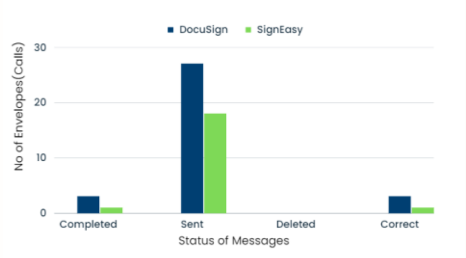


Fig.12. Envelope vs message

Fig 12 compares the number of envelopes with the status of messages, including completed, sent, deleted, and accurate messages. Strict security and compliance standards DocuSign may be chosen because of its reputation for strong security procedures, allowing it to manage a greater number of envelopes containing sensitive information. So it reveals that DocuSign has more envelopes than the SignEasy technique.

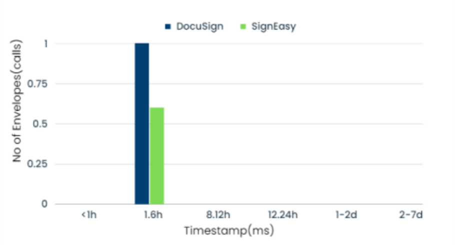


Fig.13. Envelope vs timestamp

Fig 13 compares the quantity of envelopes with the timestamp in terms of hours and days and DocuSign is well-known for its scalability, and it can manage a high amount of envelope transactions, making it ideal for enterprises that require a large number of signatures. As a result, DocuSign has more envelopes in the brief time stamp.

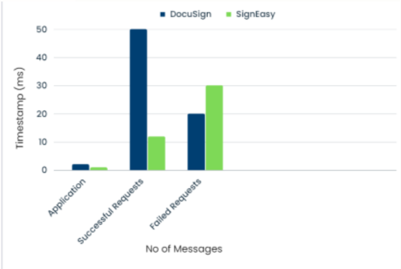


Fig.14. Timestamp vs messages

Fig 14 compares the number of messages processed in terms of application, successful requests, and unsuccessful requests with the timestamp in milliseconds. DocuSign receives more successful requests, including timestamping, implying that it has a strong and dependable infrastructure. This is especially important that rely on electronic signatures and document management to maintain constant efficiency and uptime. The figure indicates that DocuSign receives more successful requests than SignEasy. In comparison to SignEasy, DocuSign has less rejected requests.

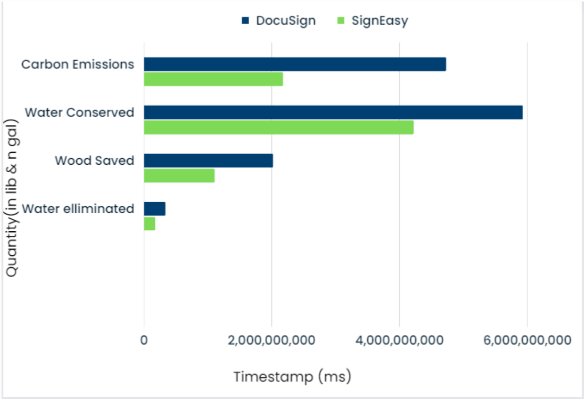


Fig. 15. DocuSign environmental impact

Fig 15 shows an analysis of the effects of DocuSign and SignEasy on the quantity assessed for carbon emissions, water conservation, wood preservation, and water elimination.

DocuSign outperforms SignEasy in terms of security, interaction with enterprise systems such as SAP, scalability, and global support. These benefits make it an appealing option for enterprises looking for a comprehensive electronic signature and document management solution, especially when dealing with high-security and high-volume document processing requirements.

4.6. Discussion

A user-friendly design and optimized e-signature functionalities are the main goals of this development. While providing the fundamental features of an electronic signature. DocuSign has an advantage in establishing trust with highly regulated businesses because of its longer history in the sector. Comparing other streamlined e-signature, the DocuSign has more regulations and trust. It offers comprehensive customer support and resources due to its larger user base.

5. Conclusion

This research paper investigates the topic of data security in SAP-enabled healthcare systems and presents a plan for addressing challenges such as data breaches, unauthorised access, and manipulation by utilising sophisticated technologies such as blockchain, digital signatures, and cryptography. With strong encryption techniques, cryptography protects confidential patient information and fortifies sensitive data. Digital signatures give an additional layer of assurance to document accuracy, lowering the chance of unauthorised changes. The proposed approach uses blockchain technology to create an immutable and decentralised ledger. According to the empirical review, security risks and data breaches have been significantly reduced. Performance benchmarks show that cryptographic procedures and digital signature verification are carried out efficiently within the SAP system. However, difficulties remain, needing careful planning and collaboration among healthcare institutions, technological professionals, and regulatory bodies. To keep healthcare data safe and secure in the ever-changing context of digital security, adaptability is critical. Our study improves healthcare data security by addressing immediate security problems while also fostering efficient, safe, and patient-centric data management, ensuring important patient and provider data remains secure and accessible.

However, the study identifies limitations and problems in applying a framework in healthcare settings, requiring coordination among institutions, technology professionals, and regulatory bodies, as well as adaptation to the ever-changing cybersecurity landscape. The revolutionary potential of this work has the potential to have a long-term impact on healthcare systems, ultimately benefiting both patients and healthcare professionals. In the future, this study investigates the feasibility of integrating public and private blockchain networks in healthcare systems to improve data access control, privacy, transparency, and scalability.

Reference

Kessler, S., Hoff, J. & Freytag, J. C. (2019). SAP HANA goes private: from privacy research to privacy aware enterprise analytics. Proceedings of the VLDB Endowment, 12(12), 1998-2009.

Spanakis, E. G., Bonomi, S., Sfakianakis, S., Santucci, G., Lenti, S., Sorella, M. & Magalini, S. (2020, July). Cyber-attacks and threats for healthcare–a multi-layer thread analysis. In 2020 42nd Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC) IEEE, 5705-5708.

Al Asad, N., Elahi, M. T., Al Hasan, A. & Yousuf, M. A. (2020, November). Permission-based blockchain with proof of authority for secured healthcare data sharing. In 2020 2nd International Conference on Advanced Information and Communication Technology (ICAICT), IEEE, 35-40.

Figueiredo, M. (2022). Developing Applications on SAP HANA Cloud. In SAP HANA Cloud in a Nutshell: Design, Develop, and Deploy Data Models using SAP HANA Cloud Berkeley, CA: Apress, 103-127.

Treiblmaier, H. & Sillaber, C. (2020). A case study of blockchain-induced digital transformation in the public sector. Blockchain and Distributed ledger technology use cases: Applications and lessons learned, 227-244.

Faccia, A. & Petratos, P. (2021). Blockchain, enterprise resource planning (ERP) and accounting information systems (AIS): Research on e-procurement and system integration. Applied Sciences, 11(15), 6792.

Fernández-Alemán, J. L., Señor, I. C., Lozoya, P. Á. O. & Toval, A. (2013). Security and privacy in electronic health records: A systematic literature review. Journal of biomedical informatics, 46(3), 541-562.

Chandrasekhar, S., Ibrahim, A. & Singhal, M. (2017). A novel access control protocol using proxy signatures for cloud-based health information exchange. Computers & security, 67, 73-88.

Al Omar, A., Bhuiyan, M. Z. A., Basu, A., Kiyomoto, S. & Rahman, M. S. (2019). Privacy-friendly platform for healthcare data in cloud based on block chain environment. Future generation computer systems, 95, 511-521.

Xu, B., Xu, L. D., Wang, Y. & Cai, H. (2022). A distributed dynamic authorisation method for Internet+ medical & healthcare data access based on consortium block chain. Enterprise Information Systems, 16(12), 1922757.

Roy, M. & Singh, M. (2021, April). Analytical Study of Block Chain Enabled Security Enhancement Methods for Healthcare Data. In IOP Conference Series: Materials Science and Engineering, IOP Publishing, 1131(1), 012002.

He, Y., Aliyu, A., Evans, M. & Luo, C. (2021). Health care cybersecurity challenges and solutions under the climate of COVID-19: Scoping review. Journal of medical Internet research, 23(4), e21747.

Kumar, A., Singh, A. K., Ahmad, I., Kumar Singh, P., Anushree, Verma, P. K. & Tag-Eldin, E. (2022). A novel decentralized blockchain architecture for the preservation of privacy and data security against cyberattacks in healthcare. Sensors, 22(15), 5921.

SAP News, Enterprise Threat Detection Cloud [Online]. Available: https://news.sap.com/2021/07/sap-enterprise-threat-detection-cloudbased-managed-service/

SAP Help, Enterprise Threat Detection Cloud Edition [Online]. Available: https://help.sap.com/docs/SAP_ENTERPRISE_THREAT_DETECTION_CLOUD_EDITION.

Jiang, S., Cao, J., McCann, J. A., Yang, Y., Liu, Y., Wang, X. & Deng, Y. (2019, July). Privacy-preserving and efficient multi-keyword search over encrypted data on blockchain. In 2019 IEEE international conference on Blockchain (Blockchain), IEEE, 405-410.

Avizheh, S., Nabi, M., Safavi-Naini, R. & Venkateswarlu K, M. (2019, November). Verifiable computation using smart contracts. In Proceedings of the 2019 ACM SIGSAC Conference on Cloud Computing Security Workshop, 17-28.

Maddali, L. P., Thakur, M. S. D., Vigneswaran, R., Rajan, M. A., Kanchanapalli, S. & Das, B. (2020, January). VeriBlock: A novel blockchain framework based on verifiable computing and trusted execution environment. In 2020 International Conference on COMmunication Systems & NETworkS (COMSNETS) IEEE, 1-6.

Jia, B., Zhang, X., Liu, J., Zhang, Y., Huang, K. & Liang, Y. (2021). Block chain-enabled federated learning data protection aggregation scheme with differential privacy and homomorphic encryption in IIoT. IEEE Transactions on Industrial Informatics, 18(6), 4049-4058.

Wood, A., Najarian, K., & Kahrobaei, D. (2020). Homomorphic encryption for machine learning in medicine and bioinformatics. ACM Computing Surveys (CSUR), 53(4), 1-35.

AUTHOR BIOGRAPHIES

Mrs Sonali Shwetapadma Rath (PhD) is a Research Scholar in the Department of Computer Science and Engineering at JSS Academy of Technical Education, Bengaluru, India. She has completed her M.Tech in Computer Science from Visveswaraya Technological University, Belagavi, and B.E in Information Technology from BPUT, Odisha. Her area of research interest includes Cyber & Information Security/Threats, Block chain, and SAP BTP Integration Suite. She has published 7 papers in International/National conferences and Journals. She has published 1 book chapter in Springer edited Book. She has acted as a resource person for various FDP/seminars/workshops on DevOps approach, Docker, Block chain, SAP Integration Suite,Dell Boomi . She has experience of mentoring critical projects sponsored by universities. She has been awarded as STAR faculty & MOUNTAIN MOVER. She has certified in Sales & CRM Overview from Salesforce Pathstream, certified in Robotic Process automation from Blueprism University, certified in Block chain Technologies, Smart Contract. She also received certificate as placement trainer under TCSion.

Dr. Prabhudev Jagadeesh is a Professor in the Department of Computer Science and Engineering at JSS Academy of Technical Education, Bengaluru, India. He completed his PhD in Computer Science from the University of Mysore, M.Tech in Software Engineering from Visveswaraya Technological University, Belagavi, and B.E in Computer Science & Engineering from the University of Mysore. His area of research interest includes Information Security, Machine learning, and Deep Learning. He has published 20 papers in International conferences and Journals. He has experience executing sponsored research projects in computer vision applications. He has worked on a consultancy project to explore a framework for a comprehensive sensor data processing engine. He has experience as General Chair and editor of proceedings of the 3rd International Conference on Cognitive Computing and Information Processing published by Springer.

Performance evaluation of deep learning models for detecting deep fakes

Aishwarya Rajeev[1,2]*, Raviraj P[3]

[1]Research Scholar, Geetha Shishu Shikshana Sangha Institute of Engineering & Technology for Women, Mysuru, Karnataka, Affiliated to VTU Belagavi, 570016.

[2]Department of CSE, Coorg Institute of Technology, Ponnampet, Karnataka, Affiliated to VTU Belagavi, 571216.

[3]Professor, Geetha Shishu Shikshana Sangha Institute of Engineering & Technology for Women, Mysuru, Karnataka, Affiliated to VTU Belagavi, 570016.

*Corresponding author mail: aishwaryarajeev@gmail.com

(Received 23 October 2023; Final version received 04 December 2023; Accepted 27 December 2023)

Abstract

The proliferation of deep fake content in multimedia has necessitated the development of robust detection mechanisms. In this study, a comparative analysis of four state-of-the-art deep learning models for detecting deep fakes is conducted: CNN+RNN, DAFDN, Hybrid Inception ResNet v2, and Xception. The evaluation focuses on their performance metrics, emphasizing accuracy as a primary measure. Through extensive experimentation and evaluation on a comprehensive dataset, the findings reveal notable distinctions among these models. The CNN+RNN architecture demonstrates a commendable accuracy of 94.8%, providing a solid baseline for comparison. Surpassing this, the DAFDN model achieves an accuracy of 97.8%, showcasing superior discriminatory capabilities in identifying manipulated content. Furthermore, the CNN model stands out with an accuracy of 98%, exhibiting remarkable effectiveness in distinguishing between genuine and deep fake media. The comparative analysis delves into the strengths and weaknesses of each model, shedding light on their respective performance levels in detecting sophisticated deep fake content. The observed accuracies underscore the nuanced differences in their architectures and training methodologies, offering insights crucial for selecting appropriate models based on specific detection requirements.

Keywords: Face Forensics, Convolutional neural network, recurrent neural network, DAFDN, Resnet v2, Xceptio

1. Introduction

Deep fake films are modified videos that use machine learning-based algorithms to swap out humans for other objects or actors in an existing image or video. Three categories of deep fake videos exist: lip-syncing, face swapping, and head puppetry. The art of head puppetry involves using a source video person's head to manipulate a video of a specific human's head and upper shoulder so that the altered person looks exactly like the source (Shad et.al.,2021). Face swapping is changing a person's face to that of another while keeping the same expression on their face. Since lip-syncing merely modifies the lip area of a video, the target person says something that isn't actually true. Although some deep fakes can be produced using classic visualization techniques or computer graphics, the most recent and widely used deep learning techniques for producing deep fake videos are auto encoders and generative adversarial networks (GAN) (Rahman et.al.,2022).

These models are used to synthesize face images of people with comparable expressions and movements based on the analysis of a person's facial emotions and movement. For deep fake technologies to train a model to create photorealistic photos and movies, a significant amount of image and video data sets are typically required. Politicians and celebrities are the first targets of profound fakes due to the sheer volume of their films and photographs that are readily available online (Nguyen et.al.,2019) In pornographic pictures and films, the heads of famous people and political figures have been replaced with deep fakes. In the first deep fake movie, a celebrity's visage was changed to that of a porn, it was released in 2017. Deep fake movies are an increasing concern to global security since they are increasingly being used to produce false speeches by world leaders (Bode 2021).

In response to this challenge, researchers have explored various deep learning architectures to enhance detection accuracy. Comparative analyses have been conducted to assess the efficacy of different models, including CNN+RNN, DAFDN (Deep Adaptive Feature Distillation Network), Hybrid Inception ResNet v2, and Xception. Each architecture brings its unique strengths in identifying subtle discrepancies and patterns within manipulated content, striving to outperform adversaries' deep fake generation techniques Verdoliva 2020). The CNN+RNN model combines Convolutional Neural Networks (CNN) for feature extraction with Recurrent Neural Networks (RNN) for temporal information processing, offering a comprehensive approach to capture both spatial and sequential patterns in videos, a common format for deep fakes (Kousik et.al.,2021). DAFDN utilizes adaptive feature distillation to distill informative features and mitigate the domain gap between real and fake videos, enhancing detection accuracy.

Meanwhile, the Hybrid Inception ResNet v2 and Xception architectures leverage the power of inception modules and efficient convolutional operations, respectively, to improve feature extraction and model robustness against increasingly sophisticated manipulations (Kamaleldin et.al.,2023). However, amidst these advancements, a notable research gap persists. Despite significant progress in deep fake detection, the adaptability of detection models to new and evolving manipulation techniques remains a challenge (Guo et.al.,2021). The rapid evolution of deep fake generation methods continuously outpaces the development of detection algorithms, leading to a need for models that can generalize across diverse types of manipulations and adapt swiftly to emerging fake media tactics (George et.al.,2023, Wang et.al.,2023).

In conclusion, while various deep learning architectures have shown promise in detecting deep fakes, the dynamic landscape of fake media creation demands continuous innovation and adaptation in detection models to effectively address the ever-evolving challenges posed by deep fakes. Closing the research gap by creating more adaptable and robust detection mechanisms stands as a critical next step in the ongoing battle against misinformation and deceptive media.

The paper is organized in a systematic manner. It begins with a thorough introduction and then quickly reviews the body of research that has already been done on face forensics. The paper's main body goes into a thorough comparison study of several false face forensics designs. The salient features of the comparison are then highlighted by a summary. The article culminates with a thorough synopsis that synthesizes the knowledge and understanding acquired during the investigation.

2. Literature survey

A literature survey on deep fake detection systems encompasses a comprehensive exploration of existing research, methodologies, and advancements in the field. This survey delves into the diverse array of approaches employed to detect and mitigate the proliferation of manipulated multimedia content, specifically focusing on deep fake videos. It encompasses an analysis of various techniques, such as machine learning algorithms, neural networks, forensic analysis, and other innovative methodologies utilized to identify and combat the rising sophistication of deep fake technology. The survey aims to synthesize the current state-of-the-art methodologies, highlight their strengths and limitations, and identify potential avenues for further research and enhancement in the realm of deep fake detection systems.

(Feng Ding et al 2020) The author created a creative framework as a digital forensics tool to protect end users. It is built on deep learning and uses categorization to find assaults. The suggested model's data collecting effectiveness, resilience, and detection performance are all improved when compared to the traditional approaches, which are supported by our experiments. Additionally, our suggested approach makes use of 5G HetNets to allow high-quality real-time forensics services on edge consumer devices (ECE), such as smartphones and laptops, which has significant practical implications. Additionally, certain conversations are held to describe potential risks in the future.

(Nickson M. Karie et al 2019) The DLCF Framework, developed as a result of this research, provides a general framework for converting DL cognitive computing techniques into Cyber Forensics (CF). By imitating human decision-making in neural networks, DL uses a variety of machine learning techniques to address problems. These considerations suggest that DL has the potential to both offer forensic investigators options while also having the potential to drastically impact the field of CF in a number of ways. Examples of such remedies include minimizing prejudice in forensic investigations, contesting the admissibility of certain types of evidence in court proceedings or other civil hearings, and many others.

(Akash Chintha et al 2020) this paper offered straightforward yet effective digital forensic techniques for spoof audio and deep fake image detection. The methods combine bidirectional recurrent structures, entropy-based cost functions, and convolutional latent representations to extract semantically rich information from recordings. They are shown using the Face Forensics++, Celeb-DF, and ASVSpoof 2019 Logical Access video datasets and audio datasets, setting new standards in every category. To show generalization to other domains and learn more about the efficacy of the new designs, extensive investigations are carried out.

(Bin Wu et al 2023) to extract relevant and unusual phrases from local areas, the author presented a brand-new framework called FPCNet. For the purpose of identifying face forgery films, this system employs CNN, LSTM, CGLoss, and adaptive feature fusion. In experiments, the suggested technique's detection speed reaches 420 FPS, and the auc scores on the Raw CelebDF, Raw Face Forensics++, F2F, and NT datasets achieve the best results of 99.7%, 99.9%, 94.7%, and 82.0%, respectively. The experimental findings show that the suggested framework outperforms existing frame-level approaches in terms of time economy while also boosting detection performance.

(Ahmed Sedik et al 2022) in this study, a cyber-facial spoofing assault was combined with a deep learning methodology for video face forensic recognition utilizing convolutional neural networks (CNN) and convolutional long short-term memories (ConvLSTM). Simulation findings showed that the ConvLSTM with CNN methodology gave improved classification results in comparison to other conventional strategies. with an accuracy of 99% and up to 95%. In each technique, the classification function was handled by the SoftMax layer.

(Jiahui Wu et al 2023) remote photoplethysmography (rPPG) technique collects heartbeat signals from video recordings by analyzing the small variations in skin color induced by cardiac activity. This is a strong biological signal for deep fake detection since it develops distinct rhythmic patterns in response to various manipulation approaches. To capture both spatial and temporal differences, a two-stage network made up of a Temporal Transformer and a Mask-Guided Local Attention module (MLA) is proposed. The effectiveness of our method in comparison to all existing rPPG-based methods has been thoroughly tested on the Face Forensics ++ and Celeb-DF datasets. The proposed method's usefulness is also demonstrated through visualization.

(Davide Coccomini et al 2022) Since most algorithms are becoming more adept at creating realistic human faces, the author focused on video deep fake detection on faces in this work. We particularly combine different types of Vision Transformers with an Efficient Net B0 convolutional network used as a feature extractor, and the results are comparable to some more recent methods that also use Vision Transformers. Unlike current methodologies, the author does not employ distillation or collective approaches. Additionally, we offer a fundamental inference procedure based on a straightforward voting system for addressing several faces in a single video clip. The top model scored an F1 score of 88.0% and an AUC of 0.951 on the Deep Fake Detection Challenge (DFDC), which is very close to the state-of-the-art.

Aishwarya Rajeev et al [18] Numerous techniques, including Random Forest, Multilayer Perceptron (MLP), and Convolutional Recurrent Neural Networks (CRNN), are employed in this study to execute various kinds of forensic investigation. Also employed is image fusion, which may combine many photos to create a single image with more information and extract characteristics from the original images. According to this study's findings, the random forest has a 98.02 percent accuracy rate when it comes to producing the best results for network forensic investigation. The paper seeks to present an extensive summary of the work that has been done over the past few years to analyze current techniques and techniques for video source authentication using machine learning.

3. Face Forensics

Face forensics, commonly referred to as facial recognition forensics, is the use of forensic techniques and technology to analyze and study face photographs or videos with the aim of establishing identity, verifying a person's identity, or obtaining evidence. To extract and analyze face characteristics and patterns in order to infer or draw conclusions entails using a variety of techniques, algorithms, and tools.

Face forensics may be used in a variety of fields, including biometrics, digital forensics, law enforcement, and security. The following are some typical uses for facial forensics:

- ***Facial Identification:*** Face forensics is frequently used to identify people in surveillance footage, pictures, or videos. To uncover probable matches, face recognition algorithms compare the subject's facial traits with a database of well-known people (Aishwarya et.al.,2023).
- ***Face Authentication:*** It includes matching a person's face traits with their stored biometric information to confirm their identification. Mobile devices, access control systems, and other security applications all make use of this technology (Xiao et.al.,2019).
- ***Facial Image Analysis:*** To extract information from photos or videos or to spot modifications, forensic professionals employ face image analysis tools. This might involve analyzing facial expressions, locating locations or characteristics, and determining the veracity or integrity of a picture, among other things (Ahmadi et.al.,2021).
- ***Facial Age Progression/Regression:*** Face forensics methods can be applied to a person's face to imitate the aging or de-aging of their face based on their present or former look. Investigating missing persons or identifying people in unresolved instances may benefit from this (Ross et.al.,2020).
- ***Facial Emotion Analysis:*** In order to identify emotional states like happiness, sorrow, rage, or surprise, face recognition algorithms analyze facial expressions. Fields like psychology, market research, or human-computer interface may find a use for this (Chandaliya et.al.,2022).
- ***Facial Image Retrieval:*** Face forensics may help in searching through massive databases of pictures or videos based on particular features or traits of the face. Criminal investigations or the identification of people of interest may benefit from this (Ivanova et.al.,2020).

It's crucial to recognize that the application of face forensics involves issues of privacy and ethics. Discussions about regulation and protecting personal privacy have arisen in response to the potential for misuse or abuse of face recognition technology.

Face forensics, or face manipulation detection, is an important area of research in computer vision and deep learning. Various methods have been developed to detect manipulated or fake faces using deep learning concepts. Here are a few different techniques for face forensics:

3.1. Facial Expression Analysis

Facial expression analysis is an intriguing area that focuses on comprehending and analyzing the emotions expressed via facial expressions. Researchers and practitioners can explore the intricate world of human emotions by analyzing the minute movements, configurations, and dynamics of the face (Sikkandar et.al.,2020). In this procedure, essential face traits including the position of the eyebrows, the state of the eyes, the shape of the lips, and more are extracted from facial photos or videos. Then, these characteristics are examined using a range of techniques, including machine learning and computer vision algorithms, to categorize and interpret emotions including happiness, grief, rage, surprise, fear, and disgust. (Keshari et.al.,2019). Applications for facial expression analysis may be found in a variety of industries, including psychology, HCI, market research, and even the professional diagnosis of mental health issues. Facial expression analysis helps us better comprehend non-verbal communication and human emotions by utilizing ongoing technological and algorithmic breakthroughs (Hussain et.al.,2020).

3.2. Facial Landmark Detection

To locate and analyze important spots or landmarks on the face, a major approach used in face forensics is known as facial landmark detection. These landmarks, including the corners of the mouth, nose, and eyes, offer geometric information and serve as starting points for further study. Researchers may learn important details about the structure, position, and emotions of the face by precisely detecting and tracking facial landmarks. These insights are crucial for spotting and analyzing possible modifications (Ashwin et.al.,2019). Facial landmark detection is essential in face forensics for determining the veracity and accuracy of a face picture. Disparities or inconsistencies brought on by manipulations, such as face swapping or morphing, can be found by analyzing the locations, configurations, and motions of landmarks. When landmark spatial connections differ from what is expected, it may be a sign that the picture has been altered or tampered with. Additionally, facial landmark identification can help pinpoint regions of interest for later research. For instance, by identifying the eye landmarks, researchers may concentrate on analyzing eye-related alterations, such as changing the color of the eyes or adding digital contact lenses. Similar to this, recognizing mouth landmarks can assist in spotting possible lip-syncing or speech manipulation. It's crucial to remember that face forensics facial landmark detection might be difficult. It could be delicate to changes in facial expression, occlusions, or head posture (Agbolade et.al.,2019). Accurate landmark detection may also be hampered by the presence of cosmetics, accessories, or facial hair. As a result, to manage these complexity levels and guarantee correct outcomes, strong and precise algorithms are required (Bozkir et.al.,2023). In conclusion, facial landmark detection is an important method in face forensics that provides geometric data and helps spot any alterations. Researchers may improve their study of facial integrity by utilizing precise landmark detection, making a contribution to the fields of digital forensics and biometrics as well as assuring the reliability of face-based authentication systems.

3.3. Face Swapping Detection

Face forensics experts use the important technology of "face swapping detection" to spot instances of faces being switched or replaced in photos or videos. This method seeks to spot visual tricks when one person's face is digitally swapped out for another, which frequently produces believable but misleading results. Face swapping detection is essential for maintaining the authenticity and integrity of visual material in face forensics. Face swapping detection algorithms can spot obvious evidence of manipulation by analyzing a variety of visual signals and attributes, including facial landmarks, textures, lighting, and consistency in facial emotions (Dargan et.al.,2020). The geometric alignment and arrangement of facial landmarks before and after the swap is one typical method utilized in face swapping detection. Key facial features like the eyes, nose, and mouth may be precisely detected and tracked, making it possible to see any differences or irregularities in their locations. Face swapping may be present if there are significant differences in the way these landmarks are spaced out from one another. The uniformity and naturalness of the face textures in the swapped region may also be examined using texture analysis tools (Verdoliva 2020). Anomalies can point to the presence of face alteration, such as artificial blending or variances in lighting.

Due to improvements in face manipulation methods and the possibility for perfect blending, face swapping detection can be difficult. Face swapping algorithms based on deep learning can provide extremely convincing results that are hard to spot. In order to improve face swapping detection techniques' accuracy and sturdiness, it is essential to conduct continuing research and make improvements in the fields of deep learning, computer vision, and forensic analysis (Hashmi et.al.,2020). In order to detect instances of face replacement or swapping, face forensics must employ a fundamental method called face swapping detection. It contributes to the creation of trustworthy digital media and raises the trustworthiness of face-based authentication systems by assessing facial landmarks, textures, and other visual clues to assure the integrity and authenticity of visual material.

3.4. Deep Fake Detection

A key method in face forensics for identifying and detecting heavily altered or artificial face photos and films is deep fake detection. The term "deep fake" refers to media that has been purposefully made using deep learning algorithms. Often, this involves swapping out a person's face for another, creating very lifelike and deceptive visual results. Deep fake identification in face forensics is essential for reducing the dangers that might result from the illicit exploitation of altered material. To analyze and examine the veracity of face material, deep fake detection algorithms use a variety of methodologies. Examining visual artifacts, inconsistencies, and abnormalities that are often included throughout the deep fake creation process is involved in these procedures (Deshmukh et.al.,2020). Analyzing minute artifacts and flaws that result from the synthesis process is a typical strategy in deep fake detection. Deep fakes frequently display uneven mixing, irregular lighting, and differences in the resolution and texture of the face. Algorithms can recognize these red flags and differentiate deep fakes apart from real face material by utilizing deep learning models and computer vision techniques (Awotunde et.al.,2022).

Using sophisticated machine learning models to discover and identify patterns particular to deep fakes is an alternative strategy. These models can spot statistical discrepancies and distinctive traits related to deep fake manipulation by training on a sizable dataset of both genuine and deep fake samples. With this method, key characteristics are frequently extracted and the legitimacy of the information is categorized using convolutional neural networks (CNNs) or recurrent neural networks (RNNs). However, due to the quick advancement of deep fake-generating techniques as well as the introduction of advanced adversarial strategies, deep fake detection approaches confront difficulties. Deep fakes produced by adversarial networks may be extremely convincing and difficult to differentiate from legitimate information, making detection more difficult. In order to keep ahead of developing methods and guarantee the integrity of digital material, ongoing research, and breakthroughs in deep fake detection are essential (Byrnes et.al.,2021).

In order to recognize fabricated or altered face photos and videos, deep fake detection is an essential approach in face forensics. Deep fake detection approaches help protect against the possible abuse of deep fakes and improve the credibility of digital material by analyzing visual artifacts, and inconsistencies, and utilizing machine learning models.

3.5. Deep Texture Analysis

Deep texture analysis is a crucial tool in face forensics. It focuses on analyzing the complex patterns, specifics, and irregularities contained in the texture of a face using deep learning models and cutting-edge computer vision techniques. It seeks to identify minute texture variations that may be a sign of face switching, digital retouching, or other types of manipulation. Convolutional neural networks (CNNs) trained to identify and extract pertinent characteristics from facial textures are one popular method in deep texture analysis. Algorithms can spot differences or anomalies that may indicate tampering by analyzing the consistency and coherence of texture patterns across various face areas (Chen et.al.,2022).

However, when dealing with alterations in lighting, picture quality, or the presence of cosmetics and accessories that might affect the texture look, deep texture analysis may run into difficulties. The identification method is made more difficult by the fact that very advanced modification techniques may produce deep fakes with convincing texture features. To keep up with changing manipulation techniques and boost the precision and resilience of deep texture analysis approaches, ongoing research and development are crucial (Xi et.al.,2020).

In conclusion, deep texture analysis is an important method in face forensics for identifying and analyzing textural elements in facial photographs. It adds to the identification and detection of manipulated or altered faces by utilizing deep learning models and studying texture patterns and irregularities, improving the integrity and reliability of face-based forensic investigation and digital media authentication.

Table 1. Comparison table of different face forensics techniques using deep learning concepts

Ref.	Technique	Description	Advantages	Limitations
Jeong et.al.,2020	**Facial Expression Analysis**	Analyses facial expressions to infer emotions or intentions	Provides insights into the emotional state of a person	Difficulty in accurately interpreting complex or subtle facial expressions
Zhu et.al.,2019	**Facial Landmark Detection**	Identifies key facial landmarks for further analysis	Provides geometric information about the face	Sensitive to occlusions or variations in facial expressions
Zhang et.al.,2019	**Face Swapping Detection**	Identifies instances where faces have been swapped or replaced	Effective in detecting face replacement or swapping	Limited to face swapping techniques
Zhao et.al.,2021	**Deep Fake Detection**	Detects manipulated faces using deep learning models	Effective against deep fake videos	Limited to specific types of manipulations (e.g., deep fake videos)
Bonomi et.al.,2021	**Deep Texture Analysis**	Analyses textural features within the face using	Effective in detecting inconsistencies or	May struggle with subtle manipulations or

		deep learning models	manipulations in texture patterns	variations in image quality

4. Analysis On Face Forensics

Face forensics analysis is a broad discipline that includes a range of methods and tools for analyzing and interpreting facial characteristics in the settings of criminal investigations and legal evidence. It entails using face recognition algorithms to match and identify people according to their facial features. Experts in face forensics also use picture authentication techniques to establish the veracity and integrity of facial images, such as examining any indications of alteration or digital interference. In order to understand emotions, believability, and deceit, they also analyse the facial expressions shown in photos or movies. Techniques for estimating a person's age and predicting their anticipated appearance at various ages are used, as well as methods for comparing facial traits to databases of missing people or suspects. In order to confirm identities or determine whether persons in visual evidence are the same, facial comparison and superimposition techniques are used. In addition, forensic anthropology employs facial reconstruction methods to recreate the look of unidentified human remains' faces, aiding in the identification process. But it's important to recognize the possible drawbacks of face forensics, such as changes in picture quality, lighting, position, and the existence of barriers or disguises, which may affect the precision of studies. Therefore, face forensics analysis should be performed by qualified experts who take these elements into account and use caution when making interpretations.

In this study, we concentrate on the face forensics research conducted by several researchers, compare those works, evaluate the findings, and, after comparing all the works, focus on the shortcomings of prior research and utilize those shortcomings as the foundation for our future work.

4.1. Deep Fake Video Detection Using CNN and RCNN :

(Ashifur Rahman et al 2022) deep fake videos that are convincing and increasing quickly in popularity can now fool even experienced professionals. The political, social, and personal spheres are all affected significantly by these profoundly false videos. In high quality and lengthy video data, modern machine learning experiments provide demonstrable success in spotting fraudulent movies, however, this performance is not demonstrated in low resolution and brief video clips. In this study, the authors created a model using convolutional neural networks (CNN) and recurrent neural networks (RNN) that shows mentionable accuracy in detecting fraudulent films in low-resolution and short-duration video data. In our experiment, the author employed the Kaggle Deep Fake Detection Challenge (DFDC) dataset and the Face Forensics++ dataset. When it came to identifying false videos, the model performed 94.8% accurately for the RCNN model and 94.2% accurately for the CNN model. The author compared the performance of our models to cutting-edge techniques and evaluated our models using several performance measures. Comparable performance is shown by the model. The mathematical equation for RNN has been shown in equations 1 and 2

Recurrent Neural Networks (RNNs) are a class of neural networks specifically designed to work with sequence data. They are defined by the recurrence of neural network modules, allowing them to maintain a memory of previous inputs to make decisions about the current input. Mathematically, an RNN can be represented through its forward pass equations.

Let's denote:

- x_t as the input at time step t
- h_t as the hidden state at time step t
- y_t as the output at time step t

The basic equations for a simple RNN are as follows:

$$h_t = \sigma(W_{ih} \cdot x_t + W_{hh} \cdot h_{t-1} + b_h) \quad (1)$$

$$y_t = \text{softmax}\left(W_{hy} \cdot h_t + b_y\right) \quad (2)$$

Where:

- W_{ih} is the weight matrix for input to hidden connections.
- W_{hh} is the weight matrix for hidden to hidden connections.
- W_{hy} is the weight matrix for hidden to output connections.
- b_h and b_y are the bias terms for the hidden and output layers, respectively.
- σ is an activation function, commonly the hyperbolic tangent (tanh) or the Rectified Linear Unit (ReLU).

4.1.1. Data Set:

This study utilized BlazeFace, MTCNN, and face recognition DL libraries to extract faces swiftly. BlazeFace and face recognition are particularly efficient for processing numerous images. Combining these three DL libraries enhances the accuracy of face detection. They stored face images in JPEG format, sized at 224 x 224 resolution. The dataset was split into training, validation, and test sets, containing 162,174 images in total. Specifically, 112,378 were allocated for training, 24,898 for validation, and another 24,898 for testing, maintaining a 70:15:15 ratio respectively. Both the real and fake image classes were equally represented across all sets.

4.1.2. Roc Curve:

A Receiver Operating Characteristic (ROC) curve represents the performance of a binary classification model graphically. At different threshold settings, it compares the true positive rate (sensitivity) to the false positive rate (1 - specificity). The schematics of ROC curve is shown in Fig.1

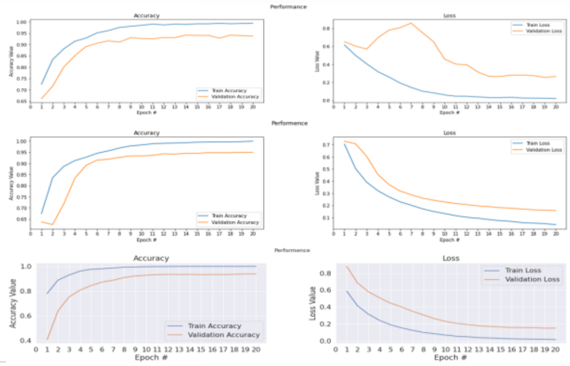


Fig. 1. ROC curve of all CNN models [38]

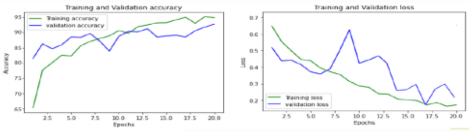


Fig. 2. ROC curve of CRNN model [38]

4.1.3. Result and Discussion:

Table 2. Result of CNN and RCNN model [38]

System Name	Architecture	Accuracy	Precision
CNN	InceptionResNetV2	93.75%	98.0%
	Mobile Net	94.2%	99.0%
	DenseNet121	93.86%	98.0%
RCNN	CNN+RNN	94.8%	94.4%

4.2. Dual Attention Network Approaches to Face Forgery Video Detection [44]:

(YI-XIANG LUO et al 2022)in this study, a Forgery Feature Attention Module (FFAM) and a Spatial Reduction Attention Block (SRAB) were integrated into the backbone network to construct a Dual Attention Forgery Detection Network (DAFDN). The two attention processes that have been presented are embedded by DAFDN, it additionally makes it possible for the convolution neural network to extract odd traces from the warped images. This study compares the effectiveness of the proposed DAFDN with other techniques using two benchmark datasets, DFDC and Face Forensics++. The results show that the proposed DAFDN approach achieves AUC values of 0.911 and 0.945, respectively, in the DFDC and Face Forensics++ datasets. These outcomes surpass those of earlier developed techniques like XceptionNet and Efficient Net-related techniques.

The whole process can be expressed as follows.

$$\begin{aligned} M_{sr}(F) &= \sigma\left(f^{7\times7}([f^{1\times1}(F); \text{MaxPool}\ (F)])\right) \\ &= \sigma\left(f^{7\times7}([F_c; F_{\max}])\right) \end{aligned} \tag{3}$$

Where σ denotes the sigmoid function; $f^{7\times7}$ and $f^{1\times1}$ represent that they were calculated via convolution, and the superscript stands in for the kernel size.

4.2.1. Data Set:

In order to assess how well DAFDN performs in identifying deep fake movies, Deep fake Detection Challenge (DFDC) and Face Forensics++ (FF++) are used as two benchmark datasets. A first-generation deep fake dataset, the FFCC dataset contains 1000 YouTube raw video sequences. The front of the face is not obscured in any of the movies because they were all manually chosen, making it possible for forgery techniques to produce lifelike forgeries. Two methods—classical computer graphics and deep learning—

can be used to generate counterfeit videos in the order they appear in the film.

4.2.2. Roc Curve:

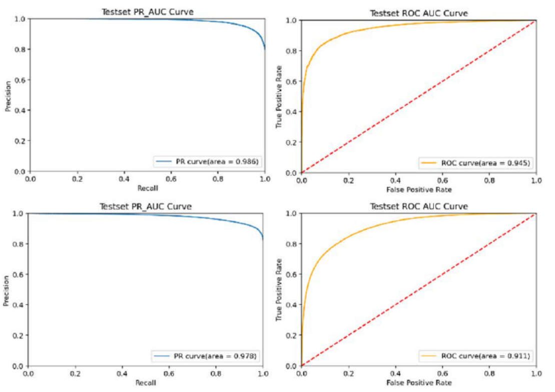


Fig. 3. PR curve and ROC curve of DAFDN on FFCC and DFDC [44].

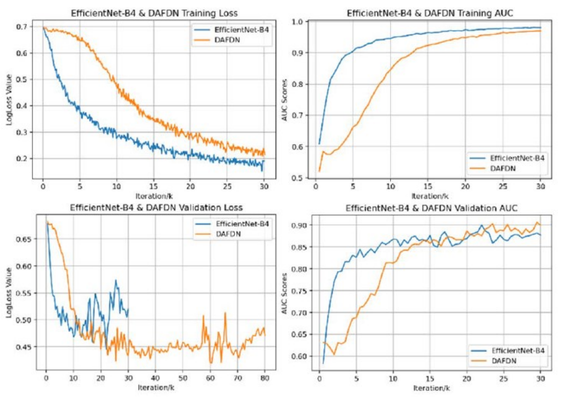


Fig. 4. Loss and AUC of EfficientNet-B4 and DAFDN [44].

4.2.3. Results of Analysis:

DAFDN received PR-AUC scores of 0.978 and 0.911 with DFDC, as well as 0.978 and 0.945 with FFCC. Both the FF++ and DFDC datasets show strong performance with DAFDN. Given that the test set is composed of data that have never been used for model training, the model's outstanding performance illustrates its great generalizability.

Table 3. Performance by DAFDN [39]

System Name	Architecture	Accuracy	Precision
DAFDN	DAFDN+ DFDC	97.8%	91.1%
DAFDN	DAFDN+ FF++	97.8%	94.5%

4.3. Deepfake Video Detection Through the Use of a Hybrid CNN Deep Learning Model :

(Sumaiya Thaseen Ikram et al 2023)with the use of numerous software programs and cutting-edge AI (Artificial Intelligence) technology, a number of fake films and images are created in the current era, leaving behind certain telltale evidence of manipulation. Videos may be used in a variety of unethical ways to intimidate, quarrel, or frighten others. Make sure that no fraudulent videos are produced using such techniques. Deep Fake is the name of an AI-based method for creating synthetic human photographs. They are produced by mixing and overlaying pre-existing videos over the original videos. In order to extract frame-level characteristics, a method that combines InceptionResnet v2 and Xception is built in this research. For experimental analysis, the DFDC deep fake detection challenge on Kaggle is used. The accuracy and training time of these deep learning-based algorithms are increased by using this dataset for both training and testing. The following results were obtained: accuracy 0.985, recall 0.96, f1-score 0.98, and support 0.968.

4.3.1. Data Set:

The DFDC dataset serves as the foundation for experiments, distinguishing itself from other deep fake datasets by collecting over 100,000 clips involving 3,426 paid actors. Unlike many other datasets that use non-consensual footage shot in controlled environments, the DFDC dataset stands out for its diverse collection of face swap videos sourced from various algorithms, including Deep fake, GAN-based, and non-learned methods. Notably, each of the 100,000 forged videos in this dataset presents a unique target/source switch, showcasing a wide array of scenarios spanning indoor and outdoor settings with different lighting conditions. Despite disruptions, DF-1.0 encompasses 1,000 distinct forged videos, contributing to the dataset's comprehensive nature and its status as the largest publicly available face swap video dataset.

4.3.2. Roc Curve:

The ROC curve for Sumaiya's suggested study is presented in Figure. 5. The genuine positive rate vs. false positive rate for various parameter cut-off points is plotted in this graph.

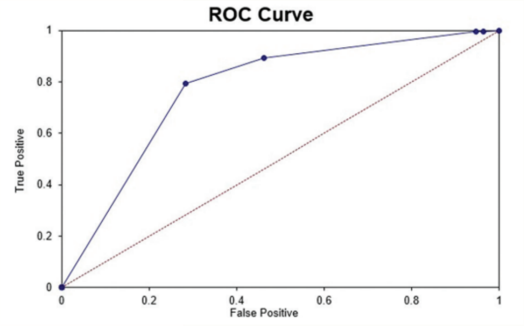


Figure. 5. ROC Curve for Training and Validation [45]

4.3.3. Result and Discussion

Table 4. Performance by CNN based face detection

System Name	Architecture	Accuracy	Precision	Recall	F1-Score
CNN	Inception	96%	97%	94%	93%
CNN	Xception	93%	98%	98%	91%
CNN	Hybrid Inception-Resnet v2 and Xception	98%	99%	97%	98%

5. Comparison of Previous Research:

Comparing previous deep learning-based face forensics results can provide insights into the developments and performance obtained by various models and methodologies.

Table 5. Comparative analysis

Author	System	Architecture	Accuracy	Precision
Ashifur Rahman et al [38]	**RCNN**	CNN+RNN	94.8%	94.4%
YI-XIANG LUO et al [39]	**DAFDN**	DAFDN+ FF++	97.8%	94.5%
Sumaiya Thaseen Ikram et al [40]	**CNN**	Hybrid Inception Resnet v2 and Xception	98%	99%

The table outlines a comparison among three distinct systems—RCNN, DAFDN, and CNN—focusing on their structure, precision, and accuracy. RCNN merges Convolutional Neural Networks (CNN) and Recurrent Neural Networks (RNN), achieving 94.8% accuracy and 94.4% precision. In contrast, DAFDN integrates DAFDN and FF++ elements, displaying exceptional performance with 97.8% accuracy and 94.5% precision. The CNN system combines Inception Resnet v2 and Xception models, yielding a remarkable 98% accuracy and an impressive precision of 99%. These statistics highlight how each system's design effectively handles tasks, offering valuable insights into their capabilities within neural networks.

Table 6. Models significance and limitation

Model	Advantages	Limitations
RCNN	Handles varied facial poses and orientations well	Computationally intensive
	Good at detecting faces even in cluttered backgrounds	Requires large datasets for training
	Can identify faces across different scales	May struggle with low-resolution or distorted images
DAFDN	Focuses on fine-grained facial features	Limited to frontal or near-frontal face orientations
	Robust against some common deepfake manipulations	May not generalize well to diverse deepfake variations
	Efficient and faster than some other deepfake models	Can be susceptible to adversarial attacks
CNN-based	Flexible architecture adaptable to different tasks	Performance highly dependent on data quality and quantity
	Can learn intricate patterns and features effectively	Prone to overfitting without proper regularization
	Handles varying lighting conditions well	Limited by training data availability and diversity

This study examined three existing deep learning methods that outperformed other existing techniques. This research assessed CNN+RNN, DAFDN, Hybrid Inception Resnet v2, and Xception, and the outcomes showed that the Inception Xception model outperforms RCNN and Dual Attention Network in terms of performance. Combining three models, such as Inception Net, ResNet, and Xception Net, generates a more complicated structure and increases computation time when compared to others. There are a few limitations in the Inception and Xception models since both are computationally costly due to their deep and complex topologies. They have a huge number of parameters and actions, making them longer to train and infer than simpler models. This computational complexity may restrict their usefulness in resource-constrained contexts or real-time applications where low latency is critical.

Pre-trained models obtained from extensive datasets (such as Image Net) are used as a starting point in transfer learning, and can help Xception Net. You may fine-tune Xception Net on face forensics datasets, which are often smaller, by exploiting the learned characteristics from these models. Furthermore, regularization approaches can aid in the prevention of over fitting and the improvement of generalization. Methods like dropout and batch normalization can help to regularize the model and limit the danger of over fitting on the training data. These methods motivate the model to learn more robust and generalizable features.

6. Future Scope:

Deep fake detection architecture holds great potential in the future, as advanced countermeasures are needed against the rise of sophisticated AI techniques that generate hyper-realistic manipulated content. Key players in this evolving landscape include Convolutional Neural Networks (CNNs), Region-based Convolutional Neural Networks (RCNNs), and the Domain Adaptive Few-Shot Detection Network (DAFDN). CNNs have been at the forefront of deep fake detection, but challenges remain in enhancing their robustness against evolving deep fake techniques. Future research may focus on improving interpretability, incorporating attention mechanisms, and exploring novel architectures to detect subtle artifacts and inconsistencies introduced by deep fake algorithms. RCNNs capture spatial relationships in images, but adapting to real-time processing and handling video streams efficiently remains a challenge. DAFDN, designed to adapt to domain shifts and few-shot scenarios, presents a promising direction for future research. However, challenges persist, such as adversarial attacks and ethical considerations surrounding privacy and consent in the deployment of deep fake detection technologies. In conclusion, the future of deep fake detection architecture is poised for continued innovation and refinement, with researchers navigating the evolving landscape of deep fake generation techniques, enhancing detection networks' speed and efficiency, and addressing ethical considerations to ensure responsible deployment in real-world scenarios.

7. Conclusion:

Only a limited number of individuals working in law enforcement, intelligence, and private investigations had any practical use for multimedia forensics. fourteen years ago. Both offense and defense had an artisanal feel and needed meticulous effort and commitment. Artificial intelligence has primarily modified these rules. Today, it appears that high-quality imitations are made on a production line, requiring extraordinary efforts from scientists and decision-makers. In actuality, today's multimedia forensics is fully developed, important organizations are supporting significant research projects, and experts from other fields are actively contributing with quick developments in concepts and techniques. This analysis will look at three studies that looked into CNN+RNN, DAFDN, Hybrid Inception Resnet v2, and Xception in relation to Face Forensics. According to the results, the Inception Xception model performs better than RCNN and Dual Attention Network.

Reference:

Shad, H. S., Rizvee, M. M., Roza, N. T., Hoq, S. M., Monirujjaman Khan, M., Singh, A. &amp; Bourouis, S. (2021). Comparative analysis of deepfake image detection method using convolutional neural network. Computational Intelligence and Neuroscience, 2021.

Rahman, A., Islam, M. M., Moon, M. J., Tasnim, T., Siddique, N., Shahiduzzaman, M. & Ahmed, S. (2022). A qualitative survey on deep learning based deep fake video creation and detection method, Aust. J. Eng. Innov. Technol, 4(1), 13-26.

Nguyen, T. T., Nguyen, C. M., Nguyen, D. T., Nguyen, D. T. & Nahavandi, S. (2019). Deep learning for deepfakes creation and detection, arXiv preprint arXiv:1909.11573, 1(2), 2.

Bode, L. (2021). Deepfaking Keanu: YouTube deepfakes, platform visual effects, and the complexity of reception, Convergence, 27(4), 919-934.

Verdoliva, L. (2020). Media forensics and deepfakes: an overview, IEEE Journal of Selected Topics in Signal Processing, 14(5), 910-932.

Kousik, N., Natarajan, Y., Raja, R. A., Kallam, S., Patan, R., & Gandomi, A. H. (2021). Improved salient object detection using hybrid Convolution Recurrent Neural Network. Expert Systems with Applications, 166, 114064.

Kamaleldin, M. G. M., Abu-Bakar, S. A. & Sheikh, U. U. (2023). Transfer Learning Models for CNN Fusion with Fisher Vector for Codebook Optimization of Foreground Features, IEEE Access.

Guo, Z., Yang, G., Chen, J. & Sun, X. (2021). Fake face detection via adaptive manipulation traces extraction network, Computer Vision and Image Understanding, 204, 103170.

George, A. S. & George, A. H. (2023). Deepfakes: The Evolution of Hyper realistic Media Manipulation, Partners Universal Innovative Research Publication, 1(2), 58-74.

Wang, T., Zhang, Y., Qi, S., Zhao, R., Xia, Z. & Weng, J. (2023). Security and privacy on generative data in aigc: A survey, arXiv preprint arXiv:2309.09435.

Ding, F., Zhu, G., Alazab, M., Li, X. & Yu, K. (2020). Deep-learning-empowered digital forensics for edge consumer electronics in 5G HetNets, IEEE consumer electronics magazine, 11(2), 42-50.

Karie, N. M., Kebande, V. R., & Venter, H. S. (2019). Diverging deep learning cognitive computing techniques into cyber forensics, Forensic Science International: Synergy, 1, 61-67.

Chintha, A., Thai, B., Sohrawardi, S. J., Bhatt, K., Hickerson, A., Wright, M., & Ptucha, R. (2020). Recurrent convolutional structures for audio spoof and video deepfake detection, IEEE Journal of Selected Topics in Signal Processing, 14(5), 1024-1037.

Wu, B., Su, L., Chen, D., & Cheng, Y. (2023). FPC-Net: Learning to detect face forgery by adaptive feature fusion of patch correlation with CG-Loss, IET Computer Vision, 17(3), 330-340.

Sedik, A., Faragallah, O. S., El-sayed, H. S., El-Banby, G. M., El-Samie, F. E. A., Khalaf, A. A. & El-Shafai, W. (2022). An efficient cybersecurity framework for facial video forensics detection based on multimodal deep learning, Neural Computing and Applications, 1-18.

Wu, J., Zhu, Y., Jiang, X., Liu, Y., & Lin, J. (2023). Local attention and long-distance interaction of rPPG for deepfake detection, The Visual Computer, 1-12.

Coccomini, D. A., Messina, N., Gennaro, C. & Falchi, F. (2022, May). Combining efficientnet and vision transformers for video deepfake detection, In Image Analysis and Processing–ICIAP 2022: 21st International Conference, Lecce, Italy, May 23–27, 2022, Proceedings, Part III, Cham: Springer International Publishing, 219-229.

Aishwarya Rajeev, A., & Raviraj, P. (2023). An Insightful Analysis of Digital Forensics Effects on Networks and Multimedia Applications, SN Computer Science, 4(2), 186.

Xiao, J., Li, S., & Xu, Q. (2019). Video-based evidence analysis and extraction in digital forensic investigation, IEEE Access, 7, 55432-55442.

Ahmadi, F., Gupta, G., Zahra, S. R., Baglat, P. & Thakur, P. (2021, March). Multi-factor biometric authentication approach for fog computing to ensure security perspective, In 2021 8th international conference on computing for sustainable global development (INDIACom), IEEE, 172-176.

Ross, A., Banerjee, S., & Chowdhury, A. (2020). Security in smart cities: A brief review of digital forensic schemes for biometric data, Pattern Recognition Letters, 138, 346-354.

Chandaliya, P. K., & Nain, N. (2022). ChildGAN: Face aging and rejuvenation to find missing children, Pattern Recognition, 129, 108761.

Ivanova, E., & Borzunov, G. (2020). Optimization of machine learning algorithm of emotion recognition in terms of human facial expressions, Procedia Computer Science, 169, 244-248.

Sikkandar, H., & Thiyagarajan, R. (2020). Soft biometrics-based face image retrieval using improved grey wolf optimisation, IET Image Processing, 14(3), 451-461.

Keshari, T., & Palaniswamy, S. (2019, July). Emotion recognition using feature-level fusion of facial expressions and body gestures, In 2019 international conference on communication and electronics systems (ICCES), IEEE, 1184-1189.

Hussain, S. A. & Al Balushi, A. S. A. (2020). A real time face emotion classification and recognition using deep learning model, In Journal of physics: Conference series, IOP Publishing, 1432(1), 012087.

Ashwin, T. S. & Guddeti, R. M. R. (2019). Unobtrusive behavioral analysis of students in classroom environment using non-verbal cues, IEEE Access, 7, 150693-150709.

Agbolade, O., Nazri, A., Yaakob, R., Ghani, A. A. & Cheah, Y. K. (2019). 3-Dimensional facial expression recognition in human using multi-points warping, BMC bioinformatics, 20(1), 1-15.

Bozkir, E., Özdel, S., Wang, M., David-John, B., Gao, H., Butler, K. & Kasneci, E. (2023). Eye-tracked Virtual Reality: A Comprehensive Survey on Methods and Privacy Challenges. arXiv preprint arXiv:2305.14080.

Dargan, S. & Kumar, M. (2020). A comprehensive survey on the biometric recognition systems based on physiological and behavioral modalities, Expert Systems with Applications, 143, 113114.

Verdoliva, L. (2020). Media forensics and deepfakes: an overview, IEEE Journal of Selected Topics in Signal Processing, 14(5), 910-932.

Hashmi, M. F., Ashish, B. K. K., Keskar, A. G., Bokde, N. D., Yoon, J. H. & Geem, Z. W. (2020). An exploratory analysis on visual counterfeits using conv-lstm hybrid architecture, IEEE Access, 8, 101293-101308.

Deshmukh, A. & Wankhade, S. B. (2020). Deepfake Detection Approaches Using Deep Learning: A Systematic Review, Intelligent Computing and Networking: Proceedings of IC-ICN 2020, 293-302.

Awotunde, J. B., Jimoh, R. G., Imoize, A. L., Abdulrazaq, A. T., Li, C. T. & Lee, C. C. (2022). An Enhanced Deep Learning-Based DeepFake Video Detection and Classification System, Electronics, 12(1), 87.

Byrnes, O., La, W., Wang, H., Ma, C., Xue, M., & Wu, Q. (2021). Data hiding with deep learning: A survey unifying digital watermarking and steganography, arXiv preprint arXiv:2107.09287.

Chen, L., Zhang, Y., Song, Y., Liu, L. & Wang, J. (2022). Self-supervised learning of adversarial example: Towards good generalizations for deepfake detection, In Proceedings of the IEEE/CVF conference on computer vision and pattern recognition, 18710-18719.

Xi, Z., Niu, Y., Chen, J., Kan, X. & Liu, H. (2020). Facial expression recognition of industrial internet of things by parallel neural networks ecombining texture features, IEEE Transactions on Industrial Informatics, 17(4), 2784-2793.

Jeong, D., Kim, B. G., & Dong, S. Y. (2020). Deep joint spatiotemporal network (DJSTN) for efficient facial expression recognition, Sensors, 20(7), 1936.

Zhu, M., Shi, D., Zheng, M., & Sadiq, M. (2019). Robust facial landmark detection via occlusion-adaptive deep networks, In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 3486-3496.

Zhang, W., & Zhao, C. (2019, November). Exposing face-swap images based on deep learning and ELA detection, In Proceedings, MDPI, 46(1), 29.

Zhao, H., Zhou, W., Chen, D., Wei, T., Zhang, W., & Yu, N. (2021). Multi-attentional deepfake detection. In Proceedings of the IEEE/CVF conference on computer vision and pattern recognition, 2185-2194.

Bonomi, M., Pasquini, C., & Boato, G. (2021). Dynamic texture analysis for detecting fake faces in video sequences, Journal of Visual Communication and Image Representation, 79, 103239.

Rahman, A. (2022). Deepfake Video Detection Using CNN and RCNN (Doctoral dissertation, Bangladesh University of Business and Technology).

Luo, Y. X. & Chen, J. L. (2022). Dual Attention Network Approaches to Face Forgery Video Detection. IEEE Access, 10, 110754-110760.

Ikram, S. T., Chambial, S., & Sood, D. (2023). A performance enhancement of deep fake video detection through the use of a hybrid CNN Deep learning model. International journal of electrical and computer engineering systems, 14(2), 169-178.

AUTHOR BIOGRAPHIES

Aishwarya Rajeev is currently working as an Assistant Professor & Head in the Department of Artificial Intelligence and Data Science at Coorg Institute of Technology, Ponnampet, Karnataka. She holds an M.E in Computer Science and Engineering with first rank and gold medal from Mahendra Engineering College, Affiliated to Anna University, Chennai, India in 2015. She also holds a MBA in IT & Systems from ICFAI University, Tripura, India in 2012. She received B. Tech Degree in Information Technology from Cochin University of Science and Technology, India in 2008. She has 15.6 years of experience in teaching has published papers, and attended various conferences and workshops. She is also a life member of professional bodies like ISTE, IE, and IAENG. She serves as Editorial Board Member and Reviewer of 2 International Journals. She can be contacted at email: aishwaryarajeev@gmail.com.

P. Raviraj completed his doctorate degree in Computer Science and Engineering in the area of Image Processing. He holds the position of Director-IQAC and Professor & Head in the Department of Computer Science & Engineering at GSSS Institute of Engineering & Technology for Women, Mysore, Karnataka. He has 19 years of teaching and research experience. He has published more than 94 papers in International journals and conferences. Five research scholars have completed their Ph.D. under his guidance at various universities. At present, he has guiding Ph.D. research scholars in the areas of Image Processing, Pervasive & Cloud computing, Bio-Inspired Algorithms Robotics etc. He has received the project grant Rs.5 Lakhs from the VGST, Govt. of Karnataka for the “Underwater Robotic Fish for Surveillance and Pollution monitoring”. He filed the patent entitled “An effective ROI based Hybrid Progressive Medical Image Transmission and Reconstruction” in the year 2021. He is serving as a Ph.D thesis Adjudicator, Doctor Committee member and Subject expert for various universities. He has served as a chairperson and keynote speaker at many National and International Conferences. He serves as an Editorial Board Member and Reviewer for more than 15 International Journals. He is also a life member of professional bodies like ISTE, CSI etc. He has received awards and recognitions such as ‘Rhastriya Gaurav Award-2015’, ‘Shri P.K.Das Memorial Best Faculty Award-2012’, ‘Young Achiever Award-2016’, ‘Best Circuit Faculty Finalist Award-2017’ in his credit.

DOI: 10.6977/IJoSI.202403_8(1).0006
H. Orovwode, S. Matthew./Int. J. Systematic Innovation, 8(1), 63-69 (2024)

Development of a solar system for charging mobile phones with customized DC chargers for rural areas in Nigeria

Hope Orovwode[1]*, Simeon Matthew[2], Oluwaseun I. Adebisi[2], Ayorinde J. Olanipekun[2], Elizabeth O. Amuta[1]
[1] Department of Electrical and Information Engineering, Covenant University, Ota, Nigeria
[2] Department of Electrical and Electronics Engineering, Federal University of Agriculture, Abeokuta, Nigeria
* Corresponding author E-mail: hope.orovwode@covenantuniversity.edu.ng

(Received 24 August 2023; Final version received 15 January 2024; Accepted 29 January 2024)

Abstract

The introduction of mobile phones has redefined the world of communication in that it has turned the world into a global village as people can now make contact through phone calls within and across countries at affordable rates. Nowadays, almost every home has a functional mobile phone, be it a conventional, android, or iPhone, among others. These phones use non-self-charging rechargeable batteries that need to be recharged from time to time to meet the demands of the users. However, access to a reliable source of power to meet the energy demand, including mobile phone charging needs, of off-grid rural dwellers remains a global challenge. As a result, this study designed and implemented a solar-powered mobile phone charging system with customized DC chargers for use in remote off-grid areas. The test results showed that the system is very effective for charging mobile phones.

Keywords: Customized DC chargers, mobile phones, Remote off-grid areas, solar system.

1. Introduction

The introduction of mobile phones has redefined the world of communication in that it has turned the world into a global village(Abbas et al., 2019)as people can now make contact through phone calls within and across countries at affordable rates. Modern mobile phones are designed to deliver smart and seamless performance with a high level of affordability, flexibility, and reliability. They are built with several mobile applications such as Wi-Fi, GPRS, HSCSD, high-definition cameras, sound and video players, and USB support systems, among others. As a result, billions of mobile phone users (be they smart or non-smartphones) exist across the world(Chaudhary & Vrat, 2018), (Panigrahi et al., 2020), (Saxena & Saxena, 2020) including remote rural dwellers.

Mobile phones are powered by rechargeable lithium-ion (Li-ion) batteries (Cui et al., 2018), (Ghiji et al., 2020)whose strengths are specified in ampere-hours (Ah). The rates at which the batteries get drained (rundown) depend on the usage as well as the applications on the phone. These batteries need frequent recharging in other to meet the demands of the users. Consequently, mobile phones are provided with alternating current (ac) based battery chargers. However, it is not every user of mobile phones that has access to a sustainable alternating current source for charging the phone. Most rural dwellers fall into the category of mobile phone users that lack access to sustainable electricity for frequent charging of their devices. Some mobile phone users depend on fossil fuel-based power generators for phone charging purposes and can go the extra mile to put on their generators at any time of the day just to recharge their mobile phone batteries thereby contributing to environmental pollution and incurring a very high cost-benefit ratio.

Furthermore, some mobile phone users in rural areas that are not privileged owners of generators may have to defer the time to charge their devices to align with the time when the privileged owners of generators are most likely to put on their generators while others visit charging booths operating on fossil fuel-based generators to pay for charging services. The charges paid per phone in such charging booths are determined by the prevailing prices as well as the availability of fossil fuels which have recently become very scarce in most developing countries. As a result, there is a need to develop alternative means for charging mobile phones. One such means is the use of solar energy which is naturally available almost everywhere with a low environmental burden(Annuk et al., 2020), (Simeon et al., 2018),(Bataev et al., 2020), (Abass & Pavlyuchenko, 2019) for phone charging purposes. Thus, this study seeks to develop a solar-powered mobile phone charging booth with customized chargers for rural dwellers in Nigeria

Literature has shown that several authors have carried out various degrees of work on solar-powered mobile phone charging booths/kiosks. For example, Louie et al(Louie et al., 2015) designed and implemented a solar power kiosk for charging mobile phones and other electronic devices in rural Zambia. Similarly, Palmiro, Rayudu, and Ford (Palmiro et al., 2015)modeled and simulated a solar-powered kiosk for charging Lithium-ion batteries. Also, the authors (Dauenhauer et al., 2019) assessed the impact of solar-powered kiosks in Zambia. Other works on solar charging kiosks are presented in (Shoarinejad & Shokri, 2016), (Munro & Christiansen, 2016), (Frame et al., 2019), (Udayalakshmi & Sheik, 2018), and (Shoewu & Salau, 2018).

Despite that, a lot of work has been carried out on solar-powered charging booths/kiosks, but the design of a charging boot with customized chargers has not been presented. Thus, in this study, an ingredient of novelty involving the re-design or configuration of waste or thrown- out ac powered mobile phone chargers whose circuit elements might have been destroyed or burnt out as dc-dc converters for tapping the solar power for charging mobile phones is presented. By so doing waste mobile phone chargers are recycled for wealth creation toward achieving the 2030 sustainable development goals of the United Nations.

5. Materials and methods

The materials used to implement the solar system for charging mobile phones with customized dc chargers include waste ac based mobile phone chargers, a solar panel, a solar battery, a Charge controller, charging outlets (13 A double sockets), connecting cables, and a wooden cabinet.

The method used in this work can be broken down into two parts. The first part deals with the design of solar-powered mobile phone charging systems while the second part deals with the sourcing and configuration of waste ac based mobile phone chargers to suit the need at hand.

2.1 The solar powered mobile phone charging system design

The block diagram of the proposed solar system for charging mobile phones is shown in Fig. 1.

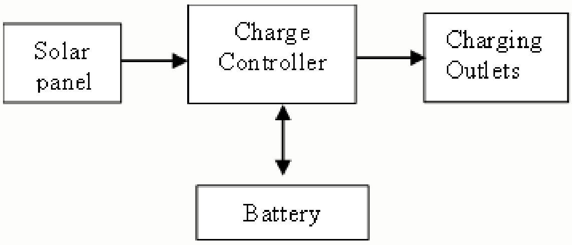


Fig. 1. Block diagram of the solar system for charging mobile phones.

Thorough and accurate sizing of the solar panel, the charge controller, and the battery is required for optimal performance and cost minimization while the choice of the charging outlets is influenced by the ease with which the customized charger can fit in. As a result, the detailed sizing of the solar panel, the charge controller, and the battery is carried out in this study

2.1.1 Load specification

The first point of reference in any solar system design is the load the system is expected to power which in this study are mobile phones whose battery specifications are used to determine the load size.

Several varieties (or makes) of mobile phones with different battery specifications exist as reported in (Liang et al., 2019) and (Diouf et al., 2015). The authors (Liang et al., 2019) reported the highest existing mobile phone battery capacity to be 3.82V, 4200 mAH (for Huawei mate 20p). To ensure that the proposed system can charge any type of mobile phone, the battery specifications for the Huawei mate 20p were used to size the components of the solar system used in the design of the mobile phone charging system. The load data determined from the battery specification is shown in Table 1.

Table 1. Load data for the solar mobile phone charging system

Type of Load	Voltage (V)	Ampere-hour (Ah)	Watt-hour (W-h)	Quantity, Q (units)	Total W-h
Mobile Phone	3.82	4.2	16.044	48	770.0112

2.1.2 Sizing of the solar battery storage

The solar battery storage for the system was sized using equation (1) (Orovwode et al., 2022).

$$B_S = \frac{T_{Wh} \times N}{B_V \times DoD \times \epsilon_{cable}} \qquad (1)$$

Where;
B_S =Minimum battery capacity (Ah)
T_{Wh} =Maximumwatt-hour required/day
N = Number of days of autonomy
B_V = DC voltage of the battery (V)
DoD = Depth of discharge allowed for the battery

Using a 12 V_{DC} battery storage system used with the other parameters in equation (1) defined as follows:
T_{Wh} = 770.0112 Wh (from table 1)

N = 1 day, the choice of which is influenced by 0% diversity

ϵ_{cable}= Efficiency of the connecting cables linking the battery and the loads = 98%

$$B_S = \frac{770.0112 \times 1}{12 \times 0.7 \times 0.98} = 93.54 \text{ Ah}$$

Therefore 100 Ah which is the nearest available size of battery was used in this design.

2.1.3 Sizing of the solar PV panel

The solar mobile phone charging system designed in this study is intended for use at Abule-Ticha which is a remote rural area in Ado Ota Local Government Area of Ogun State. Since the Local Government shares a boundary with Lagos State, the NASA data on the solar insolation for Lagos, Nigeria was used to determine the size of the required PV panel.

The size of the required PV panel was determined using equation (2) (Orovwode et al., 2018):

$$S_{PV} = \frac{T_{Wh}}{N_{ph} * \epsilon_{sys}} \qquad (2)$$

Where;
S_{PV} = Total wattage of the required panels (Watts)
T_{Wh} = MaximumWatt-hour required per day
N_{ph} = peak hours per day = $5.43 kW/m^2/day$ (NASA Data for Lagos)
ϵ_{sys} = Total system efficiency

But, T_{Wh} = $770.0112Wh$ with zero tolerance due to the adoption of 0% diversity

$$\epsilon_{sys} = \epsilon_{PV} \times \epsilon_{cable1} \times \epsilon_{cc} \times \epsilon_{Batt} \times \epsilon_{c2} \qquad (3)$$

Where;

ϵ_{PV} = $PV\ modulesEfficiency\ = 80\%$
ϵ_{cable1}= $Efficiency\ of\ the\ connecting\ cables\ linking\ the\ solar\ panel\ array\ and\ the\ battery = 95\%$
ϵ_{cc} = efficiency of charge controller = 90%
ϵ_{Batt}= $Battery\ efficiency = 90\%$
ϵ_{cable}= Efficiency of the connecting cables linking the battery and the loads = 0.95

$$\therefore \epsilon_{sys} = 0.8 \times 0.95 \times 0.9 \times 0.90 \times 0.95 = 0.58$$

$$\therefore S_{PV} = \frac{770.0112}{5.43 \times 0.58} = 244.49 \text{ W}$$

To make the system more efficient, a 280 W mono-crystalline panel was used.

2.1.4 Sizing of charge controller

The charge controller employed in the design is the Pulse Width Modulated (PWM) type whose size is given by equation (4) (Benjamin & Dickson, 2017)

$$CCS = I_{sp} \times S_f \qquad (4)$$

Where;
CCS = size of the charge controller,
I_{sp} =specified short circuit current of the panel,
S_f = safety factor = 1.25

Now, from the nameplate of the 280 W panel (shown in Fig. 2) used, I_{sp} = 10.38 A

Putting the values of I_{sp} into equation (4) will give;

$$CCS = 10.38 \times 1.25 = 13.03 \text{ A}$$

The nearest available charge controller size (20 A) was used in this design.

SUNPOWER SOLAR MODULE 280 mono solar panel	
Maximum Power/Pmax(W)	250
Maximum Power tolerance (%)	X3%
Open Circuit Voltage/ Voc(V)	36.85
Short-Circuit Current / StC(A)	10.38
Max Power Voltage/ Vmp(V)	31.5
Max Power Current/Imp(A)	8.88
Power Spectiations at STC: 1000W/m^2, AM1.5, Cell 25 C	
Weight (kg)	19.3
Dimensions(mm)	1640*952*35
Max System Voltage (V)	1000
Max Over current Protecting rating(A)	1.5
efficiency	16.65
Module Application Circuit	A

Fig. 2. Name plate of the 280 W panel used

2.2 Sourcing and reconfiguration of AC-based waste chargers

The solar charging system designed in this study is purely a direct current (DC) system requiring no inverter for energy conversion. By so doing, the system cost is minimized. However, the fact that mobile phones are provided with AC-based battery chargers becomes a serious bottleneck to overcome. To overcome the bottleneck, this study sourced and reconfigured waste/burnt-out AC-based chargers to form the desired DC chargers which are code-named 'customized DC chargers'.

The waste/burnt-out AC-based mobile phone chargers were sourced from several households at the Indomie Estate, Atan, a community in the Ado-Odo Local Government Area of Ogun State, Nigeria. From the waste chargers gathered, only the three-pin plugs were configured for use to avoid polarity interchange. The internal circuitries of the three-pin plugs waste chargers were reconfigured using the circuit diagram of Fig. 3. The choice of the LM7805 voltage regulator was motivated by the fact that a regulated voltage of 5 V is required to charge a mobile phone (Ramli et al., 2019), (Hanif et al., 2020), (Maulidyna et al., 2021).

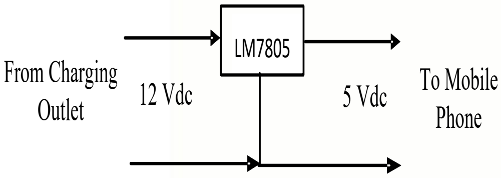


Fig. 3. The circuit diagram of the customized DC mobile phone charger

3 Implementation and testing

3.1 Implementation of the solar mobile phone charging system

The implementation of the solar charging system involved two stages including the implementation of the kiosk and that of the customized chargers

3.1.1 Implementation of the solar charging kiosk

The solar charging kiosk was implemented in a wooden enclosure shown in Fig. 4. The upper compartments (one on each side) of the enclosure were used for mounting the charging outlets (13 A double sockets) as shown in Fig. 5 while the lower compartment housed the battery and the charge controller as shown in Fig. 6. The solar panel was mounted on top of the enclosure with tilt-adjustable hangers for maximum solar energy harvesting. Provisions were also made for two wheels at the base of the kiosk for ease of mobility.

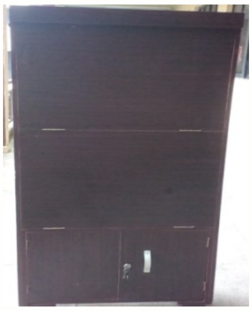


Fig. 4. The solar charging kiosk with all the doors closed

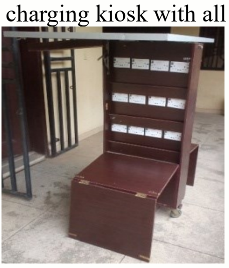


Fig. 5. The kiosk showing the arrangement of the charging Outlets

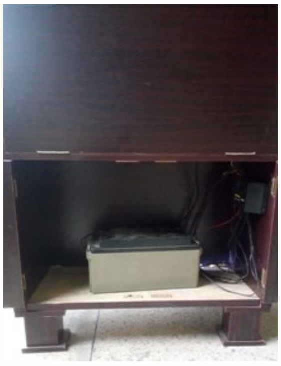


Fig. 6. The base compartment of the kiosk showing the battery and charge controller arrangement

3.1.2 Implementation of the customized DC chargers

The Customized DC chargers were implemented by opening the links between the charging ports of the waste charges and the input terminals (the plugs) and replacing them with the configuration shown in Fig. 3. The front and the back view of the internal circuitries of a sample of the implemented charger are shown in Fig. 7 and Fig. 8 respectively while the finished charger is shown in Fig. 9.

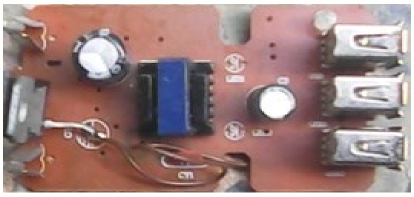


Fig. 7. The Front view of the internal circuitry of the customized DC charger

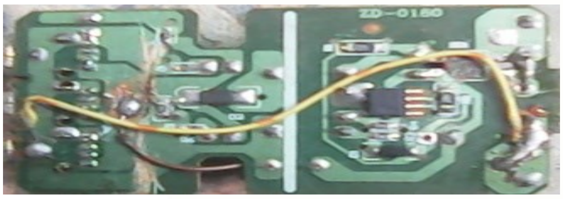


Fig. 8. The back view of the internal circuitry of the customized DC charger

Fig. 9. The customized DC charger

3.2 Testing

The Kiosk was tested to verify its usability for charging mobile phones. A sample of the photograph taken during the testing is shown in Fig. 10.

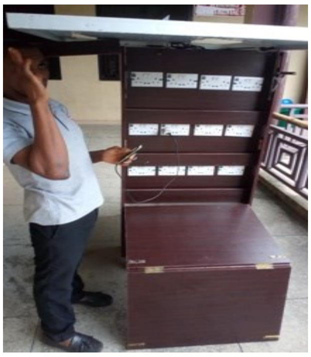


Fig. 10. Sample system testing photograph

The system was tested on six different types of Android mobile phones. The mobile phones with the batteries drained to various degrees were plugged into the system for charging and monitored for three hours. The result of the test is tabulated as shown in Table 2.

Table 2. Mobile phone charging test results

Battery Type	Initial Charge (%)	Time Used (hr)	Final Charge (%)
Umidigi A3S	10%	3	100
Itel A33	0%	3	100
Samsung Galaxy A03	20%	3	100
Infinix Hot 5	13%	3	100
Tecno Spark 8P	10%	3	100
Tecno POP 5 Pro	30%	3	100

4. Conclusion

In this study, the availability of electrical power for charging mobile phones was identified as one of the challenges faced by most rural dwellers. The increasing cost of fossil fuel and the associated dangers of carbon emission as well as the quest for conversion of waste to wealth for poverty alleviation motivated the authors to design and implement a mobile solar-powered mobile phone charging system with customized chargers. The system was designed with the capacity to charge up to 48 mobile phones at a time. The charging test conducted on the system validated the effectiveness of the system for charging mobile phones. Further works shall consider the techno-economic analysis of the system

Acknowledgment

The authors are grateful to the editorial crew of the International Journal of Systematic Innovation for thoroughly reviewing this article to improve its quality. We are more importantly grateful for publishing this article at no cost given the worst economic situation we have found ourselves in Nigeria at the moment.

References

Abass, A. Z., & Pavlyuchenko, D. A. (2019). The exploitation of western and southern deserts in Iraq for the production of solar energy. International Journal of Electrical and Computer Engineering (IJECE), 9(6), 4617–4624. https://doi.org/10.11591/ijece.v9i6.pp4617-4624

Abbas, J., Aman, J., & Nurunnabi, M. (2019). The Impact of Social Media on Learning Behavior for Sustainable Education : Evidence of Students from Selected Universities in Pakistan. Sustainability, 11, 1–23. https://doi.org/10.3390/su11061683

Annuk, A., Hovi, M., Kalder, J., Kabanen, T., Ilves, R., Märss, M., Martinkauppi, B., & Miidla, P. (2020). Methods for Increasing Shares of Self-Consumption in Small PV Solar Energy Applications. 9th International Conference on Renewable Energy Research and Application (ICRERA), 184–187. https://doi.org/10.1109/ICRERA49962.2020.9242902

Bataev, A., Potyarkin, V., Glushkova, A., & Samorukov, D. (2020). Assessment of development effectiveness of solar energy in Russia. E3S Web of Conferences, 221, 1–7. https://doi.org/10.1051/e3sconf/202022103002

Benjamin, E. A., & Dickson, E. (2017). Estimating the Solar Home System Sizing for Rural Residential Apartments Using a Panel Tilt Angle of 82 Degrees : Ilorin, Kwara State as Case Study. American Journal of Electrical and Computer Engineering, 1(3), 90–96. https://doi.org/10.11648/j.ece.20170103.13

Chaudhary, K., & Vrat, P. (2018). Circular economy model of gold recovery from cell phones using system dynamics approach : a case study of India. Environment, Development and Sustainability, 22(173–200).

Cui, Q., Zhong, Y., Pan, L., Zhang, H., Yang, Y., & Liu, D. (2018). Recent Advances in Designing High-Capacity Anode Nanomaterials for Li-Ion Batteries and Their Atomic-Scale Storage Mechanism Studies. Advanced Science, 5, 1–22. https://doi.org/10.1002/advs.201700902

Dauenhauer, P., Lauer, J. W., Louie, H., Sloughter, J. M., Lacrampe, C., Smith, C., Smith, E., Ohara, J., & Sebhat, N. (2019). Impact Assessment of Energy Kiosks in Rural Zambia. 2019 IEEE Global Humanitarian Technology Conference (GHTC), 1–8. https://doi.org/10.1109/GHTC46095.2019.9033068

Diouf, B., Pode, R., & Osei, R. (2015). Recycling mobile phone batteries for lighting. Renewable Energy, 78, 509–515. https://doi.org/10.1016/j.renene.2015.01.034

Frame, D., Dauenhauer, P., Eales, A., & Galloway, S. (2019). Sustainability of Solar PV Energy Kiosks for Off-Grid Energy Access: Malawi Case Study. IEEE Global Humanitarian Technology Conference (GHTC), 1–8. https://doi.org/10.1109/GHTC46095.2019.9033449

Ghiji, M., Novozhilov, V., Moinuddin, K., Joseph, P., Burch, I., Suendermann, B., & Gamble, G. (2020). A Review of Lithium-Ion Battery Fire Suppression. Energies, 13, 1–30.

Hanif, M. H. M., Fahmi, M. I., Wai, C. L., Aihsan, M. Z., Aminudin, A., Zhe, L. W., & Zakariya, M. Z. (2020). Maximum efficiency scheme using superimposed and Taguchi method wireless charging for mobile phone Maximum efficiency scheme using superimposed and Taguchi method wireless charging for mobile phone. Journal of Physics: Conference Series, 1432 012015. https://doi.org/10.1088/1742-6596/1432/1/012015

Liang, Y., Zhao, C., Yuan, H., Chen, Y., Zhang, W., Huang, J.-Q., Yu, D., Liu, Y., Titirici, M.-M., Chueh, Y.-L., Yu, H., & Zhang, Q. (2019). A review of rechargeable batteries for portable electronic devices. InfoMat., 1, 6–32. https://doi.org/10.1002/inf2.12000

Louie, H., Shields, M., Szablya, S. J., Makai, L., & Shields, K. (2015). Design of an off-grid energy kiosk in rural Zambia. 2015 IEEE Global Humanitarian Technology Conference (GHTC), 1–6. https://doi.org/10.1109/GHTC.2015.7343946

Maulidyna, A. N., Sudiharto, I., & Murdianto, F. D. (2021). Multi DC Load Single Port Output Adaptive Power Charge Using Fuzzy Logic Controller. IOP Conf. Series: Material Science and Engineering, 1096 012063. https://doi.org/10.1088/1757-899X/1096/1/012063

Munro, P., & Christiansen, A. (2016). charging station model in Sierra Leone. Progress in Development Studies, 16(1), 24–38. https://doi.org/10.1177/1464993415608080

Orovwode, H., Afolabi, G., Agbetuyi, F., Adoghe, A., & Temitope, M. (2022). Development and Performance Evaluation of a Solar Powered Tomatoes Storage Chamber. IOP Conf. Series: Earth and Environmental Science, 1054, 1–13. https://doi.org/10.1088/1755-1315/1054/1/012043

Orovwode, H., Wara, S., Mercy, T. J., Abudu, M., Adoghe, A., & Ayara, W. (2018). Development and Implementation of a Web Based Sustainable Alternative Energy Supply for a Retrofitted Office. 2018 IEEE PES/IAS PowerAfrica, 390–395.

Palmiro, F., Rayudu, R., & Ford, R. (2015). Modelling and simulation of a solar PV lithium ion battery charger for energy kiosks application. IEEE PES Asia-Pacific Power and Energy Engineering Conference (APPEEC), 1–5. https://doi.org/10.1109/APPEEC.2015.7381001

Panigrahi, S. K., Pathak, V. K., Kumar, M. M., Raj, U., & P, K. P. (2020). Covid-19 and mobile phone hygiene in healthcare settings. BMJ Global

Health, 12–14. https://doi.org/10.1136/bmjgh-2020-002505

Ramli, N., Rusli, M. R., Ahmad, I., Halim, A., Rahman, A., Sapiee, A., Ramli, N., Rusli, M. R., Ahmad, I., Rahman, H. A., & Sapiee, N. A. (2019). Application of water and wind energy for low cost portable mobile phone charger. AIP Conference Proceedings, 020096(July).

Saxena, M., & Saxena, A. (2020). Evolution of mHealth Eco-System: A Step Towards Personalized Medicine. In A. Khanna, D. Gupta, S. Bhattacharyya, V. Snasel, J. Platos, & A. E. Hassanien (Eds.), International Conference on Innovative Computing and Communications (pp. 351–370). Springer Singapore.

Shoarinejad, S., & Shokri, P. (2016). Charging Systems of Electronic Devices in Urban Open Spaces, A Need Analysis and Design. Modern Applied Science, 10(6), 213–218. https://doi.org/10.5539/mas.v10n6p213

Shoewu, O. O., & Salau, N. O. (2018). Design, Development, and Construction of a Solar Powered Phone Charging Box Design. Journal of Computation in Biosciences and Engineering, 3(4), 1–5.

Simeon, M., Adoghe, A. U., Wara, S. T., & Oloweni, J. O. (2018). Renewable Energy Integration Enhancement Using Energy Storage Technologies. IEEE PES/IAS PowerAfrica, 864–868. https://doi.org/10.1109/PowerAfrica.2018.8521075

Udayalakshmi, J. K., & Sheik, M. S. (2018). Design and Implementation of Solar Powered Mobile Phone Charging Station for Public Places. International Conference on Current Trends towards Converging Technologies (ICCTCT), 1–5. https://doi.org/10.1109/ICCTCT.2018.8551180

AUTHOR BIOGRAPHIES

Engr. (Dr.) Orovwode, Hope holds a Ph.D. in Electrical/Electronics Engineering from Covenant University and M. Eng and B. Eng degrees in Electrical / Electronics Engineering (Electronics & Telecommunications option) from the University of Benin, Benin City, Nigeria. He is a Senior Lecturer/Researcher in the Department of Electrical and Information Engineering of Covenant University, Ota, Nigeria
Email: Hope.orovwode@covenantniversity.edu.ng.

Engr. Matthew, Simeon is a lecturer in the Department of Electrical and Electronics Engineering, Federal University of Agriculture Abeokuta (FUNAAB). He holds B. Eng. in Electrical and Electronics Engineering from the Federal University of Agriculture, Makurdi, Benue State, Nigeria, and an M. Eng. degree in Electrical Machines and Power Systems Engineering from FUNAAB where he is currently undertaking his Ph.D. research.
email: matthews@funaab.edu.ng.

Engr. (Dr.) Oluwaseun I. Adebisi obtained his B.Eng., M.Eng., and Ph.D. degrees in Electrical and Electronics Engineering from the Federal University of Agriculture, Abeokuta (FUNAAB), Nigeria in 2010, 2013, and 2019 respectively. He specializes in power system engineering and electrical machines. He began his academic career in the Department of Electrical and Electronics Engineering, FUNAAB in 2011 as a Junior Research Fellow and he is currently a Senior Lecturer in the Department.
email: adebisioluwaseun@funaab.edu.ng.

Engr. (Dr.) Olanipekun Ayorinde Joseph had his B. Eng in Electrical and Electronics Engineering in 2000 and M. Eng with a specialty in Electronics and Control Systems Engineering in 2013. He has over Ten (10) years of experience in engineering Laboratory and 5 years of teaching experience in Electrical and Electronics engineering. His areas of interest are Energy conversion, electronics and control, smart systems, and computer-aided design.
email: olanipekunaj@funaab.edu.ng

Engr. (Dr) Elizabeth Amuta currently works at the Department of Electrical and Information Engineering, Covenant University Ota Ogun State, Nigeria. Amuta does research in Renewable energy, Electronic Engineering, and Electrical Engineering.
email: Elizabeth.amuta@covenant university.edu.ng.

INSTRUCTIONS TO AUTHORS

Submission of papers

The International Journal of Systematic Innovation is a refereed journal publishing original papers four times a year in all areas of SI. Papers for publication should be submitted online to the IJoSI website (http://www.ijosi.org) In order to preserve the anonymity of authorship, authors shall prepare two files (in MS Word format or PDF) for each submission. The first file is the electronic copy of the paper without author's (authors') name(s) and affiliation(s). The second file contains the author's (authors') name(s), affiliation(s), and email address(es) on a single page. Since the Journal is blind refereed, authors should not include any reference to themselves, their affiliations or their sponsorships in the body of the paper or on figures and computer outputs. Credits and acknowledgement can be given in the final accepted version of the paper.

Editorial policy

Submission of a paper implies that it has neither been published previously nor submitted for publication elsewhere. After the paper has been accepted, the corresponding author will be responsible for page formatting, page proof and signing off for printing on behalf of other co-authors. The corresponding author will receive one hardcopy issue in which the paper is published free of charge.

Manuscript preparation

The following points should be observed when preparing a manuscript besides being consistent in style, spelling, and the use of abbreviations. Authors are encouraged to download manuscript template from the IJoSI website, http://www.ijosi.org.

1. *Language.* Paper should be written in English except in some special issues where Chinese may be acceptable. Each paper should contain an abstract not exceeding 200 words. In addition, three to five keywords should be provided.
2. *Manuscripts.* Paper should be typed, single-column, double-spaced, on standard white paper margins: top = 25mm, bottom = 30mm, side = 20mm. (The format of the final paper prints will have the similar format except that double-column and single space will be used.)
3. *Title and Author.* The title should be concise, informative, and it should appear on top of the first page of the paper in capital letters. Author information should not appear on the title page; it should be provided on a separate information sheet that contains the title, the author’s (authors') name(s), affiliation(s), e-mail address(es).
4. *Headings.* Section headings as well as headings for subsections should start front the left-hand margin.
5. *Mathematical Expressions.* All mathematical expressions should be typed using Equation Editor of MS Word. Numbers in parenthesis shall be provided for equations or other mathematical expressions that are referred to in the paper and be aligned to the right margin of the page.
6. *Tables and Figures.* Once a paper is accepted, the corresponding author should promptly supply original copies of all drawings and/or tables. They must be clear for printing. All should come with proper numbering, titles, and descriptive captions. Figure (or table) numbering and its subsequent caption must be below the figure (or table) itself and as typed as the text.
7. *References.* Display only those references cited in the text. References should be listed and sequenced alphabetically by the surname of the first author at the end of the paper. For example:

Altshuller, G. (1998). *40 Principles: TRIZ Keys to Technical Innovation*, Technical Innovation Center.
Sheu, D. & Lee, H. (2011). A Proposed Process for Systematic Innovation, International Journal of Production Research, Vol. 49, No. 3, 2011, 847-868.

The International Journal of Systematic Innovation
Journal Order Form

Organization Or Individual Name	
Postal address for delivery	
Person to contact	Name: e-mail: Position: School/Company:
Order Information	**I would like to order ___ copy(ies) of the** *International Journal of Systematic Innovation*: **Period Start: 1st/ 2nd half ____, Year:____(Starting 2010)** **Period End : 1st/ 2nd half ____, Year:** **Price:** **Institutions: US $150 (yearly) / NT 4,500 (In Taiwan only)** **Individuals: US $50 (yearly) / NT 1500 (In Taiwan only)** (Local postage included. International postage extra) **E-mail to**: IJoSI@systematic-innovation.org or fax: +886-3-572-3210 Air mail desired □ (If checked, we will quote the additional cost for your consent)
Total amount due	**US$**
Payment Methods: 1. **Credit Card (Fill up the following information and e-mail/ facsimile this form to The Journal office indicated below)** 2. **Bank transfer** 3. **Account:** The Society of Systematic Innovation 4. **Bank Name:** Mega International Commercial BANK 5. **Account No:** 020-53-144-930 6. **SWIFT Code:** ICBCTWTP020 7. **Bank code**：017-0206 8. **Bank Address:** No. 1, Xin'an Rd., East Dist., Hsinchu City 300, Taiwan (R.O.C.)	

VISA / Master/ JCB/ AMERICAN Cardholder Authorization for Journal Order

Card Holder Information

Card Holder Name	(as it appears on card)		
Full Name (Last, First Middle)			
Expiration Date	/ (month / year)	Card Type	□ VISA □ MASTER □ JCB
Card Number	□□□□-□□□□-□□□□-□□□□	Security Code	□□□
Amount Authorized		Special Messages	
Full Address (Incl. Street, City, State, Country and Postal code)			

Please Sign your name here __________________ (same as the signature on your card)

The Society of Systematic Innovation
6 F, #352, Sec. 2, Guanfu Rd,
Hsinchu, Taiwan, 30071, R.O.C.