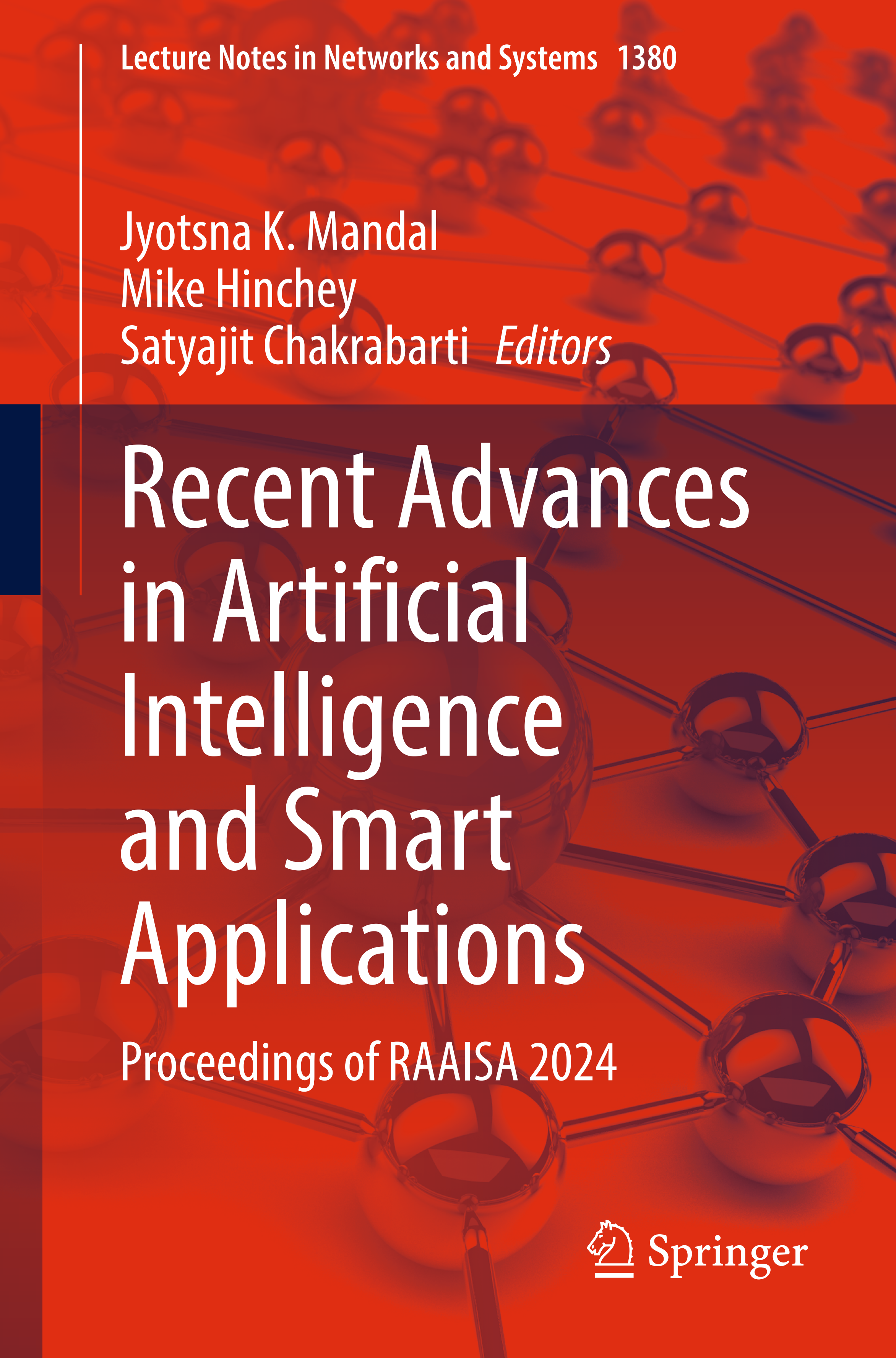


Lecture Notes in Networks and Systems 1380

Jyotsna K. Mandal
Mike Hinchey
Satyajit Chakrabarti *Editors*

Recent Advances in Artificial Intelligence and Smart Applications

Proceedings of RAAISA 2024

Springer

Lecture Notes in Networks and Systems

Volume 1380

The series "Lecture Notes in Networks and Systems" publishes the latest developments in Networks and Systems—quickly, informally and with high quality. Original research reported in proceedings and post-proceedings represents the core of LNNS.

Volumes published in LNNS embrace all aspects and subfields of, as well as new challenges in, Networks and Systems.

The series contains proceedings and edited volumes in systems and networks, spanning the areas of Cyber-Physical Systems, Autonomous Systems, Sensor Networks, Control Systems, Energy Systems, Automotive Systems, Biological Systems, Vehicular Networking and Connected Vehicles, Aerospace Systems, Automation, Manufacturing, Smart Grids, Nonlinear Systems, Power Systems, Robotics, Social Systems, Economic Systems and other. Of particular value to both the contributors and the readership are the short publication timeframe and the world-wide distribution and exposure which enable both a wide and rapid dissemination of research output.

The series covers the theory, applications, and perspectives on the state of the art and future developments relevant to systems and networks, decision making, control, complex processes and related areas, as embedded in the fields of interdisciplinary and applied sciences, engineering, computer science, physics, economics, social, and life sciences, as well as the paradigms and methodologies behind them.

Indexed by SCOPUS, EI Compendex, INSPEC, WTI Frankfurt eG, zbMATH, SCImago.

All books published in the series are submitted for consideration in Web of Science.

For proposals from Asia please contact Aninda Bose (aninda.bose@springer.com).

Jyotsna K. Mandal · Mike Hinchey ·
Satyajit Chakrabarti
Editors

Recent Advances in Artificial Intelligence and Smart Applications

Proceedings of RAAISA 2024

Springer

Editors
Jyotsna K. Mandal
Department of Computer Science
and Engineering
University of Kalyani
Kalyani, West Bengal, India

Mike Hinchey
Department of Computer Science
and Information Systems
University of Limerick
Limerick, Ireland

Satyajit Chakrabarti
University of Engineering and Management
Kolkata, West Bengal, India

ISSN 2367-3370 ISSN 2367-3389 (electronic)
Lecture Notes in Networks and Systems
ISBN 978-981-96-5821-3 ISBN 978-981-96-5822-0 (eBook)
https://doi.org/10.1007/978-981-96-5822-0

© The Editor(s) (if applicable) and The Author(s), under exclusive license to Springer Nature Singapore Pte Ltd. 2025

This work is subject to copyright. All rights are solely and exclusively licensed by the Publisher, whether the whole or part of the material is concerned, specifically the rights of translation, reprinting, reuse of illustrations, recitation, broadcasting, reproduction on microfilms or in any other physical way, and transmission or information storage and retrieval, electronic adaptation, computer software, or by similar or dissimilar methodology now known or hereafter developed.
The use of general descriptive names, registered names, trademarks, service marks, etc. in this publication does not imply, even in the absence of a specific statement, that such names are exempt from the relevant protective laws and regulations and therefore free for general use.
The publisher, the authors and the editors are safe to assume that the advice and information in this book are believed to be true and accurate at the date of publication. Neither the publisher nor the authors or the editors give a warranty, expressed or implied, with respect to the material contained herein or for any errors or omissions that may have been made. The publisher remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

This Springer imprint is published by the registered company Springer Nature Singapore Pte Ltd.
The registered company address is: 152 Beach Road, #21-01/04 Gateway East, Singapore 189721, Singapore

If disposing of this product, please recycle the paper.

Preface

The book is a collection of best-selected research papers presented at the 2nd International Conference "Recent Advances in Artificial Intelligence and Smart Applications (RAAISA) 2024." The book explores fresh concepts, practical applications, and insights shared by research engineers, scientists, industrialists, scholars, and students worldwide. It covers research contributions from Artificial Intelligence and Smart Applications which are from the areas such as Smart Agent-Based Systems, Human-Computer Interaction Technologies, Reinforcement Learning, Sentiment Analysis, Recurrent Neural Networks and its Applications, Genetic Algorithm Neural Networks, Deep Learning-Based Smart Systems Security, Privacy, and Trust issues in Smart Computing, Smart Computing using generative AI, Augmented Reality/Virtual Reality, AI/ML, and IoT.

The 2nd International Conference "Recent Advances in Artificial Intelligence and Smart Applications (RAAISA) 2024" aimed at making the academic sector visualize and understand the industry and its need, by acting as a bridge between the two sectors. The organizers successfully conducted the first conference RAAISA 2023 with 17% acceptance rate and published in the Springer book series *Innovations in Sustainable Technologies and Computing*.

There where tutorials, Keynote addresses, invited, talks, and general speeches of interest, talks from industries, paper, presentations, and panel discussions in multiple tracks covering diverse topic of interest. The proceedings of selected and presented papers to be published in LNNS series of Springer nature by the Springer at free of cost. We are thankful to the authority of springer nature for their active association with this conference.

The papers are checked for similarities multiple times through authentic, followed by updation by the authors and draw double blinded refuse. Thirty-eight quality papers are uploaded after presentation of the same by the authors.

The editors express their gratitude to the reviewers of the articles for their enormous AirPods. Dhar also like to thanks the authors for submitting their papers in this conference, the editors also express is there gratitude to the organisers of this conference.

We hope that this volume will be useful and contributes actively towards the holistic progress of the civilisation. This volume will be value added bits and pieces for academicians, researchers, and young promising engineers.

Kalyani, India — Dr. Jyotsna K. Mandal
Limerick, Ireland — Dr. Mike Hinchey
Kolkata, India — Dr. Satyajit Chakrabarti

Contents

Application of Chaotic Map and DNA Encoding for Securing Stereo Audio Data

Mousomi Roy and Kalyani Mali

Abstract The importance of audio data in different areas such as telecommunications, entertainment, and confidential information shows the need for advanced encryption methods. In the ever-evolving information security environment, the need for strong and innovative encryption technologies is critical. Traditional encryption methods have long been the cornerstone of digital data security. However, as technology advances, so do the challenges of sophisticated attackers. The combination of chaos and DNA encoding not only hardens audio data against potential breaches but also ushers in a new dimension of secure communication. This study applies chaos theory and DNA encoding to audio encryption to demonstrate the importance of protecting audio data. In an era of growing concerns about privacy and data integrity, this innovative approach is a powerful solution for protecting audio information. Using chaos theory, it is possible to incorporate deterministic yet unpredictable dynamics to embed audio information in a chaotic system, which would be difficult to decode by traditional methods. To improve this strategy, DNA encoding is introduced to convert audio data into genetic code, using the information density and flexibility of DNA. Converting audio data to DNA code adds a layer of security and data density. The robustness of genetic material and the complex dance of chaos create a powerful encryption strategy that resists traditional cryptographic attacks.

Keywords Audio security · Chaos · DNA encoding

M. Roy (✉)
Department of Computer Science and Engineering, Hooghly Engineering & Technology College, Hugli-Chuchura, India
e-mail: iammouroy@gmail.com

K. Mali
Department of Computer Science and Engineering, University of Kalyani, Kalyani, India

J. K. Mandal et al. (eds.), *Recent Advances in Artificial Intelligence and Smart Applications*, Lecture Notes in Networks and Systems 1380,
https://doi.org/10.1007/978-981-96-5822-0_1

1 Introduction

In a world where communication is transcendently advanced and the transmission of touchy data may be a scheduled undertaking, the requirement for impenetrable encryption strategies has never been more basic. With the rise of quantum computing and the consistent advancement of cyber dangers, conventional encryption methods are progressively helpless to compromise. As a reaction to these challenges, this paper proposes a cutting-edge sound encryption approach that leverages chaos theory and DNA encoding, offering a novel arrangement that guarantees both security and efficiency. Chaos theory starting from the domain of science and material science, has found its way into various disciplines and building applications. Within the setting of data security, chaos theory presents a component of unusualness and haphazardness that can upgrade the strength of encryption calculations. The sensitivity to the initial conditions and the non-linear flow of chaotic frameworks [4, 10, 11] give ground for creating cryptographic techniques for securing audio data. Drawing motivation from life, DNA encoding offers an interesting point of view on information capacity and recovery. The four nucleotide bases—adenine (A), thymine (T), cytosine (C), and guanine (G)—serve as nature's letter sets, encoding the diagram of life in a surprisingly thick and versatile way. The versatility and storage capacity of DNA encoding make it a captivating candidate for encoding and putting away sensitive audio information [20, 27].

In this encryption approach, the audio information is changed into a computerized representation of DNA arrangements [1, 10, 21]. Each nucleotide compares to a particular sound fragment, making a biological analog for the advanced sound flag. The natural nature of DNA includes an extra layer of complexity and security to the encoded data, as translating requires not as it were computational control but moreover a profound understanding of atomic biology. The collaboration between the chaos theory and DNA encoding shapes the foundation of our proposed audio encryption framework [17, 22, 24]. The interweaving of these two components makes the proposed encryption strategy makes it secure, robust, and resilient against various attacks. In this integrated framework, chaos makes the encoding and decoding process unpredictable, guaranteeing that indeed the foremost modern attacks can be resisted. The chaos theory and DNA encoding impose obstruction against various cryptographic attacks, showing a special challenge for intruders to breach the system. Furthermore, the chaotic nature of the proposed framework presents a level that improves the security of the encryption plot. The affectability to beginning conditions guarantees that indeed a minor annoyance within the chaotic framework comes about in an unfathomably diverse cryptographic key, obstructing endeavors at switch engineering. A comprehensive investigation is required to understand the performance of any encryption framework. Hence, the proposed chaos and DNA-based audio encryption system undergo a thorough investigation to assess its resistance against various potential attacks [6, 23, 27].

While the integration of chaos and DNA encoding in audio encryption presents a promising security framework, the computational overhead of real-time chaotic framework and DNA encoding may pose challenges for resource-constrained situations. Moreover, the organic nature of DNA encoding presents contemplations related to capacity, recovery, and the potential effect of natural factors [3, 8, 26].

The rest of the article is organized as follows: Section 2 discusses some foundational concepts related to the proposed approach, Sect. 3 illustrates the proposed approach in detail, Sect. 4 discusses the obtained results, and Sect. 5 concludes the article.

2 Foundation

2.1 DNA Encoding Technique

As the demand for secure audio communication emerges, the application of DNA encoding standards to audio security can be helpful for more reliable communication of audio data. This intersection of digital technology and audio security exemplifies the adaptability of foundational concepts, demonstrating that the principles of encoding can be applied creatively to safeguard information in diverse contexts. In the context of audio, DNA encoding refers to the unique and intricate process of representing sound waves in a digital format, preserving the essential characteristics that make each audio snippet distinct [11].

DNA consists of four nucleotide bases (adenine, thymine, cytosine, and guanine), which can be encoded into binary code. In the natural DNA sequence, the bases are often represented by the letters A, T, C, and G. To convert DNA encoding to binary, any one simple mapping can be established as given in Table 1 [2, 12]. This binary encoding serves as a digital representation of the genetic information stored in the DNA sequence. Table 2 demonstrates the DNA XOR operation.

Table 1 DNA encoding rules

Rule No. →	0		1		2		3		4		5		6		7	
	A	00	A	00	C	00	C	00	G	00	G	00	T	00	T	00
	C	01	G	01	A	01	T	01	A	01	T	01	C	01	G	01
	G	10	C	10	T	10	A	10	T	10	A	10	G	10	C	10
	T	11	T	11	G	11	G	11	C	11	C	11	A	11	A	11

Table 2 DNA XOR operation

$\oplus$	A	T	G	C
A	A	T	G	C
T	T	A	C	G
G	G	C	A	T
C	C	G	T	A

2.2 The RCM Chaotic Map

Chaos theory helps to incorporate unpredictability and randomness into security systems or processes as a strategy to enhance resilience and increase overall robustness. In this work, the RCM chaotic map is used and this map is defined in Eq. 1 [11].

$$s_i = \psi \sin\left(e^{-s_{i-1}} \times \left(1 - s_{i-1}^3\right)\right) \tag{1}$$

This chaotic property of the RCM map is already discussed and proven in [11].

3 Proposed Audio Encryption Framework

In this work, stereo audio data (i.e., with two channels) are used for the encryption purposes. Stereo data has two channels with double values. The proposed approach begins by converting each value into the IEEE 754 double-precision binary representation. Therefore, for each member of a particular sequence, there are corresponding 64 bits are generated [5] as shown in Fig. 1. Where, the first bit or the MSB is the sign bit, the next 11 bits are the exponent, and the rest of the bits, i.e., last 52 bits are mantissa. These 64 bits are divided into 8 groups.

Now, a pseudorandom bit sequence of length 16 is generated by the RCM map and the pseudorandom bit sequence generation procedure as described in [18] and this sequence is further divided into four groups of 4 bits each and it is illustrated in Fig. 2. After dividing into four groups, each group is reversed individually.

Now the third bit of the first and the third subgroups are discarded (illustrated in Fig. 3) and XOR operation is performed between the remaining three bits. These three bits are used to determine one DNA encoding rule from Table 1. The second

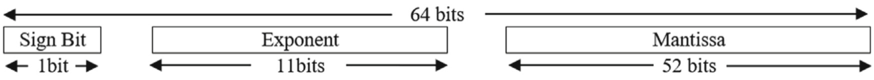


Fig. 1 IEEE 754 double-precision binary representation

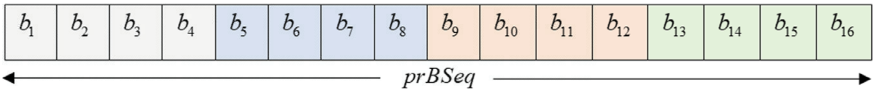
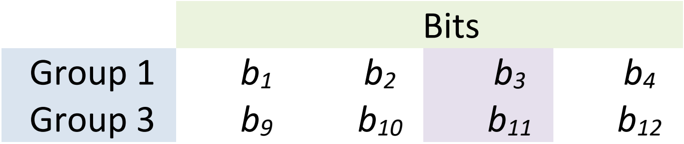


Fig. 2 Bitwise grouping of the generated pseudorandom bit sequence



	Bits			
Group 1	b_1	b_2	b_3	b_4
Group 3	b_9	b_{10}	b_{11}	b_{12}

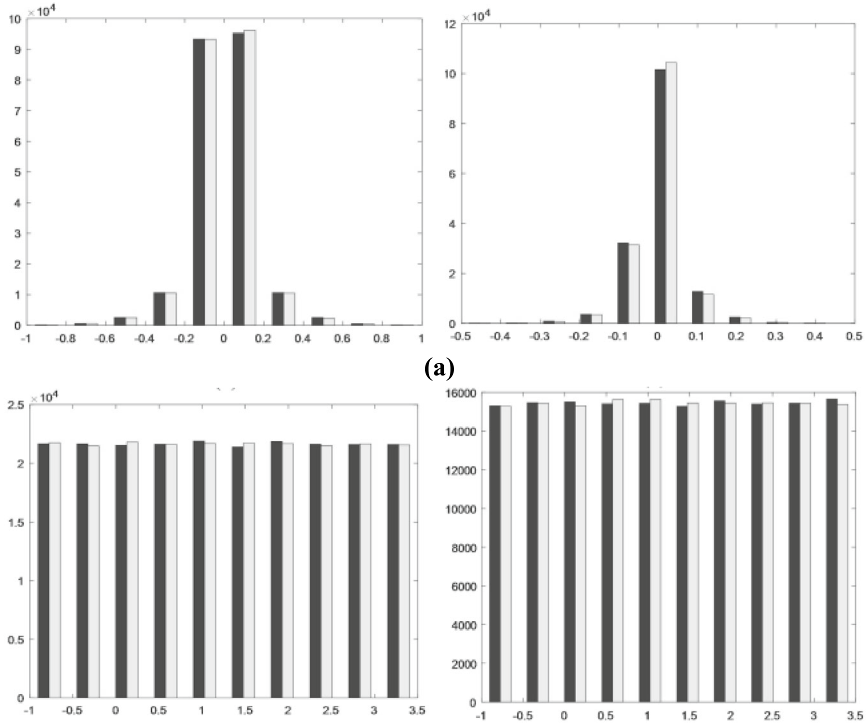
Fig. 3 Bit discarding procedure (highlighted bits are discarded)

and the fourth subgroups are combined in reverse order and generated 8 bits. Now the selected rule is applied to convert the binary values from the IEEE 754 double-precision binary representation and the merged eight bits into the DNA encoded form. The DNA XOR operation is performed between these two sequences. It is worth mentioning in this context that the process will be repeated 8 times for a certain element of the audio data (because 64/8 = 8). Moreover, the total process will be executed for each element.

A pseudorandom bit sequence $prCSeq = \{p_0, p_1, ..., p_d\}$ of length $d = 2 \cdot n$ (where n is the number of elements present in each channel) is generated using the RCM map and the pseudorandom bit sequence generation procedure as described in [18]. This pseudorandom bit sequence is sorted in descending order and a sorted sequence is prepared as $prCSeq' = \left\{p'_0, p'_1, ..., p'_d\right\}$. The two channels of the DNA XORed audio signal are merged and converted into a one-dimensional vector $v = \{v_1, v_2, v_3, ..., v_d\}$. This vector is used to construct another vector $v' = \left\{v'_1, v'_2, ..., v'_d\right\}$ using $v'_i = v_i$ where $prCSeq'_i = prCSeq_i$. This method is incorporated to scramble the sequence. Two channels are constructed from the resultant scrambled one-dimensional vector and the final encrypted audio is returned.

4 Results of the Simulation

The proposed approach is evaluated using some two-channel audio inputs. All experiments are carried out in MATLAB R2022a environment with a computer equipped with 8GB RAM, Intel i5 Processor, and 1TB HDD.



(a)

(b)

Fig. 4 Histogram analysis. **a** Original histograms of two audio signals with two channels. **b** corresponding encrypted histograms

4.1 Analysis of the Histogram

Histogram analysis is a technique commonly used in signal processing to analyze the distribution of signal amplitudes. For a good encryption technique, the uniform distribution of the histogram is essential. The histogram analysis is presented in Fig. 4. In this figure, the original histograms of two audio signals with two channels are presented. In the second row, the corresponding encrypted histograms are reported. Obtained histograms of the encrypted audio signals are nearly uniform.

4.2 Analysis of the Correlation Coefficients

The correlation coefficient is a statistical measure that quantifies the degree to which two variables are linearly related. An unencrypted meaningful audio signal typically has a high correlation because one element of a channel is almost nearly similar to its neighbors. Typically, correlation coefficients can take any value from the interval $[-1, +1]$, with maximum or minimum values being generated in the worst scenario, i.e., a completely negative or positive linear relationship [7, 13, 14]. A zero value of the correlation coefficient is the ideal situation and indicates no linear relationship between adjacent elements. The value of the correlation coefficient can be computed using Eq. 2 [15].

$$CC_{i,j} = \frac{CovF(i,j)}{\sqrt{\mathrm{var}(i) \cdot \mathrm{var}(j)}} \tag{2}$$

$CovF(i, j)$ is the covariance function between two samples i and j. $\mathrm{var}(x)$ is the function that determines the variation of a sample x. $CovF()$ and var() are defined in Eqs. 3 and 4 respectively.

$$CovF(i,j) = \frac{\sum_{z=1}^{P}(i_z - E(i))(j_z - E(j))}{P} \tag{3}$$

$$\mathrm{var}(x) = \frac{1}{P}\sum_{z=1}^{P}(x_z - E(x))^2 \tag{4}$$

Here, $E(x)$ is the mean, and it is defined in Eq. 5. P denotes the count of the paired data.

$$E(x) = \frac{1}{P}\sum_{z=1}^{P} x_z \tag{5}$$

The outcome of the analysis of the correlation coefficient is reported in Fig. 5.

The value of the correlation coefficient of the left and the right channel of both original and encrypted signals are reported in Table 3.

4.3 Analysis of the Differential Attack

The attackers and cryptanalysts try to determine the relation between the original signal and the encrypted signal. They change some portion of the original signal and observe the encrypted output. It is essential for any encryption approach to resist differential attacks successfully. To understand the strength of encryption, the Number of Sample Change Rates (NSCR) and Unified Average Changing Intensity (UACI) values are often investigated. These two parameters are defined in Eqs. 6 and 7 respectively [8, 25].

$$NSCR = 100 \times \frac{\sum_{m,n} D(m,n)}{L} \tag{6}$$

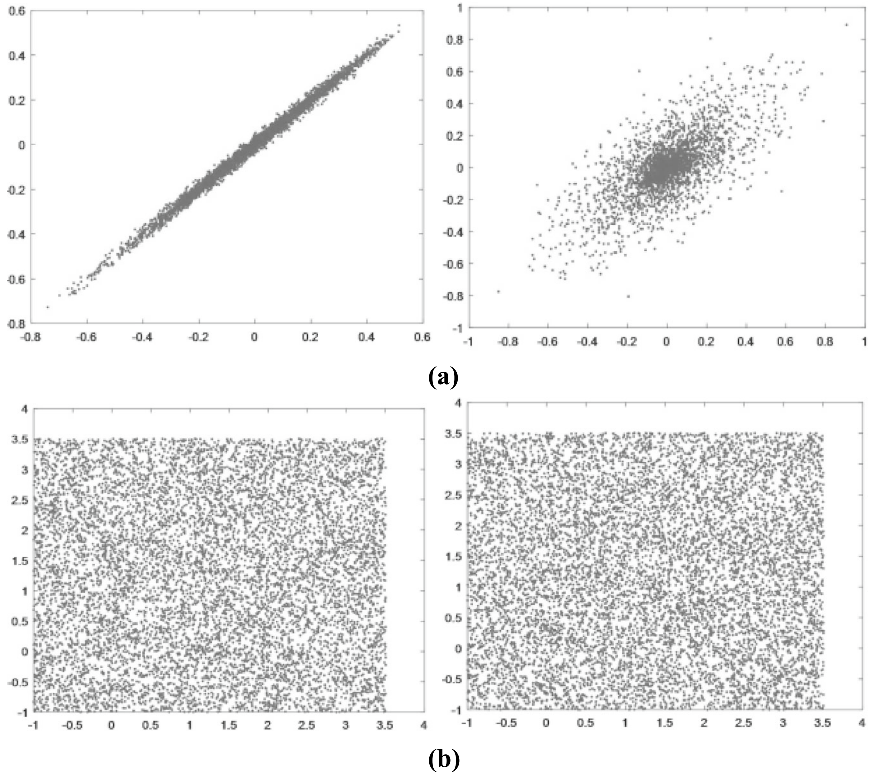


Fig. 5 Analysis of the correlation coefficient. **a** Correlation coefficient of the original channels. **b** Correlation coefficient of the corresponding encrypted channels

Table 3 Analysis of the correlation coefficient

Audio sample	Original signal ($L_{original}/R_{original}$)	Encrypted signal ($L_{encrypted}/R_{encrypted}$)
Sample 1	0.9537	0.0003
Sample 2	0.8829	0.0059
Sample 3	0.9091	0.0002
Sample 4	0.9962	0.0007
Sample 5	0.9718	0.0071

$$UACI = \frac{1}{L}\sum_{m,n}\frac{|c_1(m,n) - c_2(m,n)|}{2^b - 1} \times 100 \tag{7}$$

Here, L is the length of the original audio signal and b is the bit count required to represent the audio signal. c_1 and c_2 are the two cipher signals that are produced by

Table 4 Analysis of the PSNR and UACI values

Audio sample	NSCR (%)	UACI (%)
Sample 1	98.39	33.3966
Sample 2	99.26	33.6833
Sample 3	99.21	33.8077
Sample 4	98.99	33.5077
Sample 5	99.17	33.9007

changing a single value in the input audio signal. If $c_1(m, n) = c_2(m, n)$ then $D(m, n) = 0$. Table 4 reports the NSCR and UACI values.

5 Conclusion

The proposed audio encryption approach is tested rigorously. Experimental outcomes show that the proposed approach can be effective enough to be deployed in various real-life scenarios. The proposed approach can successfully defend against various cryptographic attacks. The proposed approach can be further extended by applying various chaotic maps. Moreover, the proposed approach can also be extended with the help of the hyperdimensional chaos theory. The strength of the proposed approach can be further extended with the help of quantum cryptographic techniques. It is possible to combine multiple chaotic maps to improve randomness and avoid predictability. Moreover, it is possible to introduce multi-dimensional chaotic maps for increased complexity. It is possible to Opt for low-complexity maps like the Logistic Map or Piecewise Linear Chaotic Map (PWLCM) which require fewer computational resources compared to higher-dimensional maps (e.g., Lorenz, Henon). The use of fixed-point arithmetic instead of floating-point can reduce the computational overhead and memory usage while maintaining sufficient precision. The load of the proposed system can be further reduced by Pre-computation of Chaotic Sequences. However, the encrypted audio signal produced by the proposed approach is secured enough the proposed approach can be deployed in real-life scenarios reliably.

References

1. Akkasaligar, P. T., & Biradar, S. (2020). Selective medical image encryption using DNA cryptography. *Information Security Journal, 29*(2), 91–101. Taylor and Francis Inc. https://doi.org/10.1080/19393555.2020.1718248
2. Amin, S. T., Saeb, M., & El-Gindi, S. (2006). A DNA-based Implementation of YAEA encryption algorithm. *Computational Intelligence*. http://www.gutenberg.org/
3. Chakraborty, S., Seal, A., Roy, M., & Mali, K. (2016). A novel lossless image encryption method using DNA substitution and chaotic logistic map. *International Journal of Security and Its Applications*, *10*(2), 205–216. https://doi.org/10.14257/ijsia.2016.10.2.19

4. Farajallah, M., El Assad, S., & Deforges, O. (2016). Fast and secure chaos-based cryptosystem for images, *26*(2). https://doi.org/10.1142/S0218127416500218
5. Havaldar, S., & Gurumurthy, K. S. (2017). Design of vedic IEEE 754 floating point multiplier. In *2016 IEEE International Conference on Recent Trends in Electronics, Information and Communication Technology, RTEICT 2016—Proceedings* (pp. 1131–1135). https://doi.org/10.1109/RTEICT.2016.7808008
6. Hu, J., & Han, F. (2009). A pixel-based scrambling scheme for digital medical images protection. *Journal of Network and Computer Applications, 32*(4), 788–794. https://doi.org/10.1016/J.JNCA.2009.02.009
7. Indrasena Reddy, M., Siva Kumar, A. P., & Subba Reddy, K. (2020). A secured cryptographic system based on DNA and a hybrid key generation approach. *Bio Systems, 197*, Article 104207. https://doi.org/10.1016/j.biosystems.2020.104207
8. Mali, K., Chakraborty, S., & Roy, M. (2015). A study on statistical analysis and security evaluation parameters in image encryption. *IJSRD-International Journal for Scientific Research & Development*, *3*, 2321–0613. www.ijsrd.com
9. Mali, K., Chakraborty, S., Seal, A., & Roy, M. (2015). An efficient image cryptographic algorithm based on frequency domain using haar wavelet transform. *International Journal of Security and Its Applications*, *9*(12), 279–288. https://doi.org/10.14257/ijsia.2015.9.12.26
10. Roy, M., Chakraborty, S., & Mali, K. (2020). A robust image encryption method using chaotic skew-tent map. In S. Chakraborty & K. Mali (Eds.), *Applications of advanced machine intelligence in computer vision and object recognition: emerging research and opportunities*. https://doi.org/10.4018/978-1-7998-2736-8.ch001
11. Roy, M., Chakraborty, S., & Mali, K. (2021). A chaotic framework and its application in image encryption. *Multimedia Tools and Applications*, 1–42. https://doi.org/10.1007/s11042-021-10839-7
12. Roy, M., Chakraborty, S., & Mali, K. (2021). The MSK: a simple and robust image encryption method. *Multimedia Tools and Applications*, 1–31. https://doi.org/10.1007/s11042-021-10761-y
13. Roy, M., Chakraborty, S., & Mali, K. (2023). An evolutionary image encryption system with chaos theory and DNA encoding. *Multimedia Tools and Applications, 82*(22), 33607–33635. https://doi.org/10.1007/S11042-023-14948-3/METRICS
14. Roy, M., Chakraborty, S., & Mali, K. (2023). Metaheuristic-supported image encryption framework based on binary search tree and DNA encoding. *Multimedia Tools and Applications*, 1–29. https://doi.org/10.1007/S11042-023-16471-X/METRICS
15. Roy, M., Chakraborty, S., & Mali, K. (2024). Audio encryption framework based on chaotic map and DNA encoding. *Applied Acoustics, 224*, Article 110152. https://doi.org/10.1016/J.APACOUST.2024.110152
16. Roy, M., Chakraborty, S., Mali, K., Banerjee, A., Ghosh, K., & Chatterjee, S. (2020). Biomedical image security using matrix manipulation and DNA encryption. *Advances in Intelligent Systems and Computing, 1065*, 49–60. https://doi.org/10.1007/978-981-15-0361-0_4
17. Roy, M., Chakraborty, S., Mali, K., Mitra, S., Mondal, I., Dawn, R., Das, D., & Chatterjee, S. (2019). A dual layer image encryption using polymerase chain reaction amplification and dna encryption. In *2019 International Conference on Opto-Electronics and Applied Optics, Optronix 2019*. https://doi.org/10.1109/OPTRONIX.2019.8862350
18. Roy, M., Chakraborty, S., Mali, K., & Roy, D. (2021). *Utilization of Hyperchaotic Environment and DNA Sequences for Digital Image Security* (pp. 289–297). Springer, Singapore. https://doi.org/10.1007/978-981-15-9433-5_28
19. Roy, M., Chakraborty, S., Mali, K., Roy, D., & Chatterjee, S. (2021). A robust image encryption framework based on DNA computing and chaotic environment. *Microsystem Technologies*, 1–11. https://doi.org/10.1007/s00542-020-05120-0
20. Roy, M., Chakraborty, S., Mali, K., Roy, D., & Chatterjee, S. (2021). *An image security method based on low dimensional chaotic environment and DNA encoding* (pp. 267–277). Springer, Singapore. https://doi.org/10.1007/978-981-15-9433-5_26

21. Roy, M., Chakraborty, S., Mali, K., Swarnakar, R., Ghosh, K., Banerjee, A., & Chatterjee, S. (2020). Data security techniques based on DNA encryption. In *Advances in intelligent systems and computing* (Vol. 1065). https://doi.org/10.1007/978-981-15-0361-0_19
22. Roy, M., Mali, K., Chatterjee, S., Chakraborty, S., Debnath, R., & Sen, S. (2019). A study on the applications of the biomedical image encryption methods for secured computer aided diagnostics. In *Proceedings—2019 Amity International Conference on Artificial Intelligence, AICAI 2019*. https://doi.org/10.1109/AICAI.2019.8701382
23. Seal, A., Chakraborty, S., & Mali, K. (2017). A new and resilient image encryption technique based on pixel manipulation, value transformation and visual transformation utilizing single–Level haar wavelet transform. In *Advances in intelligent systems and computing* (Vol. 458). https://doi.org/10.1007/978-981-10-2035-3_61
24. Tedmori, S., & Al-Najdawi, N. (2014). Image cryptographic algorithm based on the Haar wavelet transform. *Information Sciences, 269*, 21–34. https://doi.org/10.1016/j.ins.2014.02.004
25. Wang, J., Yu, W., Wang, J., Zhao, Y., Zhang, J., & Jiang, D. (2019). A new six-dimensional hyperchaotic system and its secure communication circuit implementation. *International Journal of Circuit Theory and Applications, 47*(5), 702–717. https://doi.org/10.1002/CTA.2617
26. Zare, A., Aghagolzadeh, A., & Kazemitabar, S. J. (2019). A novel metaheuristic based visual cryptography. *Journal of Advanced Defense Science and Technology*, *10*(3), 297–306. https://adst.ihu.ac.ir/article_204333.html
27. Zhang, Y., Zhou, J., Chen, F., Zhang, L. Y., Wong, K. W., He, X., & Xiao, D. (2016). Embedding cryptographic features in compressive sensing. *Neurocomputing, 205*, 472–480. https://doi.org/10.1016/J.NEUCOM.2016.04.053

Disease Classification of Star Fruit (Averrhoa carambola L.) Using Deep Learning

Samanawaya Datta, Sakyojit Banerjee, and Susovan Jana

Abstract This research presents an automated system for carambola disease diagnosis using Convolutional Neural Networks (CNNs), which contains two modules, i.e., Carambola Identification and Disease Classification. It accurately identifies carambola objects and categorizes diseases through extensive real-world dataset experiments. It also presents early detection and enhanced crop management. This system promises to transform agricultural diagnostics. Through the research, the developed Carambola Identification Model achieved an impressive F1 score of 1, while the Disease Classification Model achieved an F1 score of 0.98. This system holds significant potential for sustainable carambola cultivation practices, benefiting farmers, researchers, and policymakers alike. This innovation represents a significant step toward optimizing agricultural productivity and ensuring food security in carambola cultivation regions. Its scalability and adaptability suit broader agricultural applications.

Keywords Agricultural image analysis · CNN · Disease diagnosis · Plant pathology · Carambola

1 Introduction

One of the most popular subtropical fruits in Southeast Asian and Eastern countries is the Starfruit, also known as Carambola. Low in calories and high in vitamin C, it is also used in various dishes like jams, jellies, salsas, and more. However, it is

---

S. Datta · S. Banerjee
Department of Computer Science and Engineering (IoTCSBT), Institute of Engineering & Management, Kolkata, India

S. Jana (✉)
Department of Computer Science and Engineering (IoT), Institute of Engineering & Management, University of Engineering & Management, Kolkata, India
e-mail: jana.susovan2@gmail.com

J. K. Mandal et al. (eds.), *Recent Advances in Artificial Intelligence and Smart Applications*, Lecture Notes in Networks and Systems 1380,
https://doi.org/10.1007/978-981-96-5822-0_2

very susceptible to several diseases that can lead to problems such as low output and yield, as well as rotten and low-quality fruit.

The traditional methods of disease detection tend to be extremely laborious and prone to misinputs and errors. However, with advancements in recent years in the paradigms of machine learning, deep learning, and Convolutional Neural Networks (CNNs), promising solutions have emerged that address most of the problems associated with the agriculture of this fruit. CNNs are utilized in various image recognition models that can learn to identify and perform extremely niche and complex patterns in images that may be oblivious to the naked eye. This paper aims to provide a way of preventing such problems.

This research paper is comprised of two models: the Carambola Identification Model and the Disease Classification Model. The Carambola Identification Model allows for the detection of whether the target image is a Carambola fruit or a Carambola leaf. Thus, this model allows for the identification of the entire Carambola plant as a whole dataset. This model forms the foundation for the subsequent Disease Classification Model. The Disease Classification Model is comprised of MobileNetV1, an extremely lightweight yet highly popular CNN architecture known for its robustness and efficiency. This model classifies the different forms of diseases associated with either the fruit or the leaf, utilizing various metrics like F1 scores, precision, and recall values, along with confusion matrices for accurate disease detection. Such a system will inherently improve agricultural diagnostics with early disease detection and improved crop management techniques, which will ultimately contribute to the broader agricultural management and development paradigm.

The next section describes the literature survey for fruit identification and disease detection. Section 3 contains a detailed description proposed model. Section 4 shows the experimentation results followed by the conclusion and future scope in Sect. 5.

2 Literature Study

With significant advancements in agricultural technology, various works utilize deep learning techniques, specifically the InceptionV3 model and transfer learning, to develop user-friendly tools for fruit disease detection and grading [1]. The k-means clustering algorithm is used for detecting defective portions of apple fruit [2], and methods for fruit disease detection include color, morphology, and texture features, with fuzzy c-means and k-means for segmentation [3].

Fruit images are preprocessed to separate the foreground from the background by extracting texture and statistical color features, and combining them into a single descriptor. Training is performed using a Support Vector Machine (SVM) with feature descriptors from a labeled dataset, aiming to accurately recognize carambola fruit or leaf objects, with the Disease Classification Model using CNN architecture to categorize diseases [4]. A model can also normalize fruit photographs by correcting translation, rotation, and scale issues, focusing on shape-based characteristics [5].

Comparative analysis of cutting-edge ML and DL algorithms for plant leaf diseases shows that a novel soybean leaf classification model using k-means segmentation achieves 62.53% accuracy, and Cotton plant leaf disease classification with FFBPNN achieves 95.48% accuracy [6]. A system can use CNN models for disease detection in different plant species and k-means clustering for image segmentation, aiming to identify carambola fruit or leaf objects accurately [7].

Traditional picture classification methods have limitations in generalization and accuracy. This study presents a unique system for classifying fruit images using transfer learning and deep learning methods, specifically the feature extraction and classification done by the modified Inception-V3 model. Transfer learning improves the training process [8]. The research highlights current improvements in disease detection and classification of plants using machine learning and deep learning, suggesting multi-class, multi-label deep learning algorithms for better accuracy and speed [9].

3 Proposed Model

In the above experiment, we have developed and evaluated two deep learning models: the Identification Model (Custom Model) and the Disease Classification Model (MobileNetV1 Base model). The details of the layers for both models are mentioned in Table 1. The Identification Model uses data augmentation, convolutional layers, max-pooling, dropout, Adam optimizer, categorical cross-entropy loss, and evaluation via confusion matrix and F1 score. The Disease Classification Model uses MobileNetV1 for feature extraction, custom fully connected layers with ReLU activation, dropout, Adam optimizer, categorical cross-entropy loss, and tracks training with accuracy/loss plots. Evaluation includes a confusion matrix and detailed classification report with precision, recall, and F1 scores. The base layers of both models have been frozen so as not to poison the model layers with pre-trained pre-built layers.

The primary reason for utilizing a custom model for the Identification model and MobileNetV1 model for the Disease Classification model boils down to time considerations for compilation and model creation, compared to accuracy. Due to only two classes being utilized in the Identification model, a simpler model compared to any other mainstream models fares much better in terms of overall performance. In the Disease Classification, MobileNet fares considerably better compared to any mainstream model. This reduces the need for developing custom models with a lower number of parameters with lower accuracy scores. Therefore, the identification can compile and create a model efficiently and fast considering the classes and the Disease Classification model can create an optimal model that can correctly identify the different types of diseases.

Table 1 Characteristics used for the models

Sl. no	Identification model	Disease classification model
1.	Conv2D Layer (32 filters, 3 × 3 kernel, ReLU): Detects low-level features with non-linearity	MobileNetV1 Base Model: Utilizes MobileNetV1 for feature extraction with frozen convolutional layers
2.	MaxPooling2D Layer (2 × 2): Reduces spatial dimensions and computational complexity	GlobalAveragePooling2D Layer: Averages each feature map, reducing dimensions for custom dense layers
3.	Conv2D Layer (64 filters, 3 × 3 kernel, ReLU): Extracts more complex features	Dense Layer (512 units, ReLU): Learns complex mappings between features and classes
4.	MaxPooling2D Layer (2 × 2): Further reduces spatial dimensions	Dropout Layer (0.25): Prevents overfitting by randomly dropping out 25% of neurons
5.	Conv2D Layer (128 filters, 3 × 3 kernel, ReLU): Captures intricate patterns	Dense Layer (256 units, ReLU): Further refines learned representations
6.	MaxPooling2D Layer (2 × 2): Down samples feature maps for fully connected layers	Dropout Layer (0.25): Mitigates overfitting similarly to the previous layer
7.	Flatten Layer: the feature maps are converted to one-dimensional vectors for next layer	Output Layer: Matches the number of disease classes, using softmax activation for probability distribution
8.	Dense Layer (64 units, ReLU, L2 regularization): Learns higher-level abstractions with regularization	
9.	Dropout Layer (0.25): Prevents overfitting	
10.	Dense Layer (32 units, ReLU, L2 regularization): Refines representations with regularization	
11.	Dropout Layer (0.25): Adds robustness by reducing redundancy	

4 Experimentation and Results

4.1 Dataset Description

The primary dataset utilized is Carambola Disease Recognition Dataset [10] from Mendeley Data with images also being scraped from Google Images and various other sources. After careful consideration, they have been utilized for testing and training. The images utilized are primarily been utilized from the Carambola Disease Recognition Dataset [10] and also have been acquired from various online sources. Dataset A consists of the identification images consisting of a combined total of 441 images from 2 classes resized into 299 × 299 for uniform size and faster processing.

models are trained on a variety of normal and fall scenarios to ensure accurate detection in real-world environments. Since this approach does not require users to wear any devices or have cameras installed, privacy preservation is high, and hence can be utilized in areas that privacy preservation should be highly maintained.

When a potential fall is detected by either the video analytics or CSI Wi-Fi-based system, notifications are sent to relevant caregivers and emergency responders. Additionally, the systems maintain data on fall incidents, daily activities, and mobility patterns in a secure data repository, allowing for review and analysis by stakeholders such as caregivers, healthcare professionals, and emergency teams. This ongoing data collection enables continuous enhancements in fall prevention and safety measures as well as activity recommendations to support the well-being of the monitored individuals.

3.3 System C: Telerehabilitation System for Post-Stroke Patients by Incorporating Bimanual Rehabilitation Robotic Arms

Figure 3c illustrates the developed bimanual rehabilitation system [7], designed to support post-stroke patients in motor skill recovery through customized arm exercises with robotic assistance. This system uses a master-follower robotic setup to aid the paretic arm in regaining functionality by guiding it with movements from the healthy arm, enhancing the rehabilitation process for two primary exercises: flexion & extension of elbow joint and internal & external rotation of shoulder joint. Both arms move simultaneously under bi-manual motor control, effectively simulating therapist-assisted movements. The master robotic arm supports the healthy arm, while the follower robotic arm aids the affected arm. Thus, the robot ensures position synchronization during rehabilitation exercises. Accurate force sensing is achieved by using a sensorless method called the Reaction Torque OBserver (RTOB) [7]. Each rehabilitation session is recorded, enabling further processing, analysis, and visualization of patient-specific exercises. This data allows physical therapists to monitor progress, adjust rehabilitation plans as needed, and optimize the overall benefits of the rehabilitation process.

3.4 System D: Analytics and Visualization System for Care Takers and Healthcare Professional Medical Decision Support

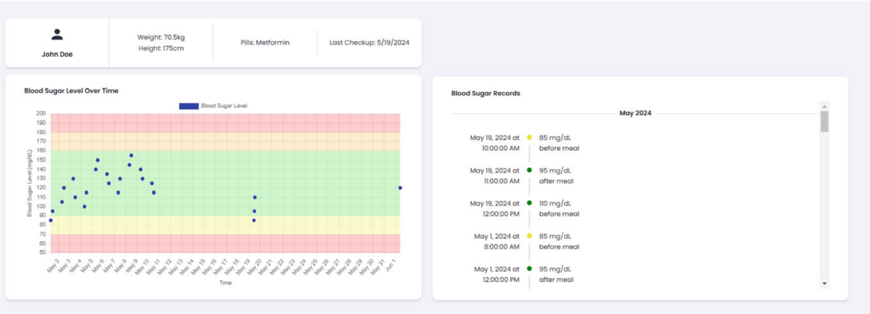
Effective monitoring and communication are essential in healthcare, particularly for patients undergoing rehabilitation or managing chronic conditions. Traditional methods, often reliant on manual data tracking and in-person consultations, can be time-consuming and may delay decision-making. As depicted by Fig. 3d, this analytics and visualization system addresses these challenges by aggregating data from other systems within the platform, including blood sugar levels, fall and activity patterns, and bimanual rehabilitation exercises, and presenting it in an easily interpretable format with charts, graphs, and visual indicators. These visual representations allow caregivers and healthcare professionals to assess individual progress, identify trends, and make informed decisions quickly. The dashboard consolidates all relevant data at a glance, supporting proactive care and efficient monitoring.

4 Prototype System Development and Demonstration

This section provides an overview of the developed prototype system of the proposed platform (System D) [11, 12], highlighting key visualizations and data representations used across various systems within the platform.

Figure 4 presents a few key *visualizations for blood glucose monitoring* (System A), displaying a diabetes patient profile and history of blood glucose measurements. The interface includes timestamps for blood glucose levels alongside meal logs, enabling patients and caregivers to observe how meals impact glucose fluctuations. A chart differentiates various glucose ranges—high, low, healthy, and unhealthy—making it straightforward to assess overall health trends.

Figure 5 demonstrates key *visualizations for activity mobility and fall analysis* (System B). Activity History (Fig. 5a) provides detailed records of activities by type location and time enabling long-term pattern analysis. The Daily Activity Timeline (Fig. 5b) offers a visual overview of activities throughout the day with filters for specific areas/camera and date range allowing customized monitoring. Fall Incident Records (Fig. 5c) focus on fall events with information on activity type location and time along with a 'View Fall' option for further investigation. The Fall Timeline (Fig. 5d) shows events before and after a fall to provide context. Together these



John Doe

Weight: 70.5kg
Height: 175cm

Pills: Metformin

Last Checkup: 5/19/2024

Blood Sugar Level Over Time

Blood Sugar Records

May 2024

May 19, 2024 at 10:00:00 AM ● 85 mg/dL before meal

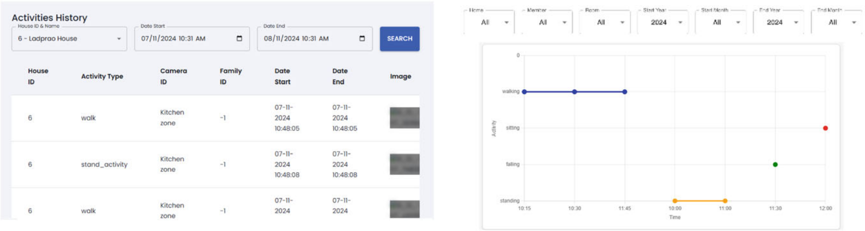
May 19, 2024 at 11:00:00 AM ● 95 mg/dL after meal

May 19, 2024 at 12:00:00 PM ● 110 mg/dL before meal

May 1, 2024 at 8:00:00 AM ● 85 mg/dL before meal

May 1, 2024 at 12:00:00 PM ● 95 mg/dL after meal

Fig. 4 Visualizations for system A: Blood glucose measurement history



Activities History

House ID & Name: 6 - Ladprao House

Date Start: 07/11/2024 10:31 AM

Date End: 08/11/2024 10:31 AM

House ID	Activity Type	Camera ID	Family ID	Date Start	Date End	Image
6	walk	Kitchen zone	-1	07-11-2024 10:48:05	07-11-2024 10:48:05	
6	stand_activity	Kitchen zone	-1	07-11-2024 10:48:08	07-11-2024 10:48:08	
6	walk	Kitchen zone	-1	07-11-2024	07-11-2024	

Home: All

Member: All

Room: All

Start Year: 2024

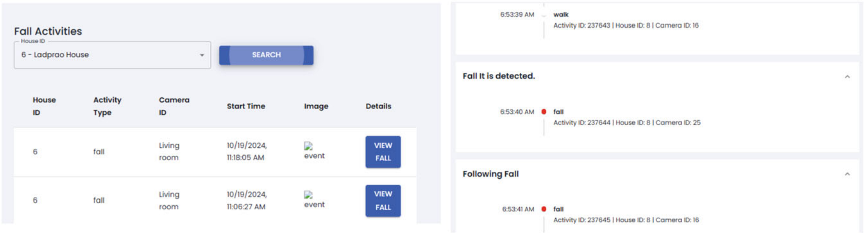
Start Month: All

End Year: 2024

End Month: All

a) Activity History

b) Daily Activity Timeline



Fall Activities

House ID: 6 - Ladprao House

House ID	Activity Type	Camera ID	Start Time	Image	Details
6	fall	Living room	10/19/2024, 11:18:05 AM	event	VIEW FALL
6	fall	Living room	10/19/2024, 11:06:27 AM	event	VIEW FALL

6:53:39 AM walk
Activity ID: 237643 | House ID: 8 | Camera ID: 16

Fall It is detected.

6:53:40 AM ● fall
Activity ID: 237644 | House ID: 8 | Camera ID: 25

Following Fall

6:53:41 AM ● fall
Activity ID: 237645 | House ID: 8 | Camera ID: 16

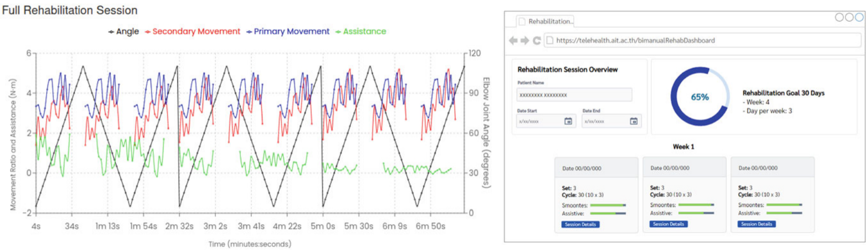
c) Fall Incident Records

d) Fall Timeline

Fig. 5 Visualizations for system B: History of activities, fall events, and detailed fall information with nearby activities

visualizations give healthcare professionals a structured overview to monitor analyze and ensure safety of elderly and disabled individuals effectively.

Figure 6 presents selected *visualizations for bimanual rehabilitation exercises and progresses* (System C). Figure 6a displays a rehabilitation session overview for a physical therapist, showing primary and secondary elbow movements during flexion, extension, and assisted exercises. Figure 6b summarizes sessions with metrics like weekly goals, progress percentages, and session details, helping patients and physiotherapists for progress tacking. Figure 6c details individual exercise cycles, highlighting flexion and extension movements with adjustments for assistance and angular motion to improve technique. Figure 6d uses a sunburst chart to illustrate flexion, and extension ranges of primary and secondary elbows, showing motion patterns and coordination. These visualizations demonstrate how the prototype system enhances monitoring, decision-making, and rehabilitation through comprehensive data visualization, supporting healthcare providers and caregivers in delivering informed and effective care.



https://telehealth.ait.ac.th/bimanualRehabDashboard

Rehabilitation Session Overview

Patient Name: xxxxxxxx xxxxxxxx

Date Start: x/xx/xxxx

Date End: x/xx/xxxx

65%

Rehabilitation Goal 30 Days
- Week: 4
- Day per week: 3

Week 1

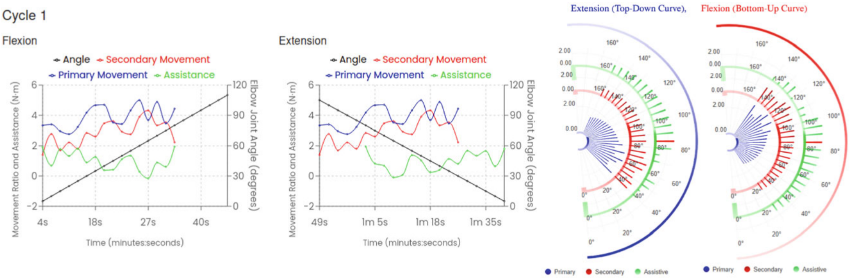
Date 00/00/000
Set: 3
Cycle: 30 (10 x 3)
Smoontes:
Assistive:

Date 00/00/000
Set: 3
Cycle: 30 (10 x 3)
Smoontes:
Assistive:

Date 00/00/000
Set: 3
Cycle: 30 (10 x 3)
Smoontes:
Assistive:

a) Overall bimanual rehabilitation recording

b) Bimanual rehabilitation session summary



c) Analysis of flexions and extensions of each exercise cycle

d) Sunburst charts for flexion and extension of primary and secondary elbow movement

Fig. 6 Visualizations for system C: Bimanual rehabilitation exercises and progresses

5 Conclusion

This paper has presented a comprehensive human-AI collaboration platform designed to address key challenges in health and wellness domains, supporting diverse stakeholders such as healthcare professionals, patients, vulnerable groups, caregivers, and clinical researchers. The platform integrates several advanced subsystems, including a wearable Raman Spectroscopy system for non-invasive blood glucose monitoring, video and CSI Wi-Fi analytics for activity and fall monitoring, a bimanual rehabilitation system, and an analytics and visualization dashboard. These systems enhance patient safety, improve decision-making, and support personalized care through real-time monitoring and actionable data insights while maintaining a high standard of privacy.

The proposed platform underscores the potential of AI to augment human capabilities, providing innovative solutions that address resource constraints and increase the accessibility of quality care. In future work, we aim to evaluate the outcomes of human-AI collaboration within the platform, focusing on usability, user satisfaction, and overall effectiveness in real-world applications. These evaluations will be essential for refining the platform to ensure it meets the practical needs of stakeholders while optimizing the quality and impact of AI-supported health and wellness services.