Hand Gesture Detection Implemented based on Long Short-Term Memory (LSTM) Method

I Gede Wiryawan
Information Technology Department
Politeknik Negeri Jember
Jember, Indonesia
wiryawan@polije.ac.id

Taufiq Rizaldi
Information Technology Department
Politeknik Negeri Jember
Jember, Indonesia
taufiq_r@polije.ac.id

Pramuditha Shinta Dewi P.
Information Technology Department
Politeknik Negeri Jember
Jember, Indonesia
pramuditha@polije.ac.id

Arvita Agus Kurniasari
Information Technology Department
Politeknik Negeri Jember
Jember, Indonesia
arvita@polije.ac.id

Abstract— The Indonesian government encourages accessibility of information that is friendly to people with disabilities, one of which is through the development of information and communication technology. Efforts to increase accessibility of information and encourage independence of people with disabilities need to be supported by the right solutions. According to the Central Statistics Agency, there were 0.68% of the total population of Indonesia in 2019, this data shows that deafness is one of the highest disabilities in Indonesia. Efforts to increase accessibility of information and encourage independence of people with disabilities need to be supported by the right solutions. One potential solution is the development of a self-service system that is friendly to the deaf. This study aims to develop a self-service system that is friendly to the deaf and helps in obtaining information and services independently. The results achieved in this study are in the application of hand signal detection using the Long Short-Term Memory method which can overcome the problem of long-distance dependency and improve performance in recognizing complex hand signal patterns. The hand signal recognition feature can be improved by overcoming the problem of long-distance dependency with a maximum user distance of 1.25 meters, the system can still recognize hand signals well. It is hoped that in the future, more in-depth studies can be carried out on long-distance dependency for variations of other hand signal recognition methods, so that people with disabilities can more easily use the self-service system.

Keywords— Deaf, Hand gestures, LSTM, Self Service

I. INTRODUCTION

Eleven million Indonesians with deafblindness currently face various barriers in their daily lives, such as lack of social equality, communication difficulties, limited access to education, and discrimination in employment. Unlike other special needs, the main limitation of deaf disabilities lies in their ability to communicate verbally. According to Moores, deafness is defined as a condition in which individuals are unable to hear, affecting their ability to understand and produce speech [1]. In 2019, according to the National Bureau of Statistics, there were 1,820,000 people in Indonesia who were deaf. This number is equivalent to 0.68% of Indonesia's total population at that time, which reached 268,100,000 people [2]. According to the official website of the Ministry of Social Affairs based on 2020 running data from the Central Bureau of Statistics (BPS), the number of people with disabilities in Indonesia reached 22.5 million or around five percent. This data shows that deafness is a significant disability in Indonesia.

The Indonesian government shows its attention to the barriers experienced by deaf people by encouraging disability-friendly information accessibility. The Indonesian government is currently paying more attention and support to people with disabilities so that they can live independently and have equal opportunities, including those who are deaf. This can be seen from the steps taken by the Ministry of Social Affairs in encouraging the

accessibility of disability-friendly information. Government efforts to improve information accessibility and encourage the independence of persons with disabilities are important steps to realize equality and inclusiveness for the entire community. One of the government's efforts in disability-friendly programs is the development of disability-friendly Information and Communication Technology (ICT) [3]. However, these efforts are still minimal in the use of disability-friendly communication technology. Seeing these problems requires concrete and sustainable efforts from various parties to overcome these obstacles and realize equality and inclusiveness for deaf people in Indonesia. Efforts to improve information accessibility and encourage the independence of people with disabilities need to be supported with the right solution. That is the background and urgency of this research.

One potential solution is the development of a deaf-friendly independent Financial Kiosk (Kiosk). Kiosk is a machine designed to facilitate buying and selling transactions between consumers and machines [4]. The disability-friendly Kiosk is designed to provide access to information and services that are easily accessible and understood by deaf people. One of the features that can be applied in disability-friendly kiosk is sign language recognition technology. The application of sign language recognition technology allows deaf people to communicate with the Kiosk directly using sign language. Some previous studies have used self service technology, one of which is to order products using hand gesture recognition technology using Kinect sensors and deep learning frameworks [5].

Electronic kiosk have become an important element in IT, offering intuitive interfaces with technologies such as facial recognition, voice, and motion sensors [6]. Its security is enhanced through biometric technology [7]. The integration of artificial intelligence and data analysis enables personalized services and operational efficiency [8]. In the healthcare sector, kiosk provides important self-service [4]. In retail and hospitality, they enhance customer experience and facilitate self-checkout. Integration with IoT enables real-time monitoring and better connectivity. In education, kiosk provide quick and independent access to learning materials [9]. Thus, kiosk continue to grow as an important solution in digital transformation in various sectors.

Long Short-Term Memory (LSTM) is a type of artificial neural network (RNN) known for its ability to capture long-term dependencies on sequential data [10]. These networks are structured with unique memory units called LSTM cells, which allow them to effectively extract time-series features from input data [11]. LSTM models have found successful applications in areas such as human activity recognition [12], fault diagnosis [13], action recognition [14], and dynamic process quality prediction [15]. One of the main advantages of LSTM networks is their ability to mitigate the problems of gradient loss and explosion commonly faced by traditional RNNs [16].

The ability of LSTM architecture to store long memories and capture complex temporal features has made it a popular choice in various fields, including behavioral modeling, digital distortion, and human action recognition [17] [18]. In addition, LSTM networks have been combined with attention mechanisms to improve memory term reduction in digital predistortion applications [19]. Overall, the adaptability and effectiveness of LSTM networks in handling sequential data make them a valuable tool in various fields, ranging from healthcare to industrial processes [20] [15] [21] [22].

Hand gesture recognition using the Long Short-Term Memory (LSTM) method has been an interesting research topic in the field of information technology, especially in the context of implementation on kiosk [23]. This technology offers the potential to improve interactivity, efficiency, and overall user experience in a kiosk environment. Several related studies have also made important contributions in expanding the understanding and application of this technology. For example, the first study reference [24] presents a hand gesture recognition system using Kinect sensors and a deep learning framework, which can recognize various hand gestures with a high degree of accuracy, enabling intuitive interaction between humans and computers.

The second article introduced a real-time hand gesture recognition system based on Convolutional Neural Networks (CNN), which is suitable for interactive applications in various domains due to its ability to recognize hand gestures quickly and accurately [5]. Furthermore, the third study discusses the use of LSTM in hand gesture recognition, where LSTM can overcome the problem of long-range dependencies and improve performance in recognizing complex hand gesture patterns [25]. The last article discusses a hand gesture recognition system for human-computer interaction in smart kiosk applications, focusing on smart retail applications to provide a more intuitive and efficient shopping experience for kiosk users [26]. The integration of findings from these studies provides a strong foundation for the development of an implementation of hand gesture detection with LSTM in the context of k-stores to improve interaction and service effectiveness.

Based on the literature review above, it can be concluded that the Long Short Term Memory (LSTM) method can be used as a solution for people with deaf disabilities to be able to interact, especially in placing orders on transactions in various business fields. So that this method can overcome problems in improving information accessibility and encourage the independence of people with disabilities need to be supported with the right solution. The utilization of hand gesture recognition and KiosK can be applied as a solution to help people with deaf disabilities. The output of this research is in the form of kiosk integration with hand gesture recognition using the Long Short-Term Memory (LSTM) method which aims to facilitate deaf people to get the information and services they need independently. So that later this deaf-friendly Kiosk can be applied to Kedai Susu Tuli (K-Suli) which is an MSME run by people with deaf disabilities [27]. With this solution, it is hoped that deaf people can live more independently and have the same opportunities as other people.

II. METHOD

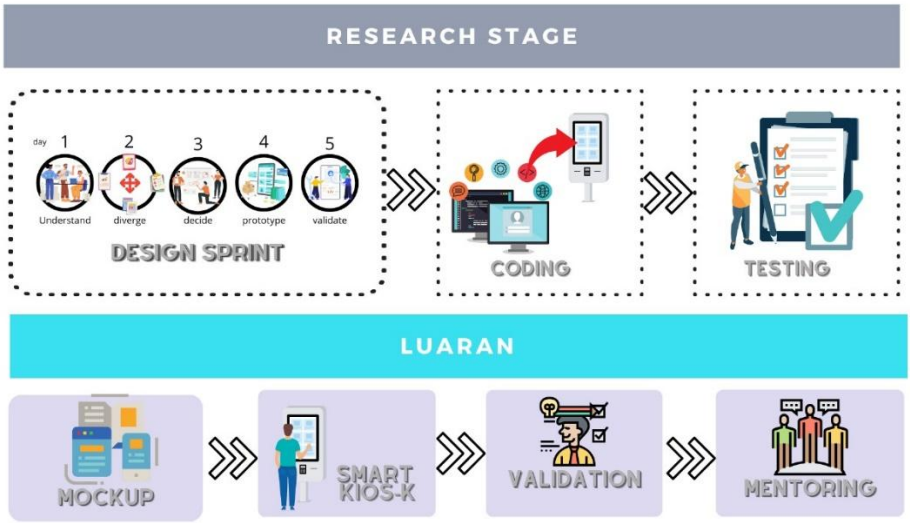

Figure 1. Hand gesture detection development stages

This research aims to improve efficiency and affordability in sales at Kedai Susu Tuli (K-SULI) through the integration of Kiosk-K, hand gesture detection, and a sign language-sensitive self-service system. Research stage shown on figure 1 above. The initial phase included an in-depth understanding of the business landscape and consumer needs, followed by a divergence of ideas to generate diverse solution concepts. Selected concepts were then realized in the form of prototypes that visually and functionally depicted the solution implementation. We validated the prototypes through testing with users and relevant experts, before proceeding to the application coding and deployment stage. The

final step involved assisting in the implementation of the product at Kedai Susu Tuli. As a result, a smart Kiosk product design mockup was produced that has been validated by experts, as a guide for implementation and further development.

Figure 2. Design Sprint stages

A design sprint consists of five stages which are: understand, diverge, decide, prototype and validate. Figure 2 shows the stage of design sprint. The following is an explanation of each stage of the design sprint:

A. Understand

The initial stage of Understand is to equalize a perception of a research discussion. At this stage, exploring information needs by interviewing a number of experts or experts in the field of Special Education. The resource person (expert) in this research is Bhennita Sukmawati, S.Psi., M.Psi Special Education Study Program IKIP Jember. The resource person is an academic who has a field of expertise in psychology.

B. Diverge

In the diverge stage of this research, the focus is to gather as many ideas and potential solution concepts as possible to address the challenges of efficiency and affordability in the sales of Kedai Susu Tuli (K-SULI). During this stage, the research team will hold intensive and collaborative brainstorming sessions, both individually and in groups, to create a variety of innovative and creative ideas.

The frontend technology uses React JS for a dynamic and responsive user interface, and Tailwind CSS to quickly design layouts and design styles. MySQL is used as a relational database to store sales data, inventory, and operational information of Kedai Susu Tuli. For the AI technology and server, Media Pipe is used for hand gesture detection, Flask as the Python back-end framework for business logic and API, and Socket.IO for real-time communication between server and client. OpenCV is also used for image processing of hand gesture detection.

C. Decide

In the Decide stage, the research team will evaluate all the ideas and solution concepts generated during the diverge stage. The goal is to select one solution concept that is most suitable and has the greatest potential to address the challenges of efficiency and affordability in sales at Kedai Susu Tuli (K-SULI). The evaluation was conducted by considering various factors, including technical feasibility, business needs, user needs, potential impact, and resource availability. In-depth discussions and analysis were conducted to consider the advantages and disadvantages of each solution concept.

D. Prototype

This stage will design the interface of the system to be created based on the approved system design. In Figure 3, we can see the prototype of the smart kiosk that has been successfully developed. The kiosk interface displays the available menu options with clear details about each menu item. Users can easily select the desired menu and order the desired quantity through an intuitive interface. In addition, the system also provides information

about the total payment directly, making it easier for users to make transactions efficiently. Thus, this kiosk not only provides practical self-service, but also improves efficiency and affordability in the sales process at Kedai Susu Tuli (K-SULI).

E. Validate

In the last stage, the prototype is tested directly to potential users. The goal is to get feedback from users by knowing whether the system built is in accordance with the design and process. If the results obtained do not match the initial design, the team will adjust according to the wishes of potential users. This stage minimizes the occurrence of errors during the program code implementation process. So this stage is very important to do to get maximum results.

III. RESULTS

Based on the stages of the design sprint method described in the previous section, this study begins with the understanding stage. This first stage involves an expert in the field of psychology who aims to gain knowledge about the characteristics and types of disabilities, self-development in disabilities that can avoid adverse effects and problems. This study involves several characteristics and types of deaf disabilities, and according to the results of discussions with experts found some of the most frequently experienced problems and also along with the ranking. Table 1 below shows the ranking of the most common problems experienced by people with deaf disabilities.

Tabel 1. Ranking and The most common problems experienced by people with deafblindness

Ranking	The most common problems experienced by people with deafblindness
1	Social interaction (communication with friends and community)
2	Internal problems (stress, depression, insecurity, etc.)
3	Not being able to participate in talent and interest development activities so that there is a lack of connection with other people
4	Time management in doing assignments
5	Related to assistance (the companion is not present and the companion comes from a different department so it is difficult to understand the material)
6	Vocabulary
7	Environmental support for deaf disabilities
8	Procedures and access
9	Learning strategy
10	There are no sign language interpreters available when participating in activities

The result of this understand stage is an understanding of ten problems and especially with the highest rank, namely social interaction in the form of communication with friends and society. The many problems that may be experienced by people with deaf disabilities are continued to the diverge stage, where in the second stage of the design sprint a brainstorming session is held with the development team until creative ideas are obtained to solve the problems of people with deaf disabilities.

Especially in the world of food and beverage businesses, through this study, an assistive software is developed that can facilitate people with deaf disabilities in transactions in food and beverage businesses. The concept of self-service using kiosk is the main thing in the development of this software. People with deaf disabilities are enough with hand gestures to be able to order food and drinks. The solution is the result of the next stage, namely the decide stage. At this stage, a decision is made regarding the solution which is the result of the diverge stage of several problems that have been obtained at the understanding stage.

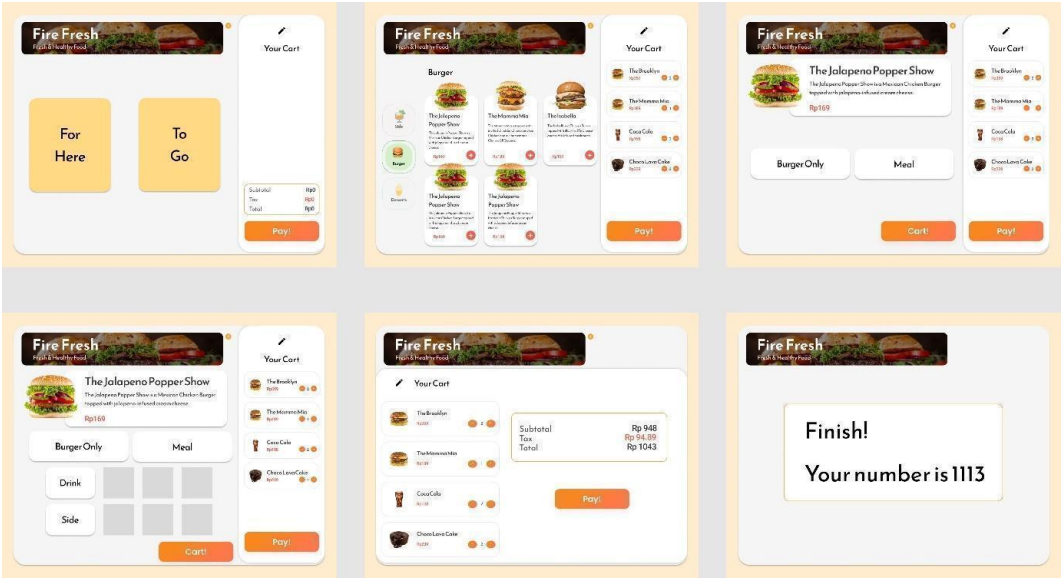

Figure 3. Prototype stage results

After the understanding, diverge, and decide stages, this study proceeds to the prototyping stage. In this stage, the development of a special kiosk system for people with deaf disabilities is carried out. The results of this prototype stage can be seen in Figure 3 above. Starting from choosing dine in or take away, to getting a queue number. This prototype was developed using several technologies including AI Server. The use of this AI Server has benefits in the kiosk system implementing hand gesture recognition using the LSTM method. AI Server consists of Media Pipe, socket.io, Flask, and OpenCV. The four assistive devices work in the system to apply the LSTM algorithm to recognize hand gestures from people with deaf disabilities.

The technology in developing the front end display of the application is used Tailwind CSS framwork and also ReactJS, these tools are used because of their flexibility in using various palforms such as computer devices, tablets, and smartphones. Finally, MySQL is used for the database. Figure 4 below shows the architecture or series of technologies used in the prototyping stage.

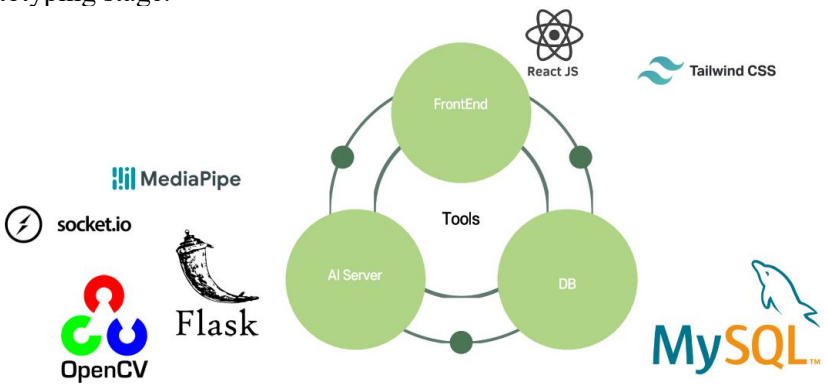

Figure 4. Architecture of Technology

The last stage in the design sprint is validate, at this stage several things were found including The success of this implementation is proven through functionality testing with a value of 90% of the overall features of the self-service system that can be used as a hand gesture detector from people with deaf disabilities. The hand gesture recognition feature can be improved by overcoming the remote dependency problem with a maximum user distance of 1.25 meters, the system can still recognize hand gestures well. It is hoped that in the future it can be studied more deeply about long-range dependencies for other variations of hand

gesture recognition methods, so that people with disabilities can more easily use self-service systems.

IV. DISCUSSION

Based on the results of the development and implementation of the hand gesture detection system in the previous section, the results of testing the effective distance to be captured by the detection system are obtained. The test results were carried out as many as 350 test repetitions with various distance sizes in centimeters. The distance measured is between the hand position of the user of the detection system and the camera of the device that has a hand gesture detection system installed. The success rate is a way to measure this effectiveness. The number of constant distance variables used to measure the level of effectiveness is 7 variables, ranging from 25 cm to 175 cm. Table 2 below shows the test results with 7 distance variables.

The test results above show that the effectiveness of this hand gesture detection system is at a distance of 1.25 meters where the success rate is at 100%, while above that the success rate has started to decline. This is the reason for taking as many as 7 constant distance variables in this test, because if continued the success rate will continue to decline. Overall, from 350 repetitions of the test, the success rate of this detection system is 92.57%. This shows that the Long Short Term Memory (LSTM) method can be used in hand gesture detection. With this LSTM method can help solve the problem of Social interaction (communication with friends and community) for people with deaf disabilities.

Tabel 2. Distance Effectiveness Testing Results

Distance (cm)	Success Detected (times)	Failed Detected (times)
25	46	4
50	48	5
75	47	3
100	48	2
125	50	0
150	45	5
175	40	10

V. CONCLUSION

Based on the results of the design sprint methodology that has been carried out, starting from the understanding to validate stage along with the technology used in the AI Server, the following things have been found. The success of this implementation is proven through functionality testing with a value succes rate of 92.57% of the overall features of the self-service system that can be used as a hand gesture detector from people with deaf disabilities. The hand gesture recognition feature can be improved by overcoming the remote dependency problem with a maximum user distance of 1.25 meters, the system can still recognize hand gestures well. It is hoped that in the future it can be studied more deeply about long-range dependencies for other variations of hand gesture recognition methods, so that people with disabilities can more easily use self-service systems.

ACKNOWLEDGEMENT

The authors would like to thank the funders of the research that supported this article, namely from the Non-Tax State Revenue (PNBP) channeled by the Center for Research and Community Service (P3M) of Jember State Polytechnic. In addition, the author also thanks those who play an active role in research and writing articles, namely, Bhennita Sukmawati, S.Psi., M.Psi.

REFERENCES

[1] Suharsiwi, *Pendidikan Anak Berkebutuhan Khusus*, I. Yogyakarta: CV Prima Print, 2017.

[2] Ministry of Health of the Republic of Indonesia, "Guidelines for the management of stimulation, detection, and early intervention in child development at the basic health service level," 2019.

[3] D. N. S. Ramadhan, "Kementerian Sosial mudahkan penyandang disabilitas mengakses informasi," www.antaranews.com. Accessed: Mar. 20, 2024. [Online]. Available: https://www.antaranews.com/berita/3691179/kementerian-sosial-mudahkan-penyandang-disabilitas-mengakses-informasi

[4] A. Jayanthi, M. Rifky Gunawan, H. Ashaury, and U. Jenderal Achmad Yani Jl Terusan Sudirman, "Sistem Informasi Kiosk Antrean Pelayanan Pada Klinik Kesehatan dengan Metode Prototyping," in *Seminar Nasional Informatika dan Aplikasinya (SNIA)*, Cimahi, 2021, p. 2021.

[5] G. B. Prananta, H. A. Azzikri, and C. Rozikin, "Deteksi dan Pengenalan Gesture Tangan Secara Real-Time Menggunakan Jaringan Saraf Tiruan Konvolusional," *METHODIKA*, vol. 9, no. 2, pp. 30–34, 2023, [Online]. Available: https://ejurnal.methodist.ac.id/index.php/methodika/article/view/1911/1578

[6] S. Ariyanti, S. S. Adi, and S. Purbawanto, "Sistem Buka Tutup Pintu Otomatis Berbasis Suara Manusia," *Elinvo (Electronics, Informatics, Vocat. Educ.*, vol. 3, no. 1, pp. 83–91, 2018, doi: 10.21831/elinvo.v3i1.19076.

[7] N. Hartono, A. Rizaldy, and Ni. A. Lestari, "Studi Literature Sistem Keamanan Biometrik Untuk Verifikasi dan Transaksi Dompet Digital," *J. Software, Hardw. Inf. Technol.*, vol. 2, no. 2, pp. 10–14, 2022, doi: 10.24252/shift.v2i2.30.

[8] M. Siska, I. Siregar, A. Saputra, M. Juliana, and M. T. Afifudin, "Kecerdasan Buatan dan Big Data dalam Industri Manufaktur: Sebuah Tinjauan Sistematis," *Nusant. Technol. Eng. Rev.*, vol. 1, no. 1, pp. 41–53, 2023, doi: 10.55732/nter.v1i1.1119.

[9] Christian, Y. Wijaya, and B. Mario, "Rancang Bangun Sistem Kiosk Untuk Daftar Hadir Seminar Dan Workshop," *Zo. J. Sist. Inf.*, vol. 5, no. 2, pp. 380–391, 2023, doi: 10.31849/zn.v5i2.13857.

[10] H. Okut, "Deep Learning for Subtyping and Prediction of Diseases: Long-Short Term Memory," in *Deep Learning Applications*, 1st ed., no. tourism, Pier Luigi Mazzeo and Paolo Spagnolo, Eds., London, United Kingdom: IntechOpen, 2021, ch. 3, pp. 57–79. [Online]. Available: http://dx.doi.org/10.1039/C7RA00172J%0Ahttps://www.intechopen.com/books/advanced-biometric-technologies/liveness-detection-in-biometrics%0Ahttp://dx.doi.org/10.1016/j.colsurfa.2011.12.014

[11] W. Lei, J. Luo, F. Hou, L. Xu, R. Wang, and X. Jiang, "Underground cylindrical objects detection and diameter identification in GPR B-scans via the CNN-LSTM framework," *Electron.*, vol. 9, no. 11, pp. 1–16, 2020, doi: 10.3390/electronics9111804.

[12] S. Wilhelm Pienaar and R. Malekian, "Human Activity Recognition Using LSTM-RNN Deep Neural Network Architecture," in *2019 IEEE 2nd Wireless Africa Conference*, Pretoria, South Africa: IEEE, 2019, pp. 1–5. doi: 10.1109/ICDSE47409.2019.8971476.

[13] X. Li, K. Su, Q. He, X. Wang, and Z. Xie, "Research on Fault Diagnosis of Highway Bi-LSTM Based on Attention Mechanism Indexed by," *Eksploat. i Niezawodn.*, vol. 25, no. 2, pp. 0–3, 2023, doi: 10.17531/ein/162937.

[14] K. Hu, F. Zheng, L. Weng, Y. Ding, and J. Jin, "Action recognition algorithm of spatio–temporal differential lstm based on feature enhancement," *Appl. Sci.*, vol. 11, no. 17, pp. 1–15, 2021, doi: 10.3390/app11177876.

[15] X. Yuan, L. Li, Y. Wang, C. Yang, and W. Gui, "Deep learning for quality prediction of nonlinear dynamic processes with variable attention-based long short-term memory network," *Can. J. Chem. Eng.*, vol. 98, no. 6, pp. 1377–1389, 2020, doi: 10.1002/cjce.23665.

[16] Y. Li, X. Chen, C. Li, G. Tang, Z. Gan, and X. An, "A Hybrid Deep Interval Prediction Model for Wind Speed Forecasting," *IEEE Access*, vol. 9, pp. 7323–7335, 2021, doi: 10.1109/ACCESS.2020.3047903.

[17] W. Lu, H. Rui, C. Liang, L. Jiang, S. Zhao, and K. Li, "A method based on GA-CNN-LSTM for daily tourist flow prediction at scenic spots," *Entropy*, vol. 22, no. 3, pp. 1–18, 2020, doi: 10.3390/e22030261.

[18] Y. Ma, "Two models for predicting stock prices in combination with LSTM," *Highlights Business, Econ. Manag.*, vol. 5, pp. 664–673, 2023, doi: 10.54097/hbem.v5i.5256.

[19] G. Javid, D. Ould Abdeslam, and M. Basset, "Adaptive online state of charge estimation of evs lithium-ion batteries with deep recurrent neural networks," *Energies*, vol. 14, no. 3, 2021, doi: 10.3390/en14030758.

[20] Z. Cao, Y. Wang, C. C. Chu, and R. Gadh, "Scalable distribution systems state estimation using long short-term memory networks as surrogates," *IEEE Access*, vol. 8, pp. 23359–23368, 2020, doi: 10.1109/ACCESS.2020.2967638.

[21] X. Wang, Z. Miao, R. Zhang, and S. Hao, "I3D-LSTM: A New Model for Human Action Recognition," in *IOP Conference Series: Materials Science and Engineering*, Zhuhai, China, 2019. doi: 10.1088/1757-899X/569/3/032035.

[22] H. Yu, G. Xu, T. Liu, J. Huang, and X. Zhang, "A Memory Term Reduction Approach for Digital Pre-Distortion Using the Attention Mechanism," *IEEE Access*, vol. 7, pp. 38185–38194, 2019, doi: 10.1109/ACCESS.2019.2906370.

[23] A. A. Kurniasari, I. G. Wiryawan, P. Shinta, and D. Puspitasari, "Implementasi Deteksi Gerakan Tangan untuk Sistem Interaktif Kios menggunakan Metode Long Short-Term Memory (LSTM) Implementation of Hand Gesture Detection for Kiosk Interactive System using Long Short-Term Memory (LST," *Komputika J. Sist. Komput.*, vol. 14, no. April, pp. 31–39, 2025, doi: 10.34010/y895rc89.

[24] A. Graves and J. Schmidhuber, "Offline handwriting recognition with multidimensional recurrent neural networks," *Adv. Neural Inf. Process. Syst. 21 - Proc. 2008 Conf.*, pp. 545–552, 2009.

[25] H. M. Putri, F. Fadlisyah, and W. Fuadi, "Pendeteksian Bahasa Isyarat Indonesia Secara Real-Time Menggunakan Long Short-Term Memory (LSTM)," *J. Teknol. Terap. Sains 4.0*, vol. 3, no. 1, pp. 663–675, 2022, doi: 10.29103/tts.v3i1.6853.

[26] S. S. Makahaube, A. M. Sambul, and S. R. Sompie, "Implementation of Gesture Recognition Technology for Automated Education Service Kiosk," *J. Tek. Inform.*, vol. 16, no. 4, pp. 465–472, 2021.

[27] I. Nawawi, "K-Suli, Kedai Susu Tuli Tongkrongan Kawan Disabilitas Di Jember, Sediakan Bahasa Isyarat," Surya travel. Accessed: Mar. 20, 2024. [Online]. Available: https://suryatravel.tribunnews.com/2024/01/03/k-suli-kedai-susu-tuli-tongkrongan-kawan-disabilitas-di-jember-sediakan-bahasa-isyarat