Jurnal Ilmu Komputer dan Informasi (Journal of Computer Science and Information)
17/2 (2024), 185-194. DOI: http://dx.doi.org/10.21609/jiki.v17i2.1280

Hand Sign Interpretation through Virtual Reality Data Processing

Teja Endra Eng Tju, Muhammad Umar Shalih

Faculty of Information Technology, Universitas Budi Luhur,
Jl. Ciledug Raya, Petukangan Utara, Jakarta Selatan, 12260, Indonesia

E-mail: teja.endraengtju@budiluhur.ac.id

Abstract

The research lays the groundwork for further advancements in VR technology, aiming to develop devices capable of interpreting sign language into speech via intelligent systems. The uniqueness of this study lies in utilizing the Meta Quest 2 VR device to gather primary hand sign data, subsequently classified using Machine Learning techniques to evaluate the device's proficiency in interpreting hand signs. The initial stages emphasized collecting hand sign data from VR devices and processing the data to comprehend sign patterns and characteristics effectively. 1021 data points, comprising ten distinct hand sign gestures, were collected using a simple application developed with Unity Editor. Each data contained 14 parameters from both hands, ensuring alignment with the headset to prevent hand movements from affecting body rotation and accurately reflecting the user's facing direction. The data processing involved padding techniques to standardize varied data lengths resulting from diverse recording periods. The Interpretation Algorithm Development involved Recurrent Neural Networks tailored to data characteristics. Evaluation metrics encompassed Accuracy, Validation Accuracy, Loss, Validation Loss, and Confusion Matrix. Over 15 epochs, validation accuracy notably stabilized at 0.9951, showcasing consistent performance on unseen data. The implications of this research serve as a foundation for further studies in the development of VR devices or other wearable gadgets that can function as sign language interpreters.

Keywords: *Sign Language, Quaternion, Padded, RNN*

1. Introduction

The advancement of Virtual Reality (VR) technology has revolutionized the human interaction paradigm within the digital sphere [1]. Its ever-expanding relevance has permeated various facets of life, encompassing research across education, training, medicine, entertainment, art, and diverse applications [2], [3].

The research exploring the applications of Meta Quest 2 [4] VR technology across various domains has made remarkable strides. VR has been investigated as a vital learning aid in education and culture, especially during the pandemic, facilitating interactive and effective educational approaches [5]. VR research outcomes have facilitated simulations of dam failures [6], preservation of historical artifacts [7], cultural and touristic promotions [8], and even the preservation of Gamelan music [9]. In dental education, VR has emerged as an innovative training tool [10], [11]. Within the healthcare and nursing sectors, VR is employed to alleviate post-operative pain [12], train multiple sclerosis patients [13], [14], and as an assistive tool in surgical robotics [15]. Moreover, it contributes to the well-being of healthcare professionals, offering relaxation during the pandemic [16]. The utilization of VR research outcomes extends to mental health training [17], memory assessment [18], vertigo therapy [18], shoulder rehabilitation monitoring [19], nerve disorder examinations [20], and neck movement assessments [21].

On the other hand, VR research ventures into the industrial domain for the development of interactive robots [22]. Within the realm of entertainment, VR plays a pivotal role in game development and experiences within the Metaverse [23]. In the communication field, VR is utilized for video conferencing [24]. Even in the architecture field, VR is employed for spatial mapping and visualization, supporting the design of architectural projects [25].

Furthermore, the Meta Quest 2, a leading VR device, has garnered attention for further research and development. Efforts have been made to measure controller precision [26], hand movement accuracy [27], eye-tracking evaluations [28], integration of haptic features [29], [30], and the creation of 3D Avatar applications [31].

Interpreting hand signs is crucial, especially for those who rely on non-verbal communication methods like sign language. Recent studies have shown a growing interest in using VR to improve hand sign interpretation [32], [33]. This research is vital because it offers new ways to understand and develop technology, potentially benefiting various aspects of life. By collecting data from VR devices and using machine learning, researchers aim to enhance the accuracy of interpreting hand signs, ultimately improving communication for individuals with communication challenges.

The novelty of this research lies in its utilization of the Meta Quest 2 device to gather primary hand sign data, subsequently classified using Machine Learning to evaluate the device's capability in interpreting hand signs. The uniqueness of this study is centered on its emphasis on the initial stages, encompassing the collection of hand sign data from VR devices and the subsequent data processing to comprehend patterns and characteristics of the signs. This strong focus on these stages might offer an advantage, especially if the data processing techniques are sufficiently unique or result in more efficient and accurate solutions for interpreting hand signs.

2. Methodology

The methodology encompassed several stages to achieve the research goals, including Hand Sign Data Collection, VR Data Processing, Interpretation Algorithm Development, and Interpretation Accuracy Evaluation.

Hand Sign Data Collection process involved capturing a diverse range of hand sign data, encompassing various forms, movements, and contextual variations, utilizing the VR Meta Quest 2 device [4], as depicted in Figure 1, comprising a VR Headset and Touch Controller held by users in both hands, equipped with high-precision hand sensors for accurate measurement of hand movements [26], [34].

This data collection aimed to encompass the device's capabilities, focusing on ten frequently used hand sign gestures such as "Hello," "Bye," "Thanks", "Sorry", "Me", "You", "Good", "Bad", "Help", and "Tired". In total, 1021 data points were gathered through the performance of these ten distinct hand sign gestures, illustrated in Figure 2, using the VR device.

Fig. 1. Meta Quest 2 VR Headset and Touch Controller.

Fig. 2. Ten common hand sign movements [35], [36], [37], [38].

The data recording process utilized a simple application developed with Unity Editor [39], visualized in Figure 3, illustrating the dynamic changes in 28 input parameters during hand sign gestures.

Fig. 3. Data recording application interface.

VR Data Processing entailed analyzing the recorded data depicted in a graphical format similar to Figure 4, collected over varying time frames for each hand sign gesture. Each data contained 14 parameters from the left hand and an additional 14 parameters from the right hand, encompassing Trigger Touch, Trigger Pressed, Grip Pressed, Thumb Touch, Position (X, Y, Z), Velocity (X, Y, Z), Quaternion (W, X, Y, Z).

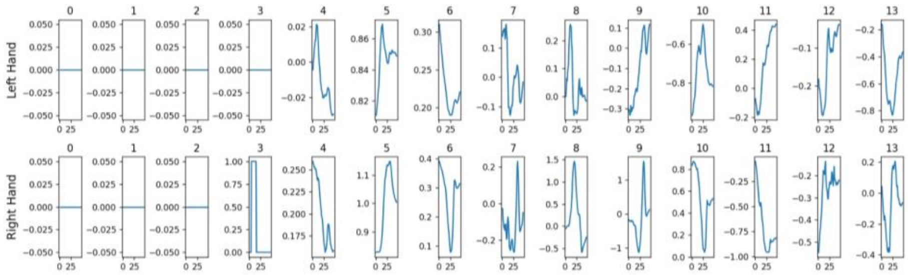

Fig. 4. Data example of a hand sign movement.

The overall graphs in Figure 4 represent the recording of a specific hand sign movement (every single data recorded with the Unity application). All the x-axis (horizontal), serving as the domain, represents the n^{th} Unity sampling. Meanwhile, the values of the y-axis (vertical) corresponding to graphs 0 to 13, along with the explanation of the 14 parameters, are detailed in Table 1.

Table 1. Data explanation of hand sign movement [39].

Graph	Parameter	Value of y-axis
0	Trigger Touch	Boolean: 0 or 1
1	Trigger Pressed	Boolean: 0 or 1
2	Grip Pressed	Boolean: 0 or 1
3	Thumb Touch	Boolean: 0 or 1
4	Position X	Meters (m)
5	Position Y	Meters (m)
6	Position Z	Meters (m)
7	Velocity X	Speed (m/s)
8	Velocity Y	Speed (m/s)
9	Velocity Z	Speed (m/s)
10	Quaternion W	Scalar
11	Quaternion X	Vector
12	Quaternion Y	Vector
13	Quaternion Z	Vector

In processing the data, it was ensured that the collected data was relative to the headset. This alignment aimed to prevent hand movements from affecting body rotation, clearly reflecting the user's facing direction without causing interference. This process involved modifying the center point location and global rotation based on the orientation of the headset. While collecting each hand sign gesture, recorded more than 100 times, variations in data length emerged due to diverse recording periods, requiring standardization through padding techniques [40], [41]. The padding technique involves adding zeros to the sequences with shorter lengths so that all sequences have the same length.

Interpretation Algorithm Development utilized a classification technique, specifically Recurrent Neural Networks (RNN) [42], [43], chosen according to the data's characteristics and the research objectives. The employed layers encompassed the Masking Layer, Long Short-Term Memory (LSTM) Layer, Dropout Layer for Regularization, Batch Normalization Layer, and Dense Layer. The computer specification used for experiments modeling RNN is AMD FX(tm)-6300 Six-Core Processor (6 CPUs) ~3.5GHz, 8192MB RAM, NVIDIA GeForce GTX 1060 3GB.

Interpretation Accuracy Evaluation was conducted using various metrics like Accuracy, Validation Accuracy, Loss, Validation Loss, and Confusion Matrix. The training encompassed 80% (816 data points) of the hand sign dataset, reserving the remaining 20% (205 data points), with the random sampling technique, for validation. Furthermore, the new dataset (100 data points) for testing. The testing set's role is crucial in assessing the model's performance on unseen samples, ensuring its effectiveness beyond the training data**.**

3. Results and Analysis

The initial data processing ensured alignment with the headset for accurate hand movement representation without affecting body rotation. Adjustments to the center point and global rotation based on the headset's orientation were made, aided by additional coding highlighted in Figure 5 within the Unity application used for data recording [44].

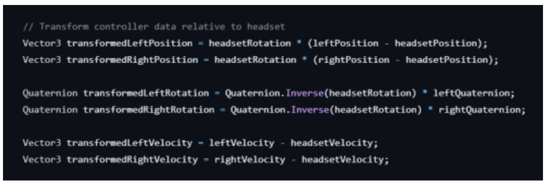

Fig. 5. Coding for headset alignment in Unity Editor.

Table 2. Data distribution for each hand sign movement.

Hand Sign	Number of Data
Hello	101
Bye	101
Thanks	103
Sorry	103
Me	101
You	105
Good	102
Bad	102
Help	101
Tired	102
Total	1.021

The data collection process amassed a total of 1021 data points from 10 distinct hand sign gestures, shown in Table 2 for data distribution (https://github.com/umaruta4/SignLanguage_MTC_Data/tree/main). Figure 6 and Figure 7 illustrate a sample of this dataset, depicting 28 parameters within each data point, including 14 from both the left and right hands.

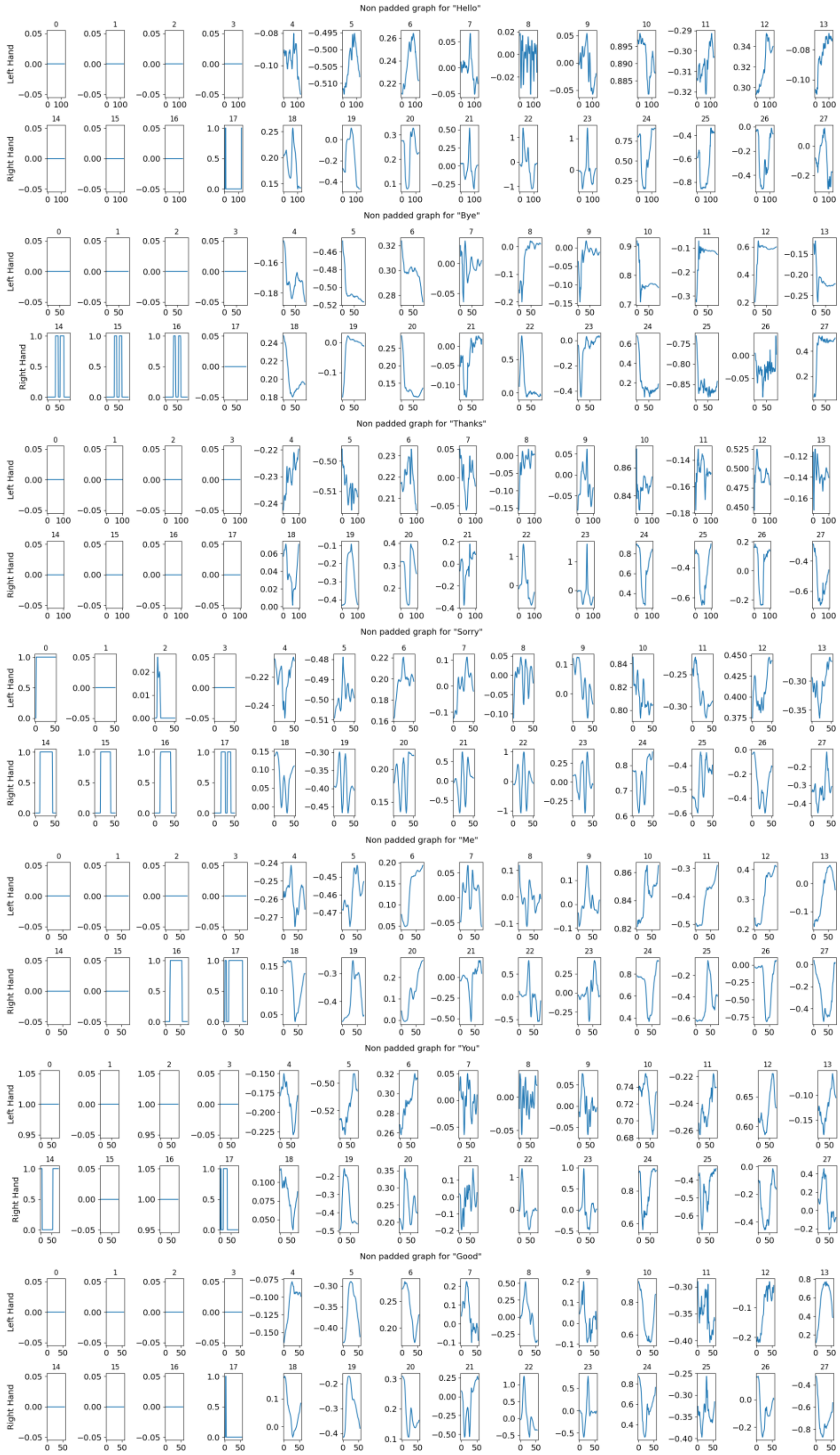

Fig. 6. Non-padded data sample of "Hello," "Bye," "Thanks", "Sorry", "Me", "You", "Good".

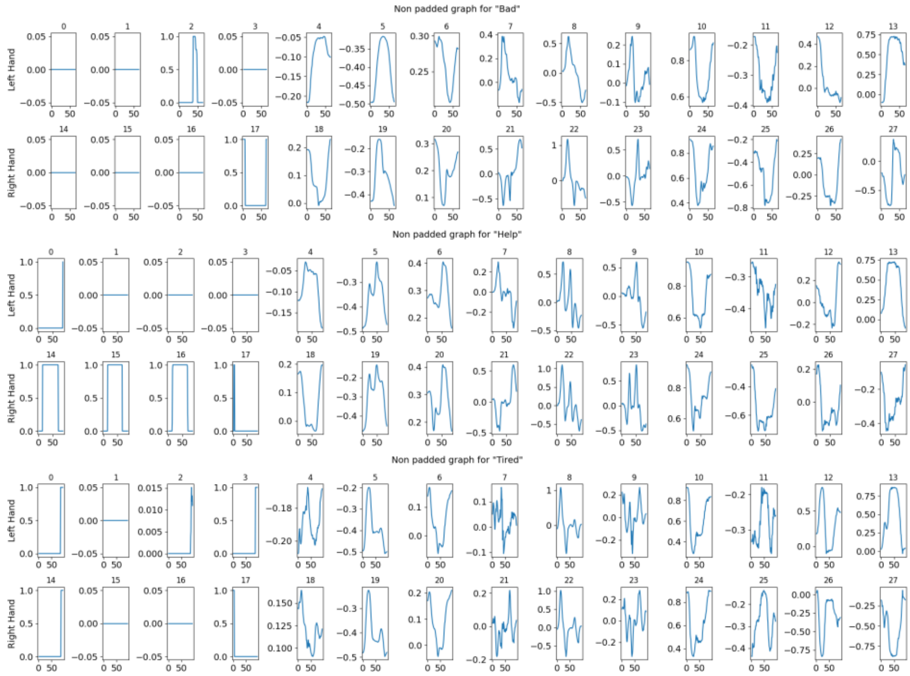

Fig. 7. Non-padded data sample of "Bad", "Help", Tired".

Data processing involved generating graphs and visual representations depicting the processed hand sign data. Notably, during the collection of each hand sign gesture, varying data lengths arose owing to diverse recording periods. To standardize these variations, padding techniques were applied, as depicted in Figure 8 and Figure 9.

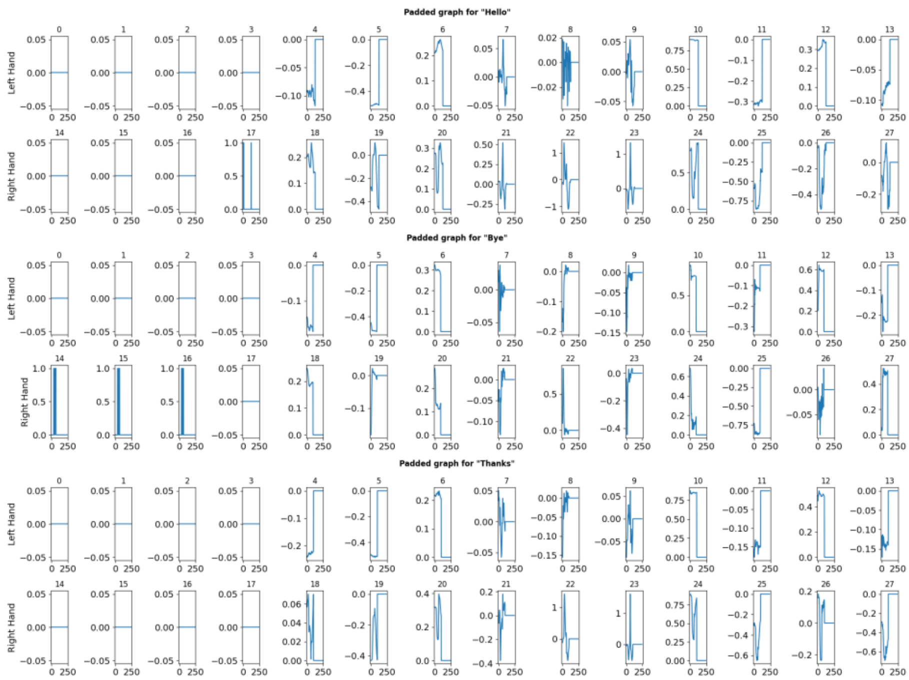

Fig. 8. Padded data sample of "Hello", "Bye", "Thanks".

Fig. 9. Padded data sample of "Sorry", "Me", "You", "Good", “Bad”, “Help”, Tired”.

The RNN model for training the dataset is structured with several essential components. It begins with a Masking layer, which is adept at handling sequences of varying lengths, using a specified mask value of 0 and an input shape denoted as (maxlen, 28), where "maxlen" signifies the maximum sequence length, and 28 represents the input dimensionality. After applied padding, the maximum sequence length becomes 239. Following this, the model integrates two LSTM layers, each comprising 64 units. Notably, the first LSTM layer is configured to return sequences, enabling it to generate an output for each input time step, while the subsequent LSTM layer consolidates temporal information without returning sequences. To prevent overfitting, two Dropout layers are incorporated, each operating with a dropout rate of 20%, thereby randomly setting a fraction of input units to 0 during training. The inclusion of Batch Normalization serves to normalize the activations of the preceding layer within each batch, thereby contributing to the stability and expeditiousness of the model training process. Finally, the model culminates with a Dense layer employing a softmax activation function, producing class probability distributions for the multi-class classification task, with the number of units in this layer aligning with the classes in the label encoder. The model is compiled using the Adam optimizer with a learning rate set at 0.001, aiming to minimize categorical cross-entropy loss and track accuracy as the evaluation metric. This comprehensive architecture is deliberately designed to effectively process sequential data, leveraging LSTM units to capture temporal dependencies while integrating regularization techniques such as dropout and batch normalization to enhance generalization and mitigate overfitting. The model architecture and its components are illustrated in Figure 10, depicting the detailed layers and their connections.

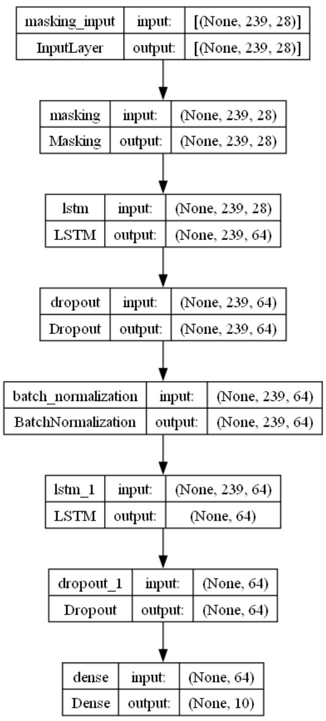

Fig. 10. Layers and Connections of RNN Model.

Throughout the training process, the RNN model underwent 15 epochs, revealing substantial insights into its performance, with each epoch taking approximately 10 to 46 seconds to complete. The accuracy data illustrated a progression, initiating at 0.5833 and steadily reaching a perfect score of 1.0000 by the 7th epoch, maintaining this high accuracy until the training's conclusion. A graphical representation of the accuracy trends across epochs is depicted in Figure 11, showcasing the model's steady improvement over the training period. Concurrently, validation accuracy started at 0.8780, stabilizing impressively at 0.9951 by the 8th epoch, demonstrating the model's consistent performance on unseen data.

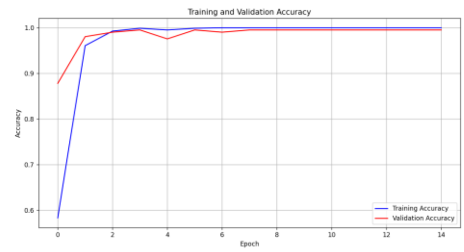

Fig. 11. Accuracy progression of the RNN model.

The model's loss metrics displayed a similar encouraging trend. Beginning at 1.5642, the loss steadily declined, reaching an impressive low of 0.0046 by the end of training. A visual representation of the loss trends throughout the epochs is presented in Figure 12. These graphical representations offer a clear visualization of the model's learning dynamics and its ability to minimize loss while maximizing accuracy over successive epochs, indicating its robustness in learning patterns within the data.

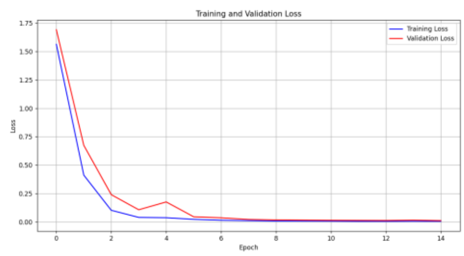

Fig. 12. Loss progression of the RNN model.

The Confusion Matrix model performance, encapsulated in Figure 13 and Figure 14, revealed highly promising evaluation metrics for the classification model. With an accuracy score of 0.9951, the model showcased remarkable precision, recall, and F1 scores across diverse classes. Precision ranged from 0.9500 to 1.0000, signifying the model's accuracy in identifying positive samples, while recall consistently ranged from 0.9231 to 1.0000, demonstrating its reliability in capturing positive samples. The F1 score, representing the harmonic mean of precision and recall, impressively ranges from 0.9600 to 1.0000, indicating a balanced and effective performance of the model in maintaining both precision and recall across different classes.

	precision	recall	f1-score	support
bad	0.95000	1.00000	0.97436	19
bye	1.00000	0.92308	0.96000	13
good	1.00000	1.00000	1.00000	23
hello	1.00000	1.00000	1.00000	22
help	1.00000	1.00000	1.00000	22
me	1.00000	1.00000	1.00000	18
sorry	1.00000	1.00000	1.00000	20
thanks	1.00000	1.00000	1.00000	22
tired	1.00000	1.00000	1.00000	17
you	1.00000	1.00000	1.00000	29
accuracy			0.99512	205

Fig. 13. Classification model performance on validation dataset.

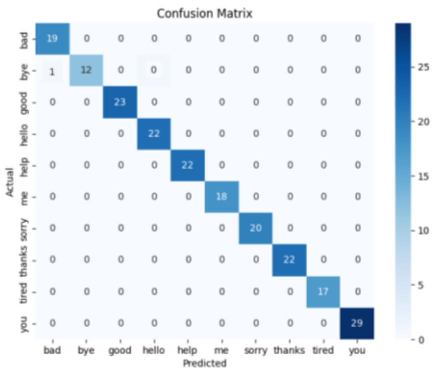

Fig. 14. Confusion Matrix heatmap visualization on the validation dataset.

The findings from this research indicate a robust and highly accurate classification model for interpreting hand sign gestures. These results hold significant implications for real-world applications, particularly in systems involving hand gesture recognition, virtual reality interfaces, or human-computer interaction. The model's exceptional accuracy, reliability in identifying specific gestures, and robustness in capturing temporal dependencies make it a valuable asset in various domains requiring accurate hand gesture interpretations.

The misclassification between 'bad' and 'bye' could occur due to similarities in the hand sign gestures for these two words. Despite efforts to differentiate them, subtle nuances or variations in hand movements may not be adequately captured by the model, leading to misclassification. Additionally, variations in individual hand gestures, differences in recording conditions, or limitations in the training data could also contribute to misclassification errors. Further refinement of the model and additional training data focusing on distinguishing between these gestures may help mitigate this issue.

Moreover, the comprehensive assessment of the model's classification performance is facilitated by the utilization of the confusion matrix shown in Figure 15 and heatmap shown in Figure 16, derived from the testing dataset comprising 100 new hand signs data. These visualizations offer detailed insights into the model's predictive accuracy across various hand sign categories, enabling a thorough analysis of its classification efficacy.

	precision	recall	f1-score	support
bad	1.00000	1.00000	1.00000	10
bye	1.00000	0.90000	0.94737	10
good	1.00000	1.00000	1.00000	10
hello	1.00000	1.00000	1.00000	10
help	1.00000	1.00000	1.00000	10
me	1.00000	0.70000	0.82353	10
sorry	1.00000	1.00000	1.00000	10
thanks	0.71429	1.00000	0.83333	10
tired	1.00000	1.00000	1.00000	10
you	1.00000	1.00000	1.00000	10
accuracy			0.96000	100

Fig. 15. Classification model performance on testing (new) dataset.

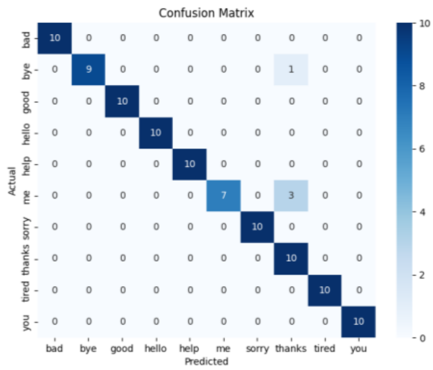

Fig. 16. Confusion Matrix heatmap visualization on the testing (new) dataset.

Based on the testing dataset results, the model demonstrates a commendable overall accuracy of 0.9600. However, it is noteworthy that the precision for the "thanks" class is at 0.7143, indicating a proportion of instances classified as "thanks" that are indeed "thanks" out of all instances classified as "thanks." Furthermore, the recall for the "bye" class stands impressively high at 0.9000, suggesting the proportion of actual "bye" instances that were correctly classified. Similarly, the recall for the "me" class is at 0.7000, reflecting the model's ability to identify a considerable portion of true "me" instances. Additionally, the F1-score, ranging between 0.8235 and 0.1000, offers a harmonic mean of precision and recall, providing a comprehensive evaluation of the model's performance across different classes.

4. Conclusion

The research embarked on an extensive investigation focusing on hand sign gesture recognition utilizing VR technology. Meticulous data collection, alignment adjustments with the headset, and sophisticated data processing techniques culminated in the construction and training of a powerful RNN model. This model exhibited exceptional performance, achieving a perfect accuracy score of 1.0000 by the 7th epoch and maintaining high accuracy throughout the 15 epochs of training. Validation accuracy stabilized impressively at 0.9951, demonstrating the model's consistent performance on new, unseen data. The model showcased a robust learning curve, minimizing loss to an impressive 0.0046 by the end of the training process. Additionally, the model's evaluation metrics—precision, recall, and F1 scores—across various classes, indicated its reliability and effectiveness in accurately identifying hand sign gestures.

Despite the remarkable performance exhibited by the model, a few considerations for future research stand out. First and foremost, overfitting remains a concern, especially with complex models like the RNN. Future investigations should focus on implementing enhanced regularization techniques or exploring alternative model architectures to mitigate overfitting while maintaining high accuracy. Furthermore, the potential for expanding the dataset through crowd data collection methods could be a promising avenue for research. Incorporating a diverse range of user-generated data might further enhance the model's generalizability and robustness in interpreting hand sign gestures across various demographics and contexts. Addressing these areas in future research endeavors could lead to more refined models capable of accurate hand gesture recognition without succumbing to overfitting, thereby improving their applicability and reliability in real-world settings.

Acknowledgement

We sincerely thank Universitas Budi Luhur for their unwavering support and the enriching opportunities provided during this research. The collaborative spirit among faculty members, students, and staff has significantly contributed to the success of our study.

References

[1] L. Hou, "Development of Virtual Reality and Computer Technology Application," in *Lecture Notes in Electrical Engineering*, vol. 827 LNEE, 2022, pp. 196–204. doi: 10.1007/978-981-16-8052-6_22.

[2] S. D. Sulistyaningrum, R. S. Putri, A. Herawati, and S. Irianto, "Trends of virtual reality for learning empirical evidence from different fields," *Journal of Education and Learning (EduLearn)*, vol. 16, no. 4, pp. 531–541, Nov. 2022, doi: 10.11591/edulearn.v16i4.20648.

[3] L. Muñoz-Saavedra, L. Miró-Amarante, and M. Domínguez-Morales, "Augmented and virtual reality evolution and future tendency," *Applied Sciences (Switzerland)*, vol. 10, no. 1, 2020, doi: 10.3390/app10010322.

[4] "Meta Quest 2: Immersive All-In-One VR Headset | Meta Store | Meta Store." Accessed: Oct. 28, 2023. [Online]. Available: https://www.meta.com/quest/products/quest-2/#overview

[5] M. Alizadeh and N. Cowie, "Self-directed learning using VR," *Pacific Journal of Technology Enhanced Learning*, vol. 4, no. 1, pp. 10–11, Feb. 2022, doi: 10.24135/pjtel.v4i1.130.

[6] H. R. Spero, I. Vazquez-Lopez, K. Miller, R. Joshaghani, S. Cutchin, and J. Enterkine, "Drones, virtual reality, and modeling: communicating catastrophic dam failure," *Int J Digit Earth*, vol. 15, no. 1, pp. 585–605, Dec. 2022, doi: 10.1080/17538947.2022.2041116.

[7] N. Albertini, J. Baldini, A. Dal Pino, F. Lazzari, S. Legnaioli, and V. Barone, "PROTEUS: an immersive tool for exploring the world of cultural heritage across space and time scales," *Herit Sci*, vol. 10, no. 1, 2022, doi: 10.1186/s40494-022-00708-3.

[8] A. Martí-Testón, A. Muñoz, L. Gracia, and J. E. Solanes, "Using WebXR Metaverse Platforms to Create Touristic Services and Cultural Promotion," *Applied Sciences (Switzerland)*, vol. 13, no. 14, 2023, doi: 10.3390/app13148544.

[9] A. Syukur, P. N. Andono, K. Hastuti, and A. M. Syarif, "Immersive and Challenging Experiences through A Virtual Reality Musical Instruments Game: An Approach to Gamelan Preservation," *Journal of Metaverse*, vol. 3, no. 1, 2023, doi: 10.57019/jmv.1172928.

[10] A. Liebermann, I. Lente, K. C. Huth, and K. Erdelt, "Impact of a virtual prosthetic case planning environment on perceived immersion, cognitive load, authenticity and learning motivation in dental students," *European Journal of Dental Education*, Apr. 2023, doi: 10.1111/eje.12910.

[11] A. Liebermann, J. K. Seefelder, K. C. Huth, and K. Erdelt, "Mobile virtual tooth morphology teaching environment for preclinical dental students," *J Dent Educ*, vol. 87, no. 1, pp. 130–138, Jan. 2023, doi: 10.1002/jdd.13098.

[12] G. E. Lysenko, A. V. Shchegolev, B. N. Bogomolov, and D. P. Meshakov, "Virtual reality therapy in patients after trauma surgery: prospective clinical trial," *Annals of Critical Care*, vol. 2023, no. 3, 2023, doi: 10.21320/1818-474X-2023-3-58-65.

[13] C. P. Kamm, R. Blättler, R. Kueng, and T. Vanbellingen, "Feasibility and usability of a new home-based immersive virtual reality headset-based dexterity training in multiple sclerosis," *Mult Scler Relat Disord*, vol. 71, p. 104525, Mar. 2023, doi: 10.1016/j.msard.2023.104525.

[14] C. P. Kamm, R. Kueng, and R. Blättler, "Development of a new immersive virtual reality (VR) headset-based dexterity training for patients with multiple sclerosis: Clinical and technical aspects," *Technology and Health Care*, pp. 1–12, Aug. 2023, doi: 10.3233/THC-230541.

[15] C. Portalés, J. Gimeno, A. Salvador, A. García-Fadrique, and S. Casas-Yrurzum, "Mixed Reality Annotation of Robotic-Assisted Surgery videos with real- time tracking and stereo matching," *Comput Graph*, vol. 110, pp. 125–140, Feb. 2023, doi: 10.1016/j.cag.2022.12.006.

[16] F. Pallavicini *et al.*, "A virtual reality home-based training for the management of stress and anxiety among healthcare workers during the COVID-19 pandemic: study protocol for a randomized controlled trial," *Trials*,

vol. 23, no. 1, 2022, doi: 10.1186/s13063-022-06337-2.

[17] S. O. Ose *et al.*, "Development of a social skills training programme to target social isolation using virtual reality technology in primary mental health care," *J Rehabil Assist Technol Eng*, vol. 10, Jan. 2023, doi: 10.1177/20556683231187545.

[18] J. Varela-Aldás, J. Buele, I. López, and G. Palacios-Navarro, "Influence of Hand Tracking in Immersive Virtual Reality for Memory Assessment," *Int J Environ Res Public Health*, vol. 20, no. 5, p. 4609, Mar. 2023, doi: 10.3390/ijerph20054609.

[19] A. Carnevale *et al.*, "Virtual Reality for Shoulder Rehabilitation: Accuracy Evaluation of Oculus Quest 2," *Sensors*, vol. 22, no. 15, 2022, doi: 10.3390/s22155511.

[20] A. Jaiswal, A. Hebri, H. R. Pavel, M. Z. Zadeh, and F. Makedon, "SmartFunction: An Immersive Vr System To Assess Attention Using Embodied Cognition," in *ACM International Conference Proceeding Series*, 2023. doi: 10.1145/3594806.3596559.

[21] M. Trinidad-Fernández, B. Bossavit, J. Salgado-Fernández, S. Abbate-Chica, A. J. Fernández-Leiva, and A. I. Cuesta-Vargas, "Head-Mounted Display for Clinical Evaluation of Neck Movement Validation with Meta Quest 2," *Sensors*, vol. 23, no. 6, p. 3077, Mar. 2023, doi: 10.3390/s23063077.

[22] M. Ostanin, S. Zaitsev, A. Sabirova, and A. Klimchik, "Interactive Industrial Robot Programming based on Mixed Reality and Full Hand Tracking," *IFAC-PapersOnLine*, vol. 55, no. 10, pp. 2791–2796, 2022, doi: 10.1016/j.ifacol.2022.10.153.

[23] S. Rostami and M. Maier, "The Metaverse and Beyond: Implementing Advanced Multiverse Realms With Smart Wearables," *IEEE Access*, vol. 10, 2022, doi: 10.1109/ACCESS.2022.3215736.

[24] D. Yoon and A. Oh, "Design of Metaverse for Two-Way Video Conferencing Platform Based on Virtual Reality," *Journal of information and communication convergence engineering*, vol. 20, no. 3, pp. 189–194, Sep. 2022, doi: 10.56977/jicce.2022.20.3.189.

[25] D. J. Kim and W. Li, "A View Direction-Driven Approach for Automatic Room Mapping in Mixed Reality," in *Proceedings of the 2023 5th International Conference on Image Processing and Machine Vision*, New York, NY, USA: ACM, Jan. 2023, pp. 29–33. doi: 10.1145/3582177.3582183.

[26] D. Pereira, V. Oliveira, J. L. Vilaça, V. Carvalho, and D. Duque, "Measuring the Precision of the Oculus Quest 2's Handheld Controllers," *Actuators*, vol. 12, no. 6, 2023, doi: 10.3390/act12060257.

[27] D. Abdlkarim *et al.*, "A methodological framework to assess the accuracy of virtual reality hand-tracking systems: A case study with the Meta Quest 2," *Behavior Research Methods*. 2023. doi: 10.3758/s13428-022-02051-8.

[28] A. S. Fernandes, T. S. Murdison, and M. J. Proulx, "Leveling the Playing Field: A Comparative Reevaluation of Unmodified Eye Tracking as an Input and Interaction Modality for VR," *IEEE Trans Vis Comput Graph*, vol. 29, no. 5, pp. 2269–2279, May 2023, doi: 10.1109/TVCG.2023.3247058.

[29] P. Kudry and M. Cohen, "Enhanced Wearable Force-Feedback Mechanism for Free-Range Haptic Experience Extended by Pass-Through Mixed Reality," *Electronics (Basel)*, vol. 12, no. 17, p. 3659, Aug. 2023, doi: 10.3390/electronics12173659.

[30] P. Kudry and M. Cohen, "Development of a wearable force-feedback mechanism for free-range haptic immersive experience," *Front Virtual Real*, vol. 3, Dec. 2022, doi: 10.3389/frvir.2022.824886.

[31] S. Ramos-cosi, L. Cardenás-Pineda, D. Llulluy-Nuñez, and A. Alva-Mantari, "Development of 3D Avatars for Inclusive Metaverse: Impact on Student Identity and Satisfaction using Agile Methodology, VRChat Platform, and Oculus Quest 2," *International Journal of Engineering Trends and Technology*, vol. 71, no. 7, pp. 1–14, Jul. 2023, doi: 10.14445/22315381/IJETT-V71I7P201.

[32] K. Sangeethalakshmi, K. G. Shanthi, A. Mohan Raj, S. Muthuselvan, P. Mohammad Taha, and S. Mohammed Shoaib, "Hand gesture vocalizer for deaf and dumb people," *Mater Today Proc*, vol. 80, 2023, doi: 10.1016/j.matpr.2021.07.324.

[33] A. Vaitkevičius, M. Taroza, T. Blažauskas, R. Damaševičius, R. Maskeliunas, and M. Woźniak, "Recognition of American Sign Language gestures in a Virtual Reality using Leap Motion," *Applied Sciences (Switzerland)*, vol. 9, no. 3, 2019, doi: 10.3390/app9030445.

[34] "Meta Quest 2: Our Most Advanced New All-in-One VR Headset | Oculus | Meta Store." Accessed: Oct. 28, 2023. [Online]. Available: https://www.meta.com/quest/products/quest-2/tech-specs/#tech-specs

[35] "'you' ASL American Sign Language." Accessed: Nov. 03, 2023. [Online]. Available: https://www.lifeprint.com/asl101/pages-signs/y/you.htm

[36] "'help' American Sign Language (ASL)." Accessed: Nov. 03, 2023. [Online]. Available: https://www.lifeprint.com/asl101/pages-signs/h/help.htm

[37] "Top 10 & 25 American Sign Language Signs for Beginners – The Most Known Top 10 & 25 ASL Signs to Learn First | Start ASL." Accessed: Nov. 03, 2023. [Online]. Available: https://www.startasl.com/top-10-25-american-sign-language-signs-for-beginners-the-most-know-top-10-25-asl-signs-to-learn-first/

[38] "'me' American Sign Language (ASL)." Accessed: Nov. 03, 2023. [Online]. Available: https://www.lifeprint.com/asl101/pages-signs/m/me.htm

[39] "Unity - Manual: Unity User Manual 2022.3 (LTS)." Accessed: Nov. 03, 2023. [Online]. Available: https://docs.unity3d.com/Manual/index.html

[40] "Padding in Neural Networks: Why and How? | by Ashish Sanjay Raut | GoPenAI." Accessed: Nov. 03, 2023. [Online]. Available: https://blog.gopenai.com/padding-in-neural-networks-why-and-how-b076ab0a4fc2

[41] "What is Padding in Neural Network? - GeeksforGeeks." Accessed: Nov. 03, 2023. [Online]. Available: https://www.geeksforgeeks.org/what-is-padding-in-neural-network/

[42] "CS 230 - Recurrent Neural Networks Cheatsheet." Accessed: Nov. 03, 2023. [Online]. Available: https://stanford.edu/~shervine/teaching/cs-230/cheatsheet-recurrent-neural-networks

[43] "What are Recurrent Neural Networks? | IBM." Accessed: Nov. 03, 2023. [Online]. Available: https://www.ibm.com/topics/recurrent-neural-networks

[44] "Unity - Scripting API: CommonUsages." Accessed: Nov. 03, 2023. [Online]. Available: https://docs.unity3d.com/ScriptReference/XR.CommonUsages.html