MYU Tokyo

S & M 4156

AI-driven Augmentative and Alternative Communication System for Communication Enhancement in Cerebral Palsy Using American Sign Language Recognition

Yuh-Shihng Chang, Chen-Wei Lin, and Chien-Chih Chen*

Department of Information Management, National Chin-Yi University of Technology, Taiwan, ROC

(Received April 28, 2025; accepted July 18, 2025)

Keywords: augmentative and alternative communication, cerebral palsy, artificial intelligence, machine learning, American Sign Language

Augmentative and alternative communication (AAC) methods are designed to facilitate effective communication for individuals with speech, language, or writing impairments. We introduce a novel AAC system designed to enhance communication for individuals with speech, language, or writing impairments, such as those with cerebral palsy (CP). Addressing the limitations of existing AI-driven AAC solutions, advanced AI and machine learning techniques are integrated in the system to minimize user effort in conveying their thoughts. Utilizing a readily available webcam as a key input modality, the system employs MediaPipe to capture hand gestures corresponding to American Sign Language signs. The resulting visual data is then processed and classified by a random forest model. By interpreting these sensor-captured gestures, the system enables users to input partial vocabulary, which subsequently prompts generative AI models to predict and complete intended text. Empirical evaluations, conducted through two distinct experiments, validate the system's viability and demonstrate its potential to significantly improve communication accessibility for individuals with CP through an accessible and intuitive gesture-based interface.

1. Introduction

Augmentative and alternative communication (AAC) methods aim to facilitate information exchange for individuals with speech, language, or writing disabilities, either by enhancing speech communication or providing alternative written communication.(1) Researchers are continuously developing new AAC technologies, including sign languages(2) and communication boards,(3) to improve communication effectiveness. In recent years, the integration of AI has led to the development of AI-based AAC methods. However, personalized AAC solutions remain a challenge.

AAC systems are broadly classified into unaided and aided categories. Unaided systems utilize the user's body language, natural gestures, manual signs,(4) and facial expressions,

*Corresponding author: e-mail: frick@ms14.hinet.net
https://doi.org/10.18494/SAM5710

ISSN 0914-4935 © MYU K.K.
https://myukk.org/

requiring no external devices. Their effectiveness depends on the receiver's familiarity with the specific system. Sign languages, a highly structured form of unaided AAC, are considered among the most developed. In contrast, aided AAC systems introduce external devices to facilitate communication. Traditional examples include communication boards featuring pictures, symbols, or words,(5) letter boards, and writing tools.

Over the past decade, AI techniques have revolutionized the AAC research domain. Individuals with communication disabilities can now utilize motion recognition applications to select symbols or words(5) for speech-generating devices, enabling synthetic speech. These technologies commonly incorporate features such as text prediction, vocabulary customization, and personalized voice options.

Individuals with cerebral palsy (CP), who experience neuromuscular disorders affecting movement and posture, often face challenges with speech production and fine motor control. Motor impairments can hinder their ability to control their hands, arms, or mouth, necessitating innovative approaches to access communication tools. Fortunately, the field of AAC offers a variety of access methods that have proven particularly beneficial for the CP community.(6)

For individuals with CP using AAC, both access methods and naming systems require careful consideration. Visual-cognitive challenges associated with CP can impede symbol recognition, necessitating the selection of symbol systems based on maximum intelligence and equivalence principles. Suitable options for individuals with moderate literacy include photographs, realistic color pictures, simplified pictographic symbols, and text.(5)

Direct access to technology remains a significant challenge for individuals with CP. Muscle tone fluctuations, involuntary movements, and difficulties in maintaining device positioning can hinder consistent equipment operation. These motor challenges, which vary throughout the day and with emotional states, necessitate dynamically adaptable access systems. Some users struggle with targeting and switch activation, while others experience fatigue during prolonged communication attempts, leading to communication endurance issues.

Dasher,(7) a data entry interface combining continuous gestures and language models to assist motion-impaired computer users, was introduced in early research. Dasher facilitates continuous text generation with minimal physical effort, reducing the precision required for effective communication. The rapid progress in computer vision and deep learning has facilitated the emergence of sophisticated human pose estimation (HPE) algorithms, with OpenPose and MediaPipe being prominent examples. For instance, Rang *et al.*(8) demonstrated the application of the MediaPipe Gesture Recognizer(9) in creating a hand gesture recognition system. Consequently, the development of effective communication tools for individuals with CP has become increasingly feasible. Many developers of mobile phones, laptops, and other communication devices have integrated large language models (LLMs) based on generative pretrained transformer (GPT) architectures to significantly enhance word prediction. Devices previously relying solely on frequency and recency now leverage semantic context to provide users with more accurate predictions, including phrases and responses, thereby reducing the number of required selections.(10)

To address communication challenges faced by individuals with CP, in this paper, we present a practical AI-based AAC system. The system employs a three-stage process: handshape recognition via MediaPipe Gesture Recognizer[9] and random forest[11] classification, text prediction using generative AI (GAI) to generate the top five sentence completions, and speech synthesis utilizing the pyttsx3 Python library. Empirical evaluations, conducted through two distinct experiments, validate the system's effectiveness and its potential to significantly improve communication access for individuals with CP.

The subsequent sections of this paper are structured as follows: Section 2 is an outline of the foundational elements of the proposed method, Sect. 3 provides a detailed explanation of the method itself, the empirical evaluation is presented in Sect. 4, and Sect. 5 concludes the paper with a summary of findings.

2. Related Work

In this section, we outline the foundational elements of our method: the American Sign Language (ASL) alphabet, MediaPipe's Gesture Recognizer, and GAI.

2.1 ASL alphabet

While ASL utilizes a manual alphabet, where each letter of the English alphabet is represented by a specific handshape, it is essential to understand that this alphabet is not the core of ASL. ASL is not simply English spelled out with hand gestures. Instead, it is a fully developed, independent language with its own grammar and syntax, distinct from English, as shown in Fig. 1.

The ASL manual alphabet, sometimes referred to as fingerspelling, is primarily used for specific purposes, such as spelling proper nouns, clarifying words with no established sign, and

Fig. 1. (Color online) The American Sign Language manual alphabet.[12]

spelling technical terms or loan words. It is important to note that the sequential nature of fingerspelling contrasts sharply with the simultaneous expression of meaning characteristic of most ASL signs. ASL signs convey concepts through a combination of handshape, location, movement, palm orientation, and nonmanual markers (facial expressions and body language). This multifaceted approach to communication highlights the complexity and richness of ASL beyond simple letter-to-hand correspondence. While the manual alphabet serves as a valuable tool within ASL, it represents only a small component of the language's overall structure and usage. ASL's independence stems from its unique linguistic framework, which is rooted in visual-spatial communication and historical development.

2.2 MediaPipe Gesture Recognizer

Developed by Google, the MediaPipe Gesture Recognizer utilizes the MediaPipe framework for perception tasks, uniquely focusing on interpreting hand movement language through multistep processing that identifies 21 specific key landmarks on the hand, mirroring skeletal points and tracking finger joints, knuckles, and palm movements, to classify gestures such as thumbs up or open palm.(8) This technology, as illustrated by the hand landmarks in Fig. 2, offers significant potential for accessibility by enabling real-time sign language translation, fostering inclusive communication between hearing and hearing-impaired individuals in the digital realm. In comparison with other HPE algorithms such as OpenPose, MediaPipe Gesture Recognizer excels in real-time performance and mobile-friendly deployment owing to its optimized models and lightweight architecture, providing a distinct advantage in competitions requiring fast inference speeds. However, while OpenPose often delivers higher accuracy and more detailed body pose estimations, MediaPipe Gesture Recognizer's focus on hand-specific gestures might limit its applicability in broader HPE scenarios that demand comprehensive full-body tracking.

2.3 Generative artificial intelligence

The evolution of GAI, as shown in Fig. 3, is a narrative of progressive innovation, commencing with the foundational Early AI & ML (machine learning) Research that laid the groundwork for subsequent advancements. Initially, AI development was characterized by

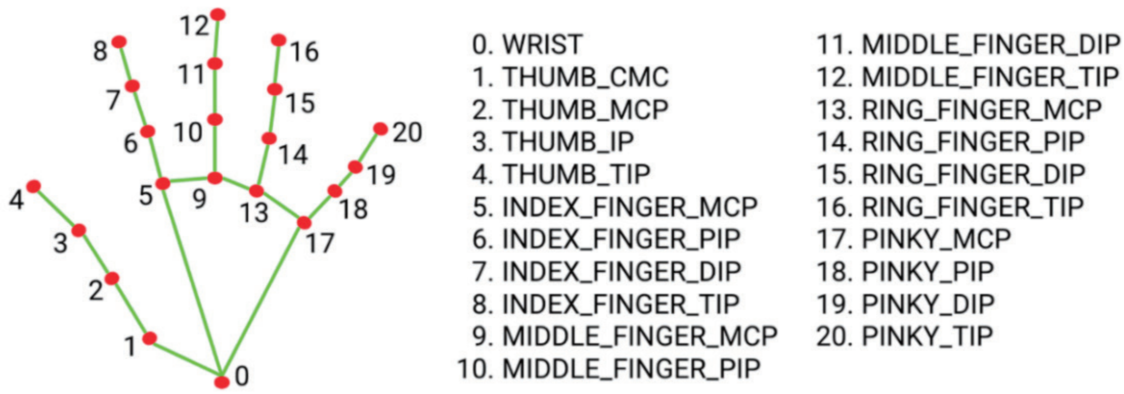

Fig. 2. (Color online) Hand keypoints defined in MediaPipe.(9)

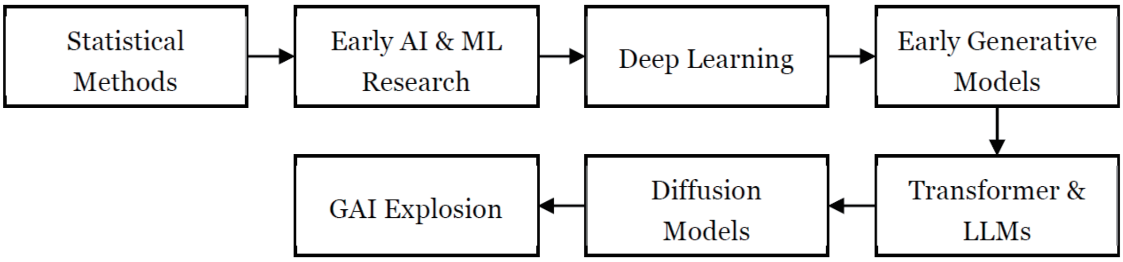

Fig. 3. Evolution of GAI.

statistical methods, which employed probabilistic models and rule-based systems to analyze and interpret data. This phase transitioned into the era of deep learning, marked by the advent of neural networks capable of learning intricate patterns from vast datasets, thereby enabling significant improvements in tasks such as image recognition and natural language processing. The emergence of early generative models, such as variational autoencoders and generative adversarial networks, represented a pivotal shift that enabled machines to generate novel content. For example, GAN has been utilized to solve the class-imbalanced learning issue.(13) However, the paradigm truly transformed with the introduction of transformer architectures and LLMs, exemplified by models such as GPT, which utilized attention mechanisms to produce coherent and contextually relevant text. Concurrently, diffusion models, as seen in systems such as DALL-E and Stable Diffusion, revolutionized image generation by iteratively refining random noise into detailed visuals. These technological breakthroughs culminated in a GAI explosion, a period of rapid proliferation and adoption of GAI across diverse applications. This trajectory, from the nascent stages of AI research to the current era of widespread GAI utilization, underscores the dynamic and transformative nature of AI, highlighting its increasing capacity to create content that mirrors human creativity and understanding.(14)

3. Proposed Method

The proposed system, illustrated in Fig. 4, comprises three distinct stages: handshape recognition, text prediction, and speech synthesis. During handshape recognition, MediaPipe is utilized to extract hand landmark coordinates, which are subsequently normalized and used to train a random forest classifier for letter identification. In the text prediction stage, a GAI model, such as Gemini, ChatGPT, or Claude, is employed to generate complete sentence hypotheses on the basis of the recognized letter sequence. The top five most probable sentence completions are then presented to the user. Finally, in the speech synthesis stage, the user-selected sentence is converted into audible speech using a Python text-to-speech (TTS) library such as pyttsx3 or gTTS.

3.1 Handshape recognition

Handshape recognition, as depicted in Fig. 5, begins with users forming ASL alphabet hand gestures. A webcam captures these gestures, and MediaPipe Gesture Recognizer extracts the

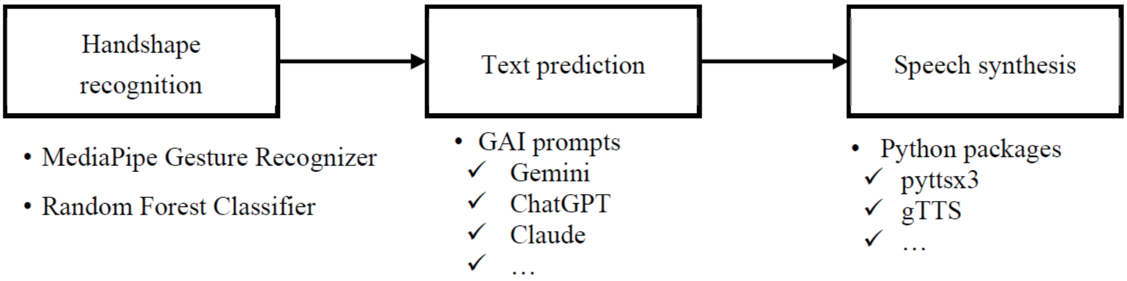

Fig. 4. Main stages of the proposed method.

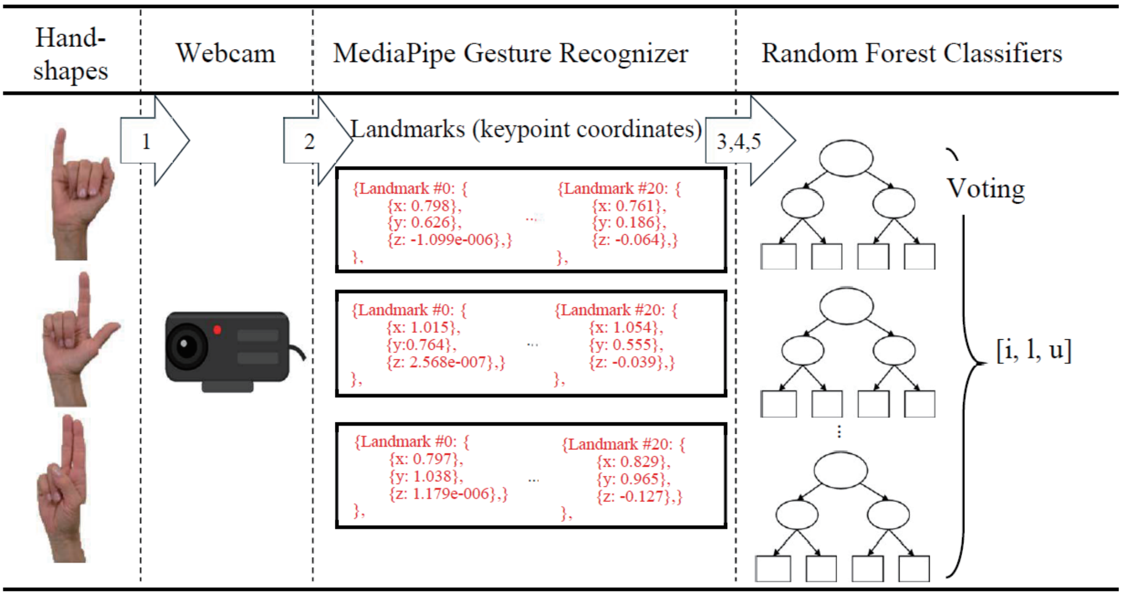

Fig. 5. (Color online) Processes in handshape recognition.

hand keypoint coordinates defined in Fig. 2. These coordinates serve as input for our pretrained random forest classifiers. Specifically, the process is as follows.

1. Gesture Capture: The webcam records the user's ASL hand gesture.
2. Keypoint Extraction: MediaPipe Gesture Recognizer outputs the coordinates of hand keypoints as a dictionary object.
3. Normalization: A Min–Max normalization transforms these coordinates into the range (0, 1):

$$x' = \frac{x - min}{max - min}, \tag{1}$$

where x' is the normalized value, x is the original coordinate, and min and max are the minimum and maximum values of the original coordinates output by MediaPipe Gesture Recognizer, respectively.

4. Classification: The normalized keypoint coordinates are then fed into the random forest classifiers.
5. Alphabet Output: The classifiers output the corresponding ASL letter.

For example, when users form the gestures for “i”, “l”, and “u” the respective keypoint coordinates are extracted, normalized, and used to accurately identify those letters.

3.2 Text prediction

A key aspect of LLM training involves masked language models (MLMs), which facilitate the development of LLM’s understanding of human cognition. In this approach, specific words or tokens within an input sequence are randomly masked. The LLM is then trained to predict these masked elements by utilizing the contextual information provided by the surrounding words. Consequently, LLMs such as ChatGPT, Gemini, and Claude can effectively infer user intent even when presented with incomplete or partial input. In this study, we utilize this capability, assuming robust MLM training, to explore the ability of LLMs to complete text from partial vocabulary or abbreviations. For instance, prompting Gemini with “Complete the top 5 possible sentences with this schema: {‘Sentence’: str} but without any explanation of whether the sentence ‘i l u’ is incomplete” yields the response shown in Fig. 6. Additionally, language translation is provided as the response shown in Fig. 7.

3.3 Speech synthesis

Depending on user needs, text-to-speech (TTS) conversion can be incorporated. Given the established maturity of TTS technology, numerous Python packages are available, such as

Fig. 6. (Color online) Gemini's output for 'i l u' sentence completion.

Fig. 7. (Color online) Gemini's output for “i l u” sentence completion in traditional Chinese.

pyttsx3 and gTTS. pyttsx3 was selected for this study because of its robust offline capabilities, cross-platform support, and straightforward operability.

4. Empirical Evaluation

In this section, we will first implement the handshape recognition and then the text prediction.

4.1 Handshape recognition

Handshape recognition involves two stages. First, MediaPipe, a Python package, extracts hand landmarks (coordinates) from webcam input. Second, a pretrained classifier identifies letters based on these coordinates. We utilize a random forest classifier[11] trained on data from Kaggle.[15] Random forests were chosen for their ability to integrate multiple weak classifiers trained on diverse datasets, mitigating issues such as overfitting and bias.

Our experimental evaluation involved individuals with CP and a control group without CP. To assess our classifier's performance, we enlisted one author with CP and ten student volunteers. Participants were instructed to hold each hand gesture for one second after ceasing movement, a condition under which all gestures could theoretically be correctly identified. The results, summarized in Table 1, indicate a recognition accuracy of 54.615% for individuals with CP and 94.808% for the control group. Lower accuracy for certain static gestures such as "A", "M", "N", "S", and "T" may be attributed to their visual similarity. Furthermore, dynamic gestures such as "J" and "Z" presented additional recognition challenges due to their inherent motion.

This initial outcome, however, did not meet our objective of enhancing communication effectiveness for individuals with CP. To address this, we collected additional training data by recording more videos from our author with CP, capturing 100 hand coordinate sets per letter using MediaPipe. After retraining our classifier with this augmented dataset, the results, presented in Table 2, showed an improved recognition accuracy of 75.769% for individuals with CP and a slightly decreased accuracy of 91.115% for the control group.

Table 1
Experimental results of our random forest classifier.

Letter	Successes (CP)	Successes (No CP)	Letter	Successes (CP)	Successes (No CP)
A	9	91	N	9	84
B	3	95	O	9	98
C	1	98	P	7	98
D	5	99	Q	8	99
E	8	98	R	5	98
F	7	99	S	8	96
G	6	96	T	5	90
H	2	96	U	6	99
I	3	96	V	7	96
J	1	88	W	6	97
K	5	90	X	4	98
L	8	96	Y	2	99
M	6	85	Z	2	86
Total	64	1227	Total	78	1238

Table 2
Experimental results of our retrained classifier.

Letter	Successes (CP)	Successes (No CP)	Letter	Successes (CP)	Successes (No CP)
A	10	82	N	9	86
B	6	86	O	9	97
C	6	90	P	7	98
D	8	89	Q	8	97
E	8	98	R	8	94
F	9	99	S	8	92
G	10	90	T	8	90
H	8	88	U	8	92
I	8	83	V	7	95
J	7	88	W	6	98
K	8	88	X	4	98
L	10	96	Y	5	89
M	6	83	Z	6	83
Total	104	1160	Total	93	1209

The handshape recognition results reveal a fundamental trade-off inherent in adapting machine learning models for individuals with CP. The initial classifier achieved 54.615% accuracy for CP users versus 94.808% for the control group, representing a substantial 40.193% gap. After retraining with CP-specific training data, we observed a significant improvement for CP users (75.769%) but a corresponding decrease for the control group (91.115%). This trade-off demonstrates that incorporating CP-specific training data creates a model specialization effect where gains for the target population come at the cost of general performance.

While the 21.154% improvement for CP users represents meaningful progress, the final performance gap (15.346%) still indicates that the system performs substantially worse for its intended users compared with the general population. This persistent disparity suggests that the motor variability and movement patterns associated with CP present fundamental challenges that cannot be fully addressed through dataset augmentation alone.

The letter-specific analysis reveals both encouraging improvements and persistent limitations. Letters such as "C", "G", "H", and "J" and "Z" showed dramatic improvements for CP users (from 10 to 60% for "C", 20 to 80% for "H", and 10 to 70% for "J"), demonstrating that targeted training data can address specific recognition challenges. However, several fundamental letters such as "B", "I", "V", and "Y" remain problematic with accuracy rates below 80%. The fact that common letters continue to show poor recognition performance suggests that the underlying feature representation may be insufficient for capturing the full spectrum of motor variations present in CP-affected sign production.

4.2 Text prediction

To evaluate the text prediction capability of GAI models, we propose an approach called partial word masking (PWM). Table 3 illustrates the PWM rate calculation with four examples. Consider the first example: for the sentence "I agree", we tested various letter combinations. We

Table 3
Examples of computing partial word masking rates.

Original sentence	Success conditions	Word masked rate	Sentence masked rate
I agree	I ag	I: (1 − 1)/1=0 agree: (5 − 2)/5=0.6	Max(0, 0.6) = 0.6
Not yet	Not yet	Not: (1 − 1)/1=0 yet: (5 − 2)/5=0	Max(0, 0) = 0
See you	See y	See: (1 − 1)/1=0 you: (5 − 2)/5=0.67	Max(0, 0.67) = 0.67
I didn't mean it	I din m i	I: (1 − 1)/1=0 didn't: (6 − 3) / 6=0.5 mean: (4 − 1)/4 = 0.75 it: (2 − 1)/2=0.5	Max(0, 0.5, 0.75, 0.5) = 0.75

observed that "I ag" was the shortest input that resulted in "I agree" being among the top five predictions by Gemini 2.0 Flash. The PWM rate for each word is calculated as the proportion of masked letters. For "I", the rate is (1 − 1)/1 = 0, as no letters were masked. For "agree", with five letters and two given ("ag"), the masked length is (5 − 2) = 3, resulting in a PWM rate of 3/5 = 0.6. The PWM rate for the entire sentence "I agree" is then defined as the maximum of the individual word PWM rates: Max(0, 0.6) = 0.6.

In this experiment, we evaluated Gemini 2.0 Flash using the first one hundred sentences from the "English for Life" quiz, sourced from https://blog.csdn.net/Wit_tang/article/details/51036733. Table 4 presents the experimental results, detailing the frequency of "hit events" instances where Gemini 2.0 Flash accurately predicted the original sentence when presented with partially masked words (PMW). The results reveal a clear trend in Gemini 2.0 Flash's predictive performance. Most notably, approximately half of the sentences (49%) were accurately inferred when the PMW rate fell within the (60%, 80%) range. This indicates a significant capability to reconstruct original text even with a substantial portion of words masked, suggesting the model utilizes contextual understanding effectively. Performance dropped considerably at lower PMW rates, with only 31% of sentences correctly predicted when masking was less than 60% [21% from (40%, 60%) and 10% from (0%, 40%)]. The data further reveals that while the model excels with moderate masking, its accuracy diminishes sharply when either minimal information (PMW rates below 40%) or excessive masking (PMW rates above 80%) is provided. The modest 20% hit rate in the (80%, 100%) range, where nearly all words are masked, underscores the challenge of high-ambiguity predictions.

These findings indicate that GAI models such as Gemini 2.0 Flash hold substantial promise for reducing user input demands in the proposed AAC system. Specifically, the high accuracy observed in the [60%, 80%] PMW range implies that users can potentially provide only a few keywords or partial words, and the system can accurately complete their intended message. This "sweet spot" of moderate masking offers a compelling pathway to significantly increase communication efficiency and reduce physical effort for AAC users.

However, a critical examination of these results also reveals limitations that warrant further consideration. While the model performs exceptionally well in the optimal PMW range, its performance at the extremes (very low or very high masking) indicates areas for improvement.

Table 4
PWM rates and their accuracy events.

PWM rate range	Hit events
[0%, 20%)	5
[20%, 40%)	5
[40%, 60%)	21
[60%, 80%)	49
[80%, 100%]	20
Total	100

The limited accuracy below 50% PMW suggests that the model might not be fully utilizing the available unmasked information in these scenarios, potentially due to an over-reliance on a "guessing" mechanism rather than sophisticated contextual reasoning when more complete words are already present. Conversely, the drop in performance at very high PMW rates (e.g., above 80%) highlights the inherent difficulty of disambiguation when extremely limited information is provided, indicating the need for robust error correction mechanisms or user-feedback loops in a practical AAC system. Future work should focus on fine-tuning the model's performance across the entire spectrum of PMW rates, perhaps by incorporating user-specific language models or dynamic masking strategies to adapt to varying levels of user input and linguistic ambiguity. Furthermore, while "hit events" provide a valuable metric, a more comprehensive evaluation should also consider the quality of "misses" (e.g., how close incorrect predictions were to the target sentence) and the computational latency of the model, which is crucial for real-time AAC applications. Addressing these limitations will be key to realizing the full potential of GAI models in clinically viable AAC solutions.

5. Conclusions

The novel AAC system presented in this research offers a promising avenue for enhancing communication accessibility for individuals with severe disabilities, particularly those with CP. By integrating hand gesture recognition powered by MediaPipe and a random forest classifier with the predictive capabilities of GAI models, this system aims to significantly reduce the physical and cognitive demands typically associated with AAC use. Our initial experimental evaluation of hand gesture recognition revealed a notable difference in accuracy between individuals with CP (54.615%) and a control group (94.808%), highlighting the challenges posed by motor impairments. Subsequent efforts to improve recognition for individuals with CP through additional training data yielded a substantial increase in accuracy to 75.769%, albeit with a slight decrease for the control group (91.115%), underscoring the need for tailored models. Furthermore, the partial word masking experiment with Gemini 2.0 Flash demonstrated the potential of GAI models to correctly infer intended sentences even with significant portions of words masked, with approximately 50% accuracy within a 60–80% masking range. This suggests that by providing even fragmented sign input, the proposed AAC system can utilize the predictive power of advanced language models to complete intended messages, thereby lowering the user burden. The combined results of these experiments validate the viability of our integrated approach and its potential to significantly improve communication effectiveness and independence for individuals who rely on AAC. Future work will be focused on refining the hand gesture recognition model to improve robustness and on exploring user-centered design principles to optimize the overall usability and real-world impact of this novel AAC system.

References

1 T. Griffiths, R. Slaughter, and A. Waller: J. Enabl. Technol. **18** (2024) 232. http://dx.doi.org/10.1108/JET-01-2024-0007
2 E. Efeoğlu and A. Tuna: Kirklareli Univ. J. Eng. Sci. **10** (2024) 219. http://dx.doi.org/10.34186/klujes.1546178
3 R. Scherer, M. Billinger, J. Wagner, A. Schwarz, D. T. Hettich, E. Bolinger, M. Lloria Garcia, J. Navarro, and G. Müller-Putz: Ann. Phys. Rehabil. Med. **58** (2015) 14. http://dx.doi.org/https://doi.org/10.1016/j.rehab.2014.11.005
4 D. Kouremenos and K. Ntalianis, Glam-Sign: Greek Language Multimodal Lip Reading with Integrated Sign Language Accessibility 2025). http://dx.doi.org/10.48550/arXiv.2501.05213
5 L. Pope and J. Light: Augment. Altern. Commun. (2025) 1. http://dx.doi.org/10.1080/07434618.2025.2458868
6 Y. Guerrier, S. Bosquet, O. Valadier, N. Cauchois, V. Delcroix, K. M. de Oliveira, and C. Kolski: Proc. HCI Int. 2024 – Late Breaking Papers (2025) 14.
7 D. Ward, A. Blackwell, and D. Mackay: Proc. 13th Annu. ACM Symp. User Interface Softw. Technol. (ACM, 2000) 129. https://doi.org/10.1145/354401.354427
8 C. N. Rang, P. Jerónimo, C. Mora, and S. Jardim: Procedia Comput. Sci. **256** (2025) 198. http://dx.doi.org/https://doi.org/10.1016/j.procs.2025.02.112
9 MediaPipe Gesture Recognizer: Mediapipe Gesture Recognizer https://ai.google.dev/edge/mediapipe/solutions/vision/gesture_recognizer?hl=zh-tw (accessed April 2025).
10 D. J. Bailey, F. Herget, D. Hansen, F. Burton, G. Pitt, T. Harmon, and D. Wingate: Aphasiology (2025) 1. http://dx.doi.org/10.1080/02687038.2024.2445663
11 L. Breiman: Mach. Learn. **45** (2001) 5. https://doi.org/10.1023/A:1010933404324
12 Lingvano: The Alphabet in Asl Sign Language https://www.lingvano.com/asl/blog/sign-language-alphabet/ (accessed April 2025).
13 C.-C. Chen, Y.-S. Lin, and H.-Y. Chen: Sens. Mater. **36** (2024) 4835. http://dx.doi.org/10.18494/SAM5224
14 L. Banh and G. Strobel: Electron. Mark. **33** (2023) 63. http://dx.doi.org/10.1007/s12525-023-00680-1
15 Kaggle: American Sign Language Fingerspelling Recognition https://www.kaggle.com/competitions/asl-fingerspelling/data (accessed April 2025).

About the Authors

Yuh-Shihng Chang is currently with the Department of Information Management at National Chin-Yi University of Technology. He has held the positions of professor and department head since 2024. In recent years, his research has been focused on the application of IoT combined with AI to support elderly care at home and improve the quality of life for sub-healthy individuals. Prior to 2022, much of his work centered on the use of augmented reality (AR) to enhance and innovate teaching methods. His current research interests include AR integration in education, AI applications, and sensor technologies. (eric_chang@ncut.edu.tw)

Chen-Wei Lin received his B.S. degree from National Chin-Yi University of Technology in Taiwan in 2024. He is currently a graduate student in the Department of Information Management at the same university. His research interest focuses on applying artificial intelligence to augmentative and alternative communication. (eexx3939@gmail.com)

Chien-Chih Chen is an assistant professor in the Department of Information Management of National Chin-Yi University of Technology, Taiwan. His current interests are focused on machine learning with small data sets. His articles have appeared in Decision Support Systems, Omega, Automation in Construction, Computers and Industrial Engineering, International Journal of Production Research, Neurocomputing, and other publications.
(frick@ms14.hinet.net)