Industry and Academic Research Review
Volume 3 Issue 1 October 2025
DOI: https://doi.org/10.53378/iarr.188

KamAI: A basic Filipino sign language recognition mobile application using deep learning

[1]Dave Lota, [2]Catherine Bhel Aguila & [3]Dayne Fradejas

Abstract

This research presents the design and development of KamAi, a mobile application that leverages deep learning to recognize basic Filipino Sign Language (FSL). Developed for Android platforms, KamAi aims to address communication challenges faced by the Deaf community by enabling real-time recognition of FSL letters, numbers, and commonly used words. The application integrates Convolutional Neural Networks (CNNs) with Google's MediaPipe for gesture detection, offering an accessible, user-friendly tool for both Deaf and hearing individuals engaged in FSL learning. Employing a developmental research design guided by the Design Thinking framework, the study engaged FSL educators, deaf and mute, hearing individuals and IT experts to ensure functionality, usability, and relevance. A dataset of over 29,000 images, combining captured and publicly available sources, was utilized to train the deep learning models. Evaluation results indicated strong model performance: the letter recognition model achieved an accuracy of 88.38%, precision of 97%, recall of 87%, and F1-score of 90%. The number model yielded 91.41% accuracy and 89% precision, though with a lower recall of 21% and F1-score of 28%, highlighting challenges in distinguishing visually similar signs. The word recognition model attained 83.08% accuracy, 96% precision, 88% recall, and 91% F1-score, demonstrating effective classification of everyday gestures. User Acceptance Testing (UAT) was conducted based on ISO 25010 software quality standards, with all six evaluated attributes, Functionality (4.73), Reliability (4.51), Compatibility (4.71), Usability (4.50), Efficiency (4.59), and Portability (4.77), receiving "Highly Acceptable" ratings. These results confirm KamAi's practical viability as a mobile-based assistive learning tool and its potential contribution to inclusive education and digital accessibility in the Philippine context.

Keywords: *android application, Filipino sign language, deep learning, gesture recognition, design thinking*

Article History:

***Received**: June 24, 2025* ***Revised**: October 18, 2025*
***Accepted**: October 25, 2025* ***Published online**: October 31, 2025*

Suggested Citation:
Lota, D., Aguila, C.B. & Fradejas, D. (2025) KamAI: A basic Filipino sign language recognition mobile application using deep learning. *Industry and Academic Research Review*, 3(1), 92-127. https://doi.org/10.53378/iarr.188

About the authors:
[1]Corresponding author. Master of Information Technology, Romblon State University. Email: dlota71@gmail.com
[2]Doctor of Information Technology, Romblon State University
[3]Master of Information Technology Romblon State University

**This paper is presented at the 6th International Conference on Multidisciplinary Industry and Academic Research (ICMLAR)-2025*

1. Introduction

The Deaf and mute community faces significant barriers in communication, limiting their participation in education, employment, and everyday social interactions. One major factor contributing to these challenges is the limited awareness and understanding of sign language among the general population, along with the lack of widely available assistive technologies to support real-time communication (Jha & Pathak, 2024). In the Philippines, the 2000 census recorded approximately 1.78 million Deaf individuals (Department of Linguistics - UP Diliman, 2023). In Odiongan municipality, reports from the local government identified 97 deaf individuals. To address the communication gap, the Philippine government passed Republic Act 11106, known as The Filipino Sign Language Act, officially recognizing Filipino Sign Language (FSL) as the national sign language for the Filipino Deaf community. FSL serves not only as a communication medium but also as a representation of the unique culture and identity of the Deaf in the Philippines.

Technological advancements in recent years have made significant contributions to sign language recognition. Systems incorporating deep learning models such as MediaPipe integrated with LSTM have demonstrated high recognition accuracy, achieving up to 98.89% (Doan, 2022), while ResNet-50 models have shown 86.7% accuracy (Montefalcon et al., 2021). However, many of these existing solutions are designed for desktop platforms, limiting their usability in real-time, everyday scenarios. The emergence of frameworks like Google's MediaPipe offers a new avenue for real-time hand gesture recognition directly on mobile devices. When combined with lightweight deep learning models such as Convolutional Neural Networks (CNNs) and LSTM architectures, these frameworks enable highly efficient gesture recognition with minimal latency on smartphones (Di Leo et al., 2025).

Mobile platforms are ideal for delivering accessible Filipino Sign Language recognition tools due to their portability, real-time responsiveness, and widespread use (Wee et al., 2024). Nevertheless, dynamic gestures still present challenges in recognition due to motion variability, background clutter, and inconsistent lighting (Ezra et al., 2024). Hybrid models, combining CNNs with LSTM or Transformer architectures, have shown potential in capturing the temporal patterns of gesture sequences more effectively (Pigou et al., 2018). Deployment frameworks like TensorFlow Lite and MediaPipe further facilitate real-time, on-device processing, allowing accurate recognition even on mid-range smartphones (Lugaresi et al., 2019).

In response to these opportunities and challenges, this study presents the development of KamAi, a mobile application that uses deep learning to recognize Filipino Sign Language hand gestures. The name KamAi blends the Filipino word kamay (hand) with AI (Artificial Intelligence), reflecting the app's purpose of using advanced technology to improve accessibility. KamAi aims to provide a user-friendly, mobile-first platform for basic learners of FSL, offering real-time recognition and interactive learning tools. By improving communication access and promoting sign language learning, KamAi contributes to fostering inclusivity and breaking down communication barriers for the Filipino Deaf community.

2. Literature Review

2.1. Importance of Filipino Sign Language (FSL) and Technological Innovations

Filipino Sign Language (FSL) plays a critical role in fostering communication and inclusion for the Deaf community in the Philippines. Despite its official recognition through Republic Act 11106, challenges remain in promoting its widespread use due to limited public awareness and the scarcity of effective technological support systems (Montefalcon et al., 2021). As FSL is an essential medium for closing communication gaps, addressing these challenges through technological innovation becomes increasingly important.

Several studies have sought to bridge these gaps through the development of assistive technologies. Abougarair and Arebi (2022) proposed a glove-based FSL translator using accelerometers to detect hand motions. Although this system demonstrated functionality, the requirement of wearing specialized gloves made it impractical for day-to-day use. To address usability issues, Evangelista et al. (2023) developed a vision-based system integrating Google's MediaPipe and Long Short-Term Memory (LSTM) networks, which showed promising results in recognizing both simple and complex gestures in real-time. However, this solution faced limitations related to scalability. Similarly, Zhang and Jiang (2024) employed MobileNetV2 for FSL recognition, achieving 93.29% accuracy, yet encountering difficulties in differentiating visually similar gestures such as "M" and "N." These studies collectively underscore the importance of developing more practical, scalable, and accurate recognition technologies for FSL.

2.2. Role of Convolutional Neural Networks (CNN) in Mobile Applications for Sign Language

Convolutional Neural Networks (CNNs) have significantly advanced the capabilities of sign language recognition systems, particularly for mobile platforms where computational efficiency is essential. Jarabese et al. (2021) achieved 95% accuracy using CNNs for FSL recognition, utilizing data augmentation and fine-tuning to enhance model robustness. However, their research highlighted the ongoing challenge of maintaining performance in diverse lighting conditions and complex backgrounds, issues that are particularly relevant to mobile application development.

Madrid et al. (2022) further extended the capabilities of gesture recognition by combining MediaPipe's hand pose detection with LSTM networks to capture the dynamic aspects of gestures, achieving over 98% accuracy. Cayme et al. (2024) emphasized the necessity of balancing computational efficiency with model performance, especially when deploying CNNs on resource-constrained mobile devices. These findings illustrate that while CNNs are highly effective for sign language recognition, continued research is needed to optimize these models for reliable, real-time performance on mobile platforms.

2.3. The Potential of Google MediaPipe in Gesture Recognition

Google MediaPipe has emerged as a robust and lightweight framework for real-time hand gesture recognition. Its cross-platform capabilities make it particularly suitable for mobile applications requiring fast and efficient processing. Madrid et al. (2022) demonstrated the effectiveness of integrating MediaPipe with LSTM networks to accurately recognize dynamic FSL gestures, while Bora et al. (2023) successfully applied MediaPipe to Assamese Sign Language, proving its versatility across different sign languages.

Evangelista et al. (2023) also leveraged MediaPipe for real-time gesture recognition in FSL systems, confirming its efficiency in capturing hand landmarks. However, despite these advancements, the potential for further improvement remains, particularly in combining MediaPipe with advanced AI architectures like transformers to improve recognition accuracy and adaptability in more diverse environments. This integration offers a promising research direction for enhancing dynamic gesture recognition on mobile platforms.

2.4. Socio-Technical Considerations and Accessibility in Mobile Application Development

The development of effective FSL recognition systems extends beyond technical challenges and includes important socio-technical considerations. Accessibility, user experience (UX), and user involvement are critical for ensuring that these systems meet the needs of the Deaf community. Familoni and Babatunde (2024) emphasized the importance of human-centered UX design in healthcare applications, advocating for usability testing and regulatory compliance to ensure broad accessibility. Sadueste and Masalinto (2023) found that accessibility is directly linked to user satisfaction in academic library services, while Joshua et al. (2023) noted that trust in mobile health applications is strongly influenced by UX and security features.

In the context of assistive technology, involving the Deaf community during the development process ensures that resulting applications are functional, inclusive, and accepted by users. Gokgoz et al. (2025) stressed the importance of community engagement to guide design decisions, while Kousis et al. (2022) and Lipsmeier et al. (2022) demonstrated how mobile deep learning apps can improve accessibility in healthcare, offering lessons for inclusive design across various domains.

Furthermore, researchers such as Oliveira et al. (2024), Choo et al. (2019), and Ahsan et al. (2024) highlighted the need for ongoing accessibility audits, especially as software updates and UI changes may inadvertently affect usability for users with disabilities. Collectively, these studies advocate for a user-informed, standards-driven, and empathetic design approach in developing mobile applications for diverse users.

2.5. Hybrid Approaches for Enhanced Model Performance

Hybrid modeling approaches, which integrate multiple AI techniques, have shown promise for improving gesture recognition performance, particularly for dynamic sign language recognition. Madrid et al. (2022) emphasized the benefits of combining MediaPipe with recurrent architectures such as LSTM for capturing temporal features of sign language gestures. Meanwhile, integrating transformer-based models and Gated Recurrent Units (GRUs) is an emerging direction for improving long-range dependency capture in gesture sequences.

While hybrid models offer superior performance, Madrid et al. (2022) also noted the importance of optimizing these models for mobile devices to maintain real-time responsiveness. Gokgoz et al. (2025) reinforced the value of incorporating user feedback throughout the development cycle to ensure the models meet practical, real-world needs. Collaborative research efforts in developing shared, annotated datasets can further support training and evaluation, ultimately enhancing model robustness and generalization across diverse users and environments.

Android Studio serves as the primary integrated development environment (IDE) for building Android-based mobile applications, offering comprehensive tools for development, testing, and optimization. Its seamless integration with machine learning frameworks such as TensorFlow Lite and Google's ML Kit enables developers to implement advanced AI-powered features like facial recognition, gesture tracking, and speech recognition within mobile applications. Android Studio's support for both Kotlin and Java allows developers to code efficiently, while its built-in emulator provides real-time testing across multiple device configurations. Furthermore, integration with Firebase enables cloud-based features such as real-time database storage, authentication, and cloud computing, expanding the potential for secure and scalable mobile applications. These capabilities make Android Studio an essential platform for developing sophisticated AI-powered mobile applications, including those aimed at FSL recognition.

Design Thinking has gained widespread recognition as a flexible, user-centered methodology for solving complex problems, particularly in educational technology. Rösch et al. (2023) outlined a structured design thinking framework that emphasizes iterative development, empathy with users, and interdisciplinary collaboration. This approach ensures that solutions are not only functional but also closely aligned with user needs. Ansori et al. (2023) demonstrated the effectiveness of design thinking in improving mobile learning applications, achieving an impressive usability score of 86% through iterative development and User Acceptance Testing (UAT). Bender-Salazar (2023) highlighted design thinking's capacity to address "wicked problems" by fostering creative solutions to complex challenges such as accessibility and inclusion. Applying design thinking to FSL recognition ensures that applications are not only technically sound but also pedagogically effective, inclusive, and widely accepted by both Deaf and hearing users.

2.6. Performance Metrics and Confusion Matrices in Model Evaluation

Accurate model evaluation is essential in ensuring the fairness and reliability of gesture recognition systems. While accuracy remains a widely used metric, researchers such as Chicco et al. (2021) argue that Matthews Correlation Coefficient (MCC) provides a more reliable measure, especially for imbalanced datasets commonly encountered in gesture recognition. Vanacore et al. (2022) further emphasized the importance of fairness in model evaluation, advocating for metrics that ensure balanced performance across all gesture categories, even those with fewer examples.

The application of these comprehensive evaluation metrics helps to avoid biased model performance, ensuring that the system effectively recognizes a wide variety of gestures, a critical factor in the practical deployment of FSL recognition systems.

2.7. The Role of User Acceptance Testing (UAT) and ISO 25010 in Mobile Educational Applications

User Acceptance Testing (UAT) is a crucial step in validating whether an educational mobile application effectively meets its intended goals and user needs. The integration of ISO 25010 software quality standards enhances the rigor of UAT by providing a structured evaluation framework encompassing eight key quality characteristics: functionality, usability, reliability, performance efficiency, compatibility, security, maintainability, and portability.

Huda et al. (2023) demonstrated the value of combining ISO 25010 and design thinking in the development of e-learning platforms, while Martino and Andry (2020) used this framework to identify functional discrepancies in business applications. Ombai et al. (2024) applied ISO 25010 in evaluating an augmented reality educational tool, achieving high usability even for novice users. Aditama and Hikmah (2022) highlighted its effectiveness in software comparisons, such as evaluating password managers. These studies illustrate how ISO 25010 strengthens UAT by ensuring that mobile applications are not only technically robust but also user-friendly, accessible, and effective for educational purposes.

3. Methodology

This section describes the research design, procedures, tools, and techniques applied in the design, development, and evaluation of KamAi, a mobile application designed to recognize Filipino Sign Language (FSL) hand signs.

3.1. Research Design

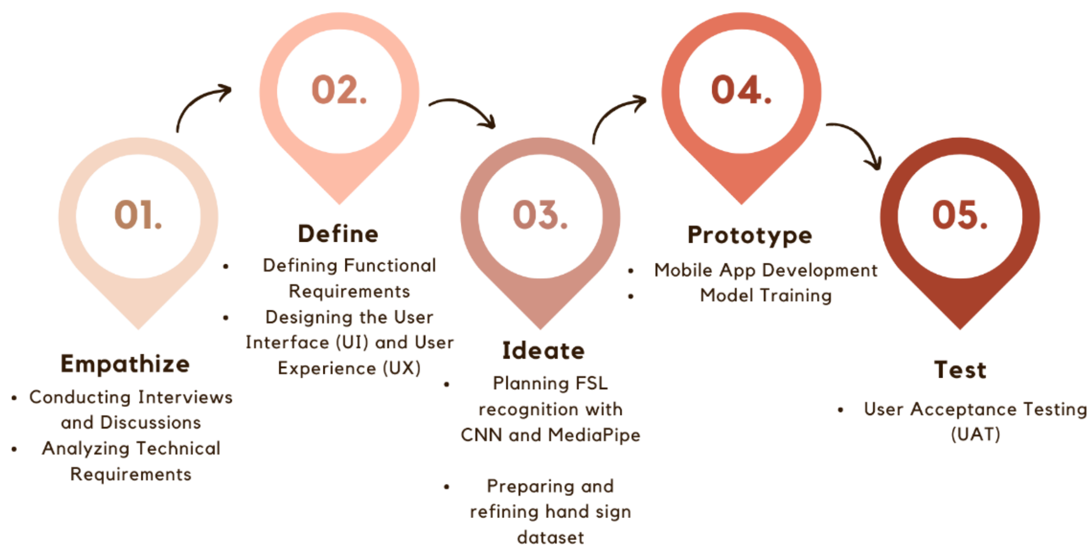
This study employed a developmental research design to guide the systematic creation and evaluation of KamAi. This approach enabled the researcher to iteratively design, prototype, and refine the mobile application in close alignment with the communication needs of Deaf and mute individuals, as well as hearing individuals seeking to learn Filipino Sign Language. The developmental research design allowed for repeated cycles of design, development, testing, and improvement, ensuring that the final product is both functional and user-centered.

To support this developmental process, the Design Thinking framework was adopted. Design Thinking is a widely recognized model that emphasizes empathy, problem-solving, and user-focused innovation. As outlined by Zulaikha et al. (2023), Design Thinking aligns technology solutions with the real-world needs of users. The process consists of five iterative phases: Empathize, Define, Ideate, Prototype, and Test, which together provided a structured path for the creation of KamAI.

3.2. Research Method

Figure 1

Design Thinking Model

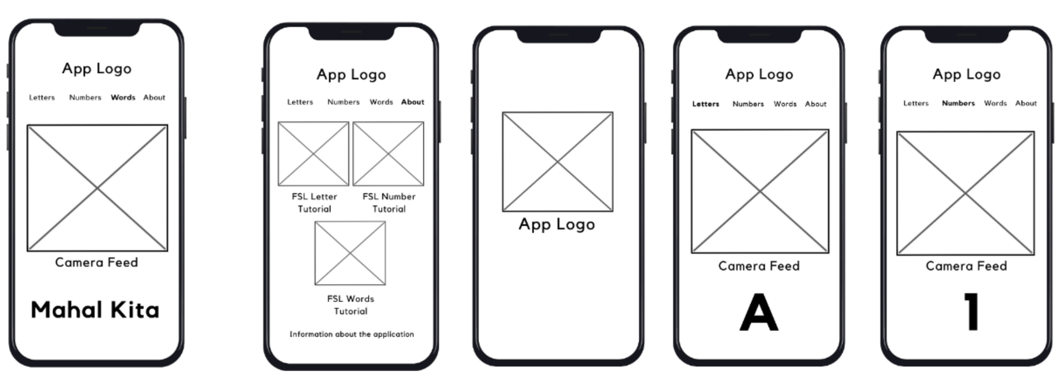


Empathize phase. In the empathize phase, the researcher-initiated engagements with the Persons with Disability Affairs Office (PDAO) in Odiongan, Romblon. The office provided demographic data on the local Deaf and mute population through formal coordination. Additionally, interviews and consultations were conducted with Filipino Sign Language (FSL) educators from Odiongan National High School and Odiongan Central Elementary School to better understand the communication challenges faced by the Deaf community. These consultations provided valuable insights into user expectations, learning difficulties, and technical needs. Functional and non-functional requirements for the application were identified, which included the need for a user-friendly interface, visual learning aids, and compatibility across various Android devices. The app was also designed to support offline usage, ensuring accessibility for users without stable internet access. The researcher also explored future possibilities of integrating real-time hand sign recognition through machine learning algorithms, although this feature was included as part of the app's long-term development roadmap.

Define Phase. The define phase focused on translating user insights into concrete functional requirements for the KamAi application. The core features identified included learning modules for FSL covering the alphabet, numbers, and basic vocabulary. Each module was supported by video demonstrations to enhance visual learning. A real-time hand sign recognition feature was also integrated, allowing users to perform hand signs that the mobile camera could detect and classify into corresponding letters, numbers, or words. Special attention was given to designing a clean and accessible user interface that would accommodate users of varying ages, literacy levels, and digital proficiencies. The user interface design was visualized using Canva, with a focus on bright colors, clear icons, and large typography to facilitate easy navigation and interaction for both Deaf users and hearing learners.

Figure 2

Screen layouts of the KamAi application



Ideate phase. The ideate phase involved selecting appropriate technical solutions to implement the real-time FSL recognition system. The application architecture integrated several key technologies. MediaPipe, an open-source framework developed by Google, was employed for real-time hand landmark detection, capturing 21 critical points on the hand. These landmarks were processed and fed into a Convolutional Neural Network (CNN) for gesture classification. TensorFlow Lite (TFLite) was used for model optimization and deployment on mobile devices. To train the CNN model, a large dataset was compiled, combining publicly available datasets from Kaggle and custom-curated data collected for this study. The dataset included 11,700 images of FSL letters, 16,500 images of FSL numbers, and 2,000 images of commonly used FSL words. To enhance the model's robustness, data augmentation techniques such as rotation, flipping, brightness adjustment, and scaling were applied, simulating various hand orientations, lighting conditions, and user environments. This comprehensive dataset ensured that the model could generalize well to real-world variations in hand gestures.

Figure 3

FSL letters

Sample dataset images of the FSL letter "A"

Sample dataset images of the FSL letter "B"

Figure 4

FSL numbers

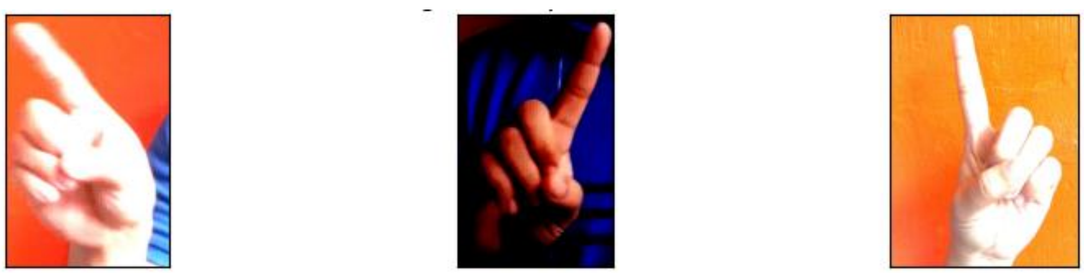
Sample dataset images of the FSL number "0"

Sample dataset images of the FSL number "1"

Figure 5

FSL words

Sample dataset images of the FSL word "Ilan"



Sample dataset images of the FSL word "Ano"

Prototype phase. In the prototype phase, the researcher developed and trained the CNN model using TensorFlow within Jupyter Notebook. The model architecture consisted of multiple convolutional and pooling layers, fully connected dense layers, and a final softmax output layer to classify gestures into their respective categories. Cross-validation techniques were applied throughout training to prevent overfitting and ensure the model's ability to generalize to new data. Evaluation metrics such as accuracy, precision, recall, and F1-score were monitored during training to optimize performance. The training process was conducted on a high-performance computing system, utilizing an NVIDIA GeForce RTX 3070 GPU, 24 GB RAM, and an AMD Ryzen 9 processor running on Ubuntu 24.04 LTS. Once training was complete, the model was converted into TensorFlow Lite format for mobile deployment.

Subsequently, the mobile application was developed using Android Studio (Koala Version) as the integrated development environment (IDE). Android Studio supported the seamless integration of the trained model into the mobile interface. The user interface was designed to allow real-time hand sign recognition, leveraging MediaPipe's Gesture Recognizer and TensorFlow Lite for on-device inference. The final mobile application was compatible with Android devices running API Level 21 and above, ensuring accessibility across a wide range of smartphones.

Table 1

Technical specifications for model training

Component	**Specification**
Operating System	Ubuntu 24.04 LTS
GPU	NVIDIA GeForce RTX 3070 (8GB VRAM)
RAM	24 GB
Processor (CPU)	AMD Ryzen 9 6900HX
Development Platform	Jupyter Notebook
Frameworks Used	TensorFlow, MediaPipe
Model Format	TensorFlow Lite (.tflite)

Table 2

Technical tools for mobile application development

Component	**Specification / Tool**
Mobile IDE	Android Studio (Koala Version)
Code Editor	Visual Studio Code (with Jupyter Plugin)
Programming Language	Python (for model training), Java/Kotlin (for Android)
Model Integration	TensorFlow Lite (.tflite)
Frameworks & Libraries	MediaPipe, TensorFlow Lite
Target Platform	Android
Minimum SDK	API Level 21 (Android 5.0 Lollipop)

Test phase. The test phase involved comprehensive evaluation of the KamAi application, guided by the ISO/IEC 25010:2011 software quality model. Six key quality attributes were assessed: functional suitability, usability, reliability, performance efficiency, compatibility, and portability. User Acceptance Testing (UAT) was conducted with a diverse group of participants, including Deaf and mute individuals, FSL educators, IT experts, and representatives from PDAO. Participants used a structured questionnaire based on ISO 25010 criteria, rating various aspects of the application using a 5-point Likert scale. This comprehensive evaluation allowed for user-centered feedback on the application's effectiveness, responsiveness, and accessibility, enabling the researcher to identify areas for improvement and refinement.

3.3. Research Locale and Participants

The research was conducted in Odiongan, Romblon, Philippines, from December 2024 to May 2025. The location was chosen due to its active Deaf community and existing support from local government units and educators. A purposive sampling method was employed to select participants who could provide meaningful feedback on both technical and functional aspects of the application. The participant group included 21 Deaf and mute individuals, 50 hearing learners and students, 4 certified FSL educators, 10 IT experts, and 2 PDAO representatives. This diversity of participants ensured comprehensive input from both end-users and domain experts, strengthening the validity of the evaluation process.

3.4. Research Instrument and Validation

The primary evaluation tool used in this study was a questionnaire based on the ISO/IEC 25010:2011 software quality model. The instrument was validated through expert review, involving three IT specialists, one grammar expert, and one statistician. The IT specialists reviewed the tool for technical accuracy, the grammar expert ensured clarity and readability, and the statistician confirmed the appropriateness of the Likert scale. A reliability test was conducted, yielding a Cronbach's alpha coefficient of 0.855, which indicates excellent internal consistency and confirms the tool's reliability for measuring software quality attributes.

3.5. Data Gathering Procedure

The data gathering process involved the collection of both secondary and primary datasets. Publicly available datasets from Kaggle were used, supplemented by new images captured locally with the assistance of FSL educators and Deaf participants. The complete dataset included images for FSL letters, numbers, and 16 commonly used words. All images were labeled and reviewed for accuracy, with variations included to represent different hand sizes, skin tones, and environmental conditions. This comprehensive dataset served as the foundation for training the recognition model.

3.6. Data Processing and Analysis

Prior to training, all images underwent preprocessing, including resizing, normalization, and noise reduction to standardize input for the CNN model. Data augmentation techniques such as rotation, flipping, brightness adjustment, and scaling were applied to enrich the dataset and improve the model's ability to handle diverse real-world scenarios. The processed images were then fed into the CNN model for training and classification. MediaPipe was used to extract hand landmarks, which served as input features for gesture recognition. TensorFlow Lite enabled lightweight deployment of the trained model on mobile devices. Model performance was evaluated using standard classification metrics derived from the confusion matrix, including accuracy, precision, recall, and F1-score. These metrics provided a comprehensive assessment of the model's ability to correctly recognize FSL gestures.

3.7. Statistical Treatment

Descriptive statistics were applied to analyze user feedback collected through the ISO 25010-based questionnaire. Mean scores were computed for each software quality attribute, while model evaluation metrics were calculated to quantify the performance of the CNN recognition model. The combined statistical analysis of both quantitative model performance and qualitative user evaluations provided a robust assessment of KamAi's technical and functional success.

3.8. Ethical Considerations

Informed consent was obtained from all participants, and personal data confidentiality was maintained throughout the study. The application does not collect or store any user data during operation. All datasets used for model training were either publicly available or collected with explicit consent. Measures were taken to prevent AI model bias by ensuring diversity in the training dataset, covering variations in skin tone, hand size, and environmental conditions. Ethical principles of beneficence and non-maleficence were observed throughout the research, ensuring that the KamAi application promotes accessibility and inclusion for the Deaf community while minimizing any potential harm or misuse.

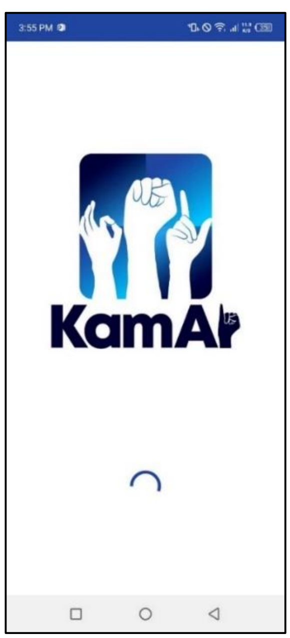
4. Results

KamAi was successfully developed using Android Studio, TensorFlow Lite, and Google's MediaPipe framework. The application features real-time FSL recognition across three learning categories: letters, numbers, and words. The application's interface includes a splash screen, home screen with live camera feed, intuitive mode selection buttons, and an "About" section offering categorized tutorial videos. Users can switch between recognition modes easily, allowing for progressive learning of the FSL alphabet, numbers, and commonly used phrases. This layered approach to learning provides an inclusive experience for beginners and more advanced learners alike

Splash Screen. The splash screen is the first visual presented to the user upon launching the KamAi application. This screen features the KamAi logo while the application loads in the background. The splash screen serves not only as a visual identity for the app but also provides a brief loading period to prepare necessary resources for seamless use.

Figure 6

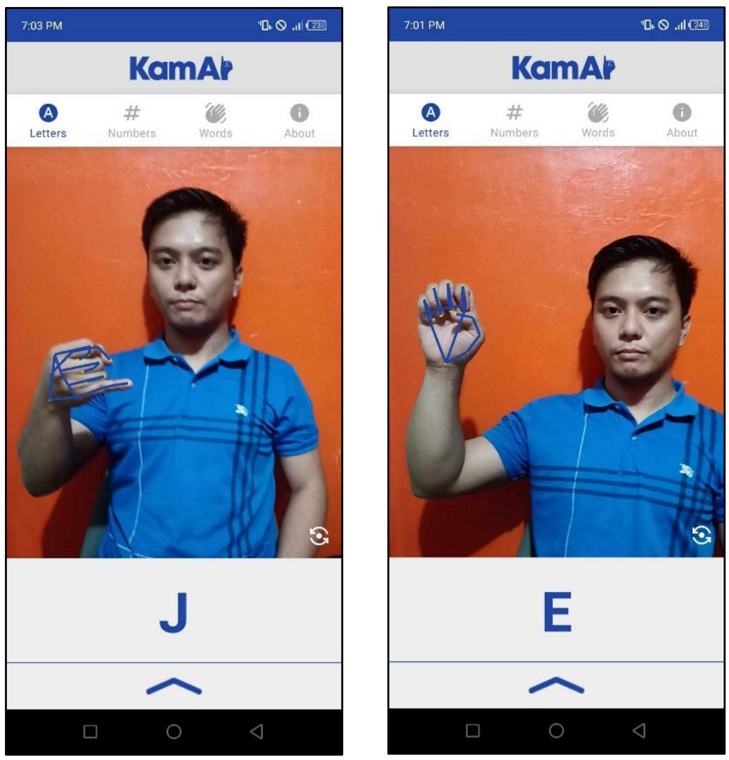
KamAi splash screen



Home screen. The home screen serves as the main interface of the application. Upon entering this screen, users are immediately presented with the camera feed, allowing them to begin using the app for gesture recognition. The Letters Screen is set as the default screen on the Home Screen, which means the app automatically opens to this mode, where users can begin practicing the recognition of the FSL alphabet (A-Z) by showing hand gestures. Additionally, the home screen displays the recognized FSL sign as a corresponding letter, number, or word. It also includes essential navigation buttons, which allow users to switch between the different modes.

Figure 7

KamAi splash screen



Mode selection. The app provides users with three primary recognition modes, designed to accommodate different learning levels and needs. These modes include:

Letters: Recognizes hand signs for the FSL alphabet (A–Z). Each correctly performed sign is shown as its corresponding letter on the screen.

Numbers: Detects hand gestures for numbers from 1 to 10 and displays the recognized number as text.

Words: Recognizes 17 words and phrases such as “Mahal kita,” “Mama,” “Papa,” and “Ate, Ano, Hello, Kailan, Lolo, Lola etc.

Figure 8

Recognition output for FSL letters A, B, C, D, E and E

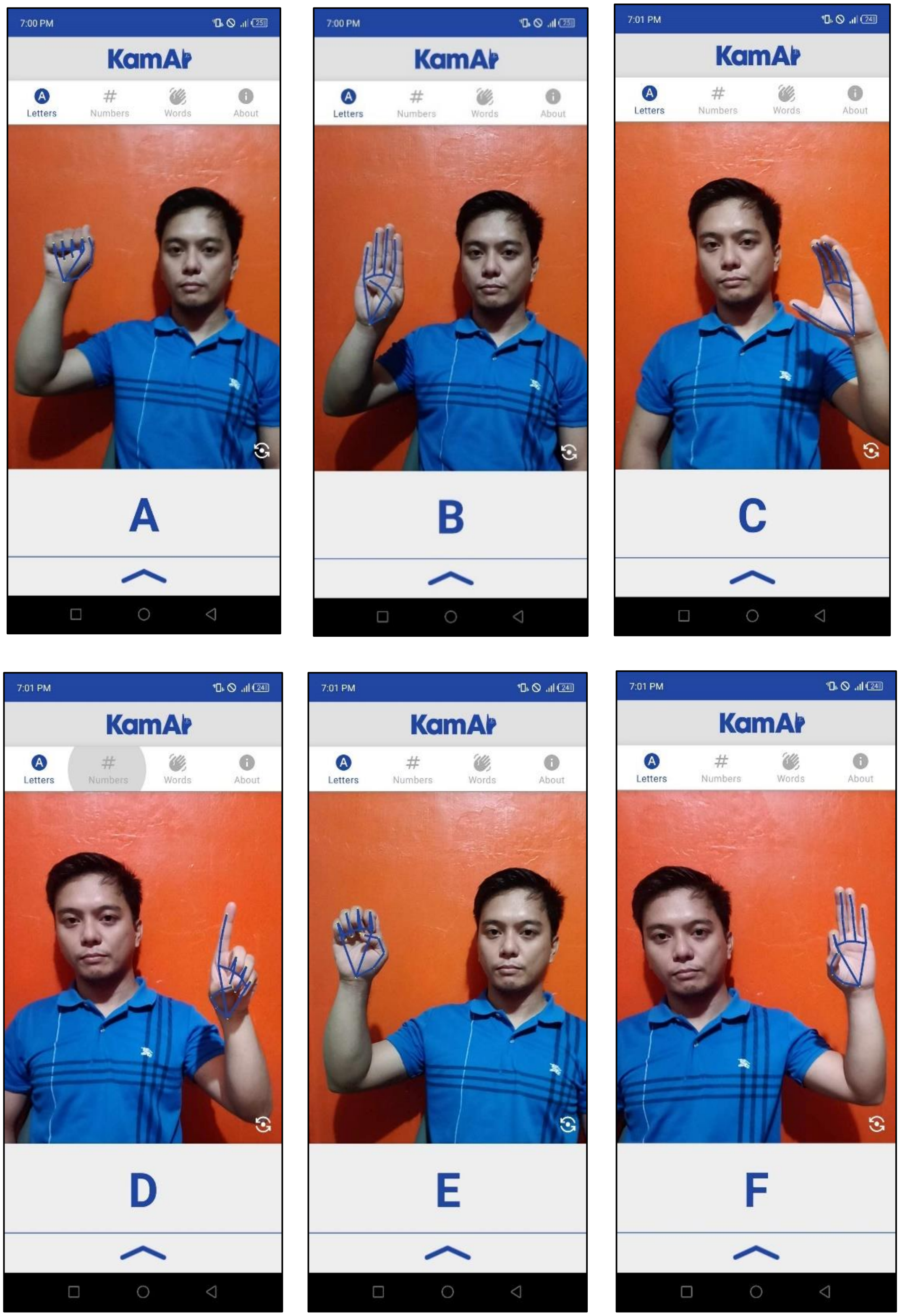


Figure 9

Recognition output for FSL letters G, H, I, J, K and L

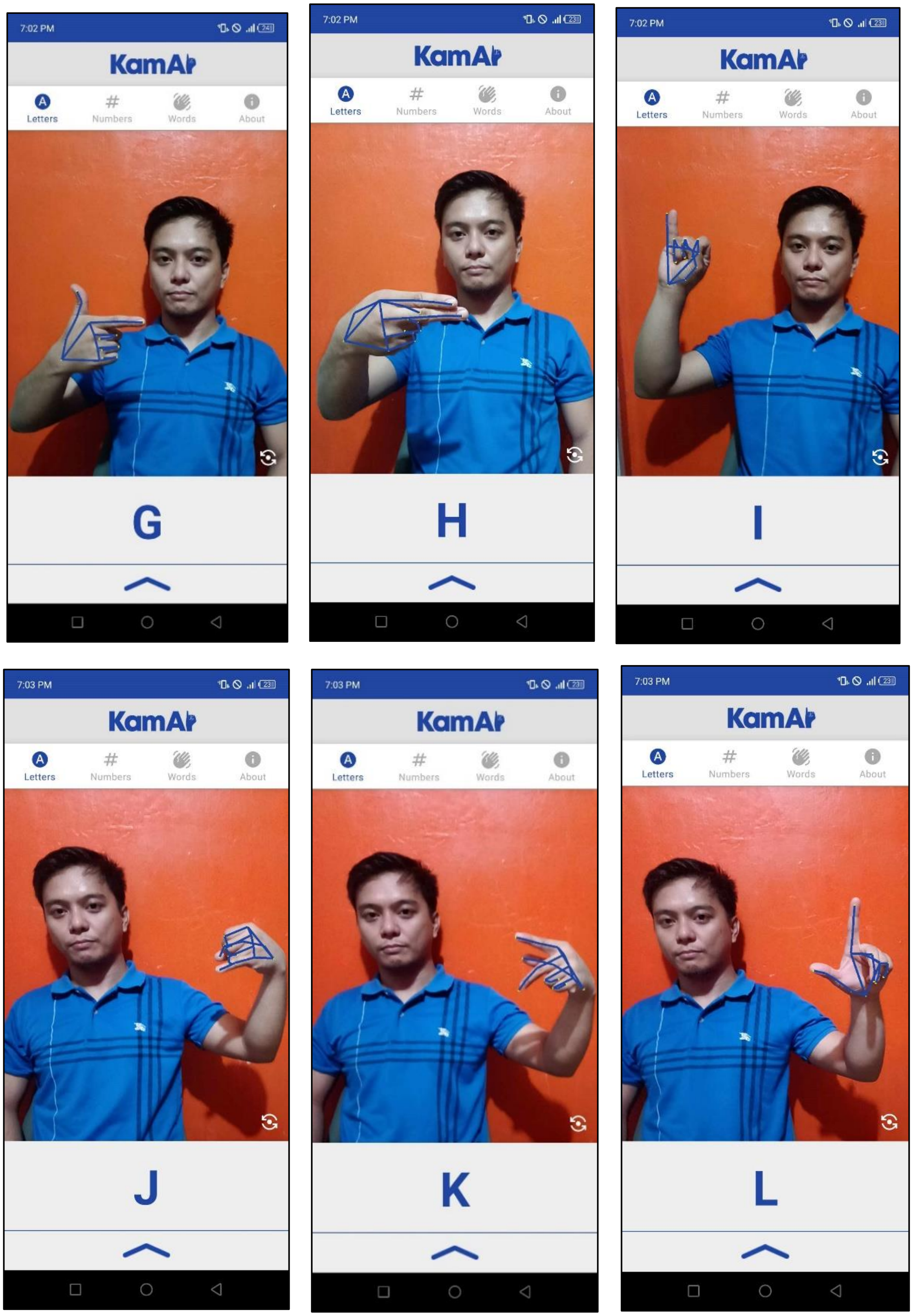


Figure 10

Recognition output for FSL letters M, N, O, P, Q and R

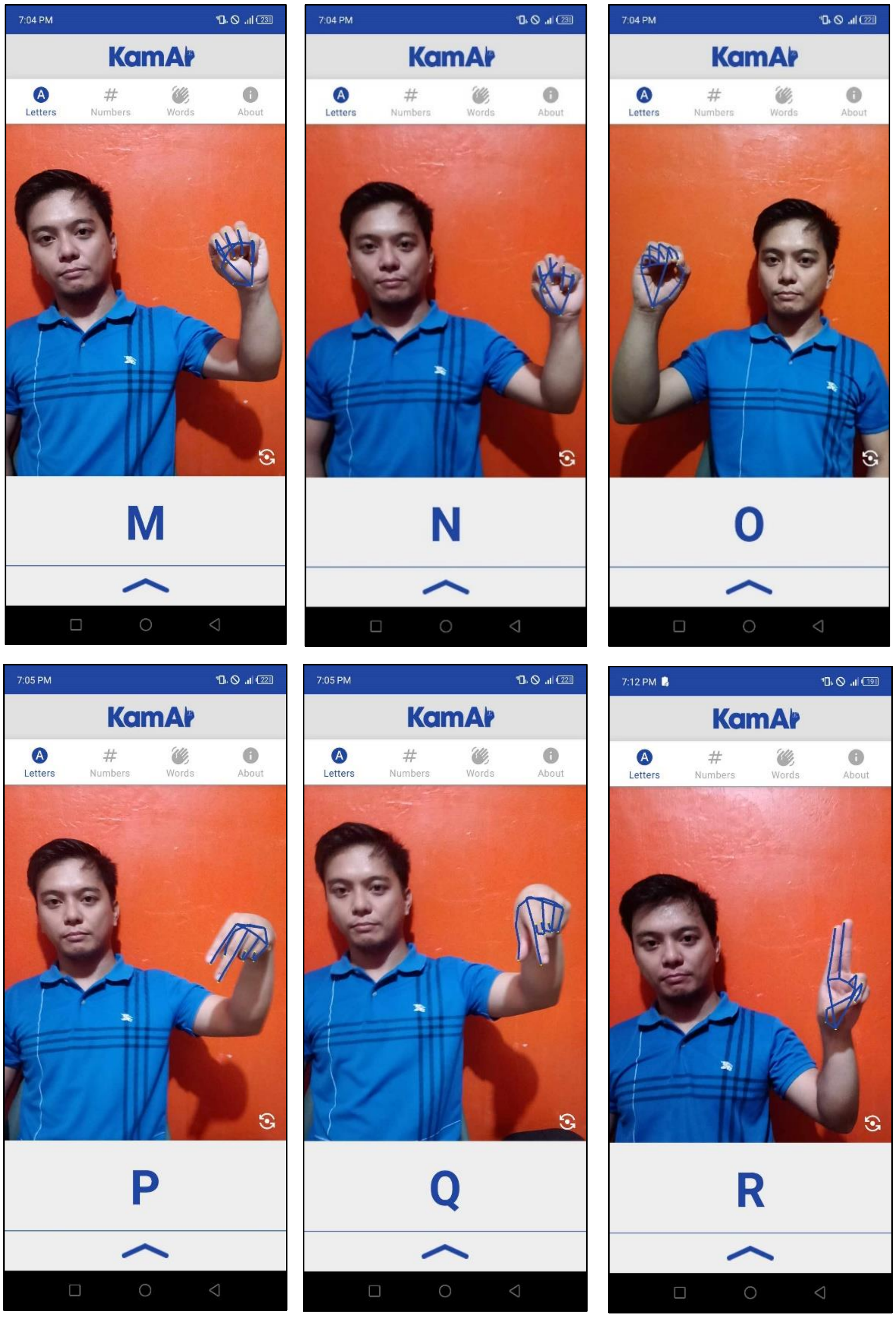


Figure 11

Recognition output for FSL letters S, T, U, V, W and X

Figure 12

Recognition output for FSL letters Y and Z

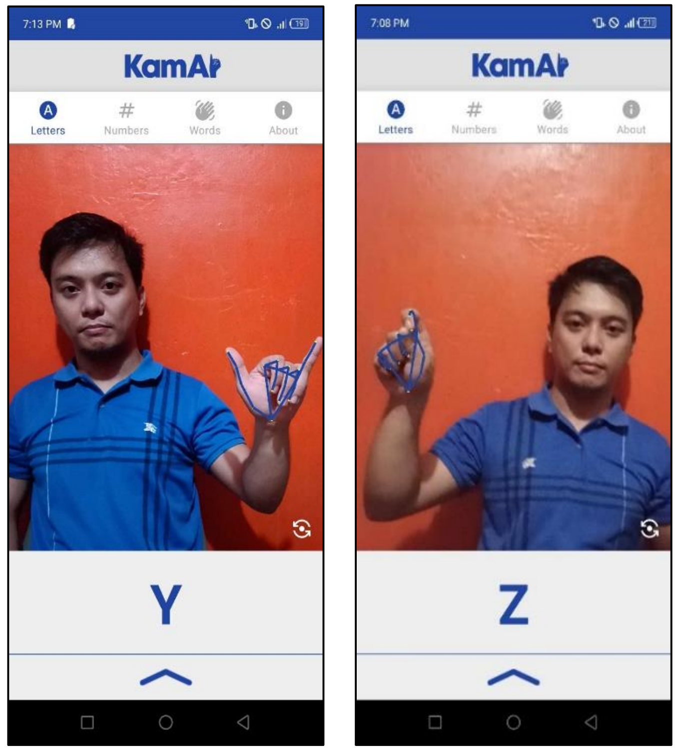


Figure 13

Recognition output for FSL numbers 1, 2 and 3

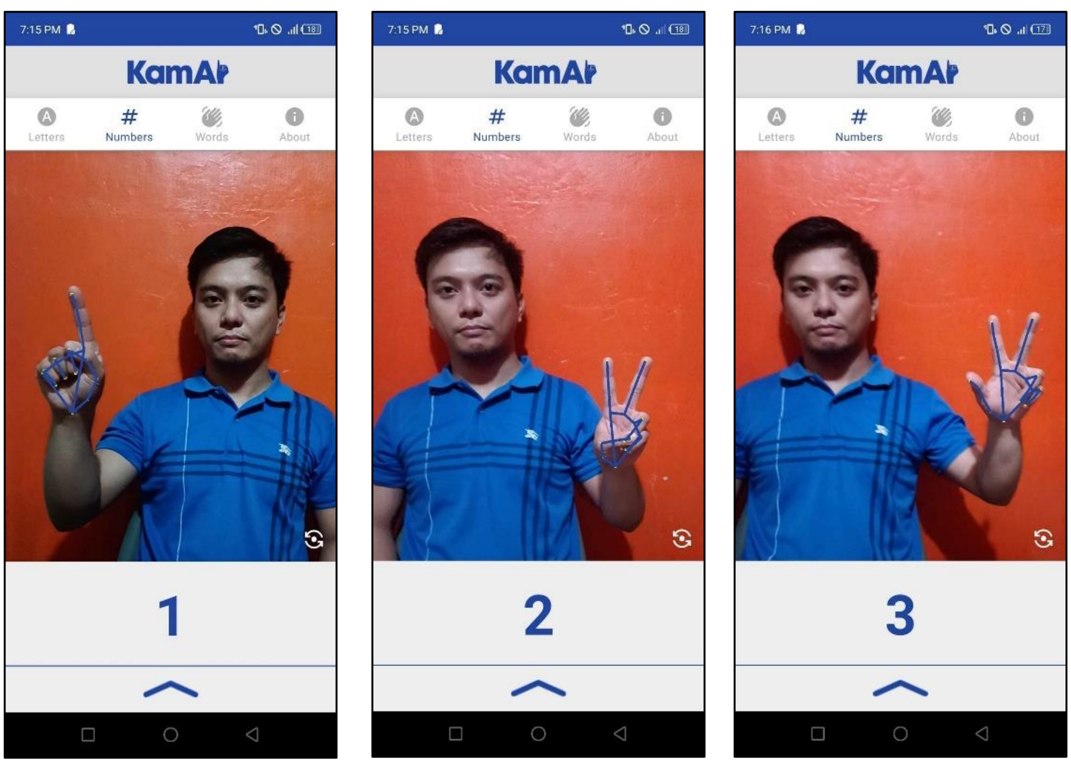


Figure 14

Recognition output for FSL numbers 4,5,6,7,8 and 9

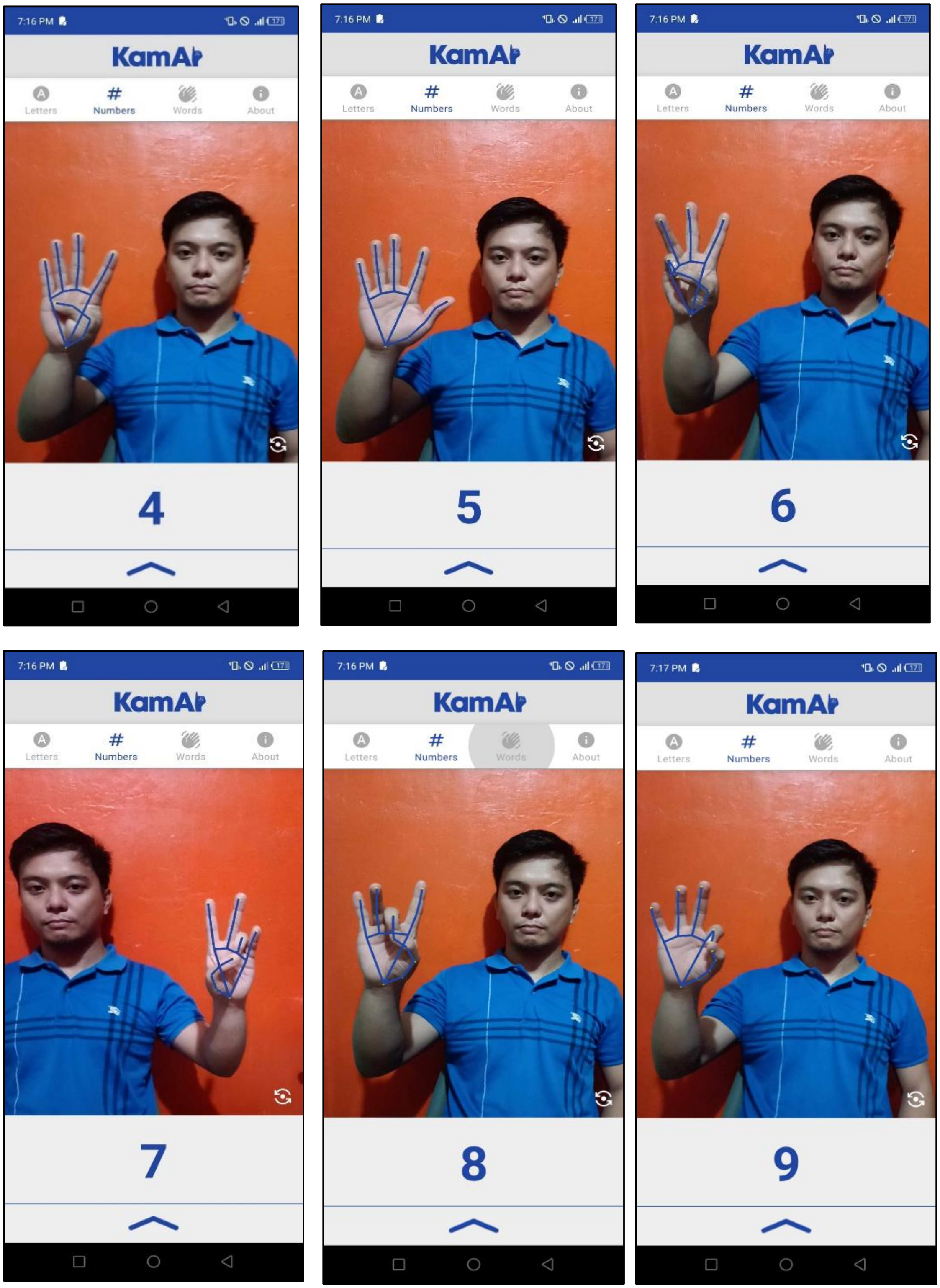


Figure 15

Recognition output for FSL number 10

Figure 16

Recognition output for FSL words lolo, kain and ilan

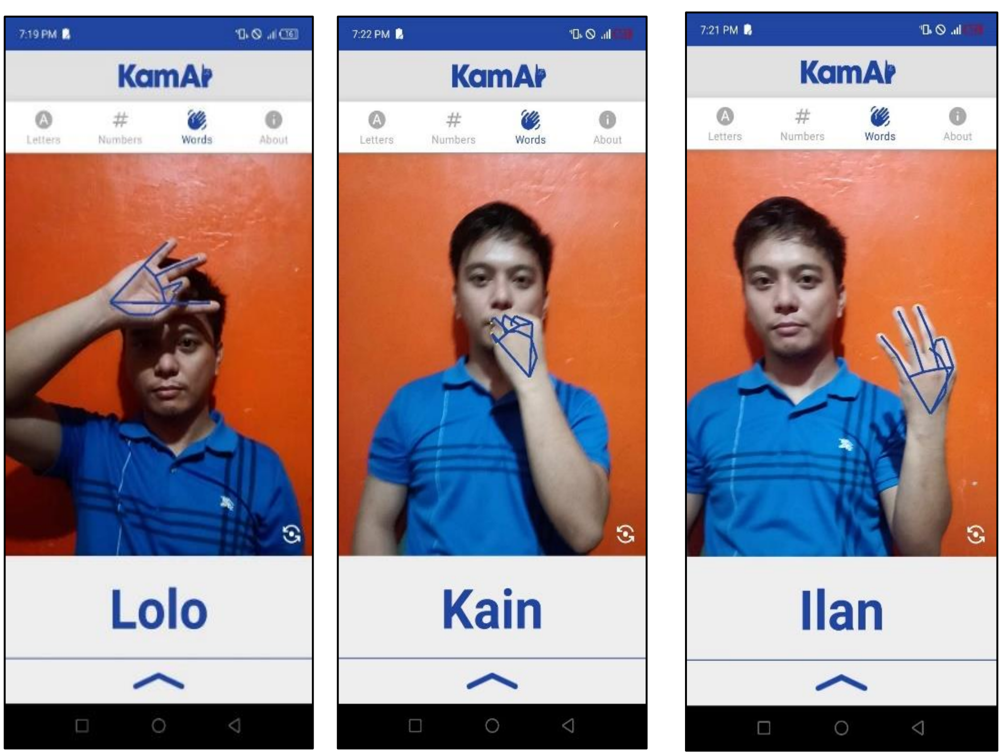


Figure 17

Recognition output for FSL words lola, baby, okay, magkano, ano, pinsan

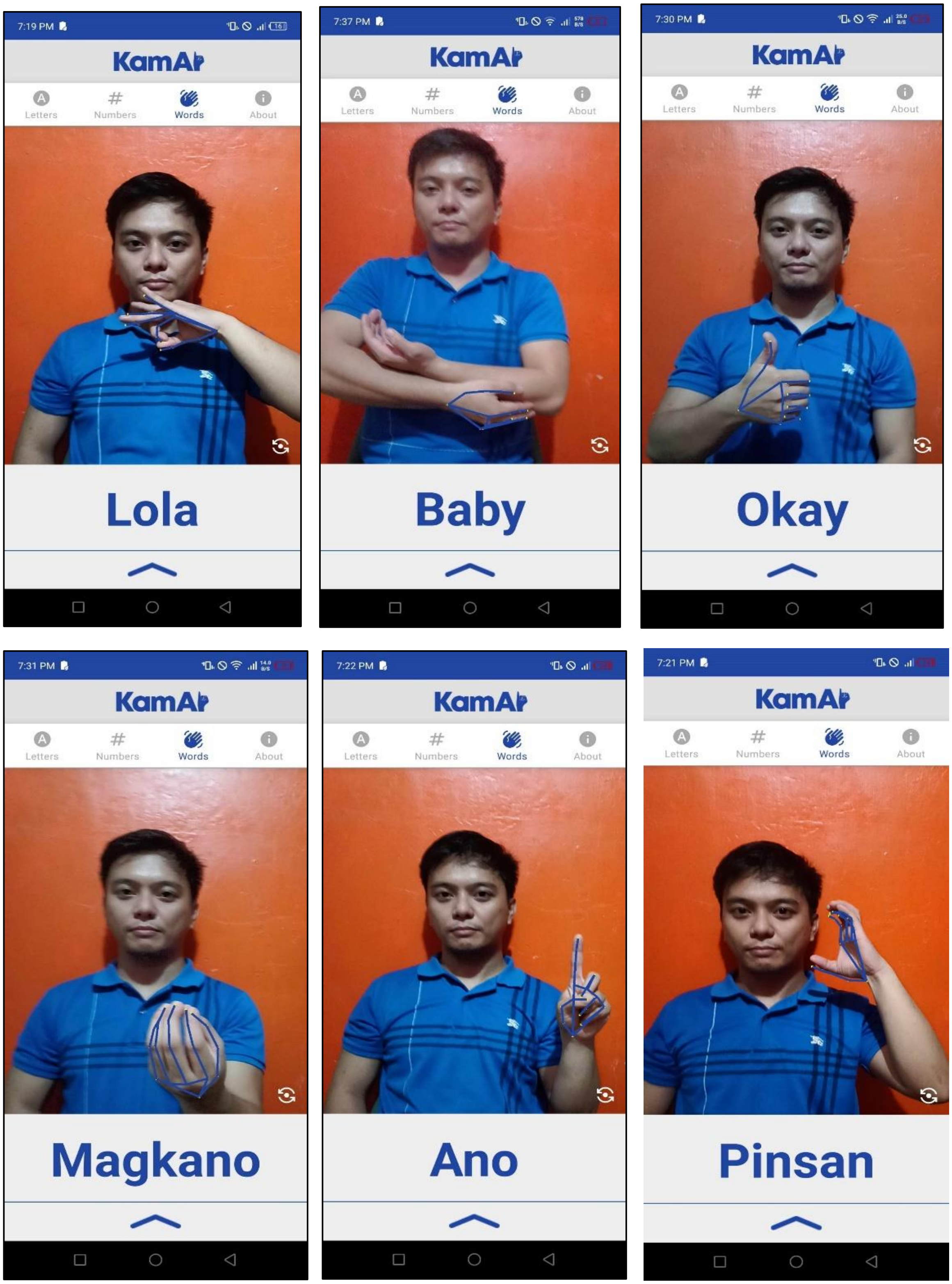


Figure 18

Recognition output for FSL words hello, mahal kita, papa, kuya, ate and mama

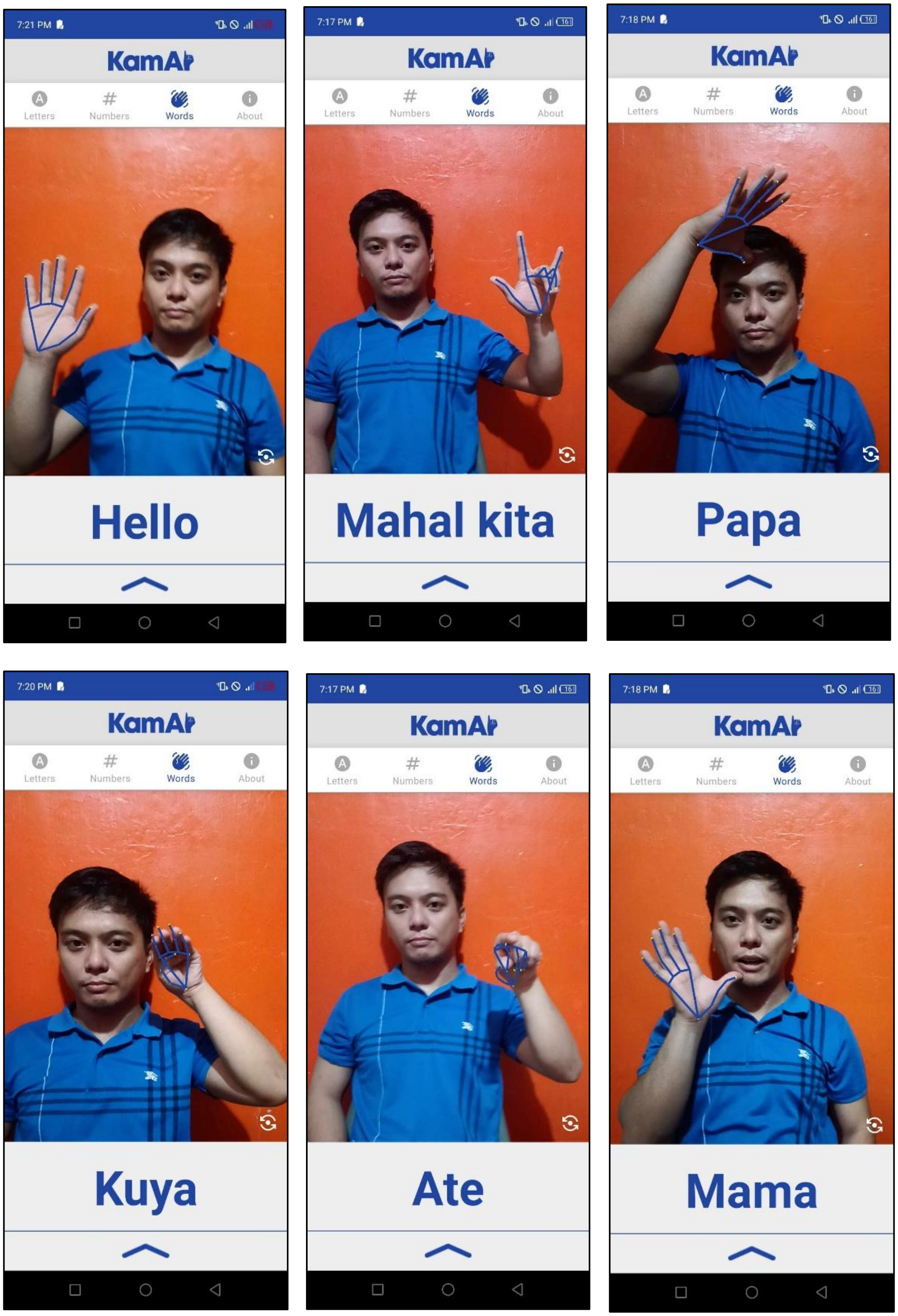


Figure 19

Recognition output for FSL word paano

Each mode is clearly represented through intuitive icons on the home screen, making it easy for users to switch between the different modes based on their learning preferences.

About Screen. The About Screen provides tutorial videos grouped into three playlists: FSL Letters, Numbers, and Words. Users need an internet connection to access the playlists, which are hosted on YouTube. They can also download the videos to YouTube Downloads and watch them offline anytime.

Figure 20

Category selection screen for FSL tutorial playlists

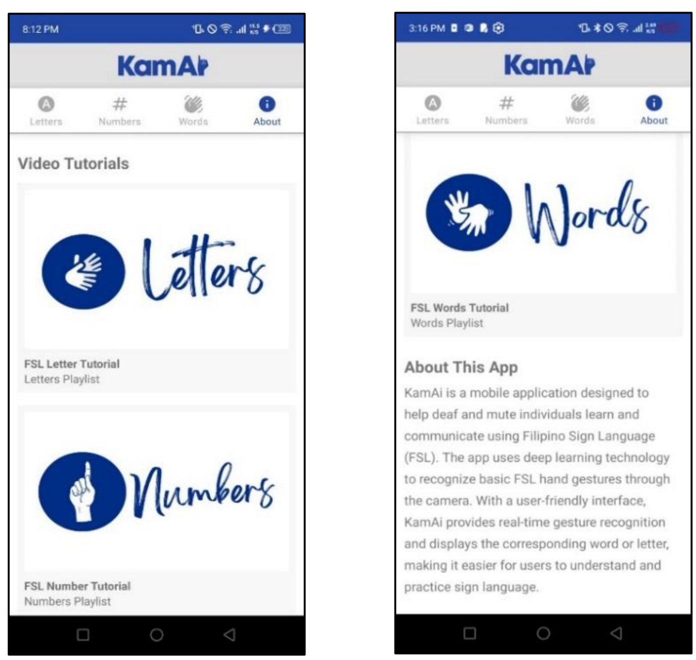
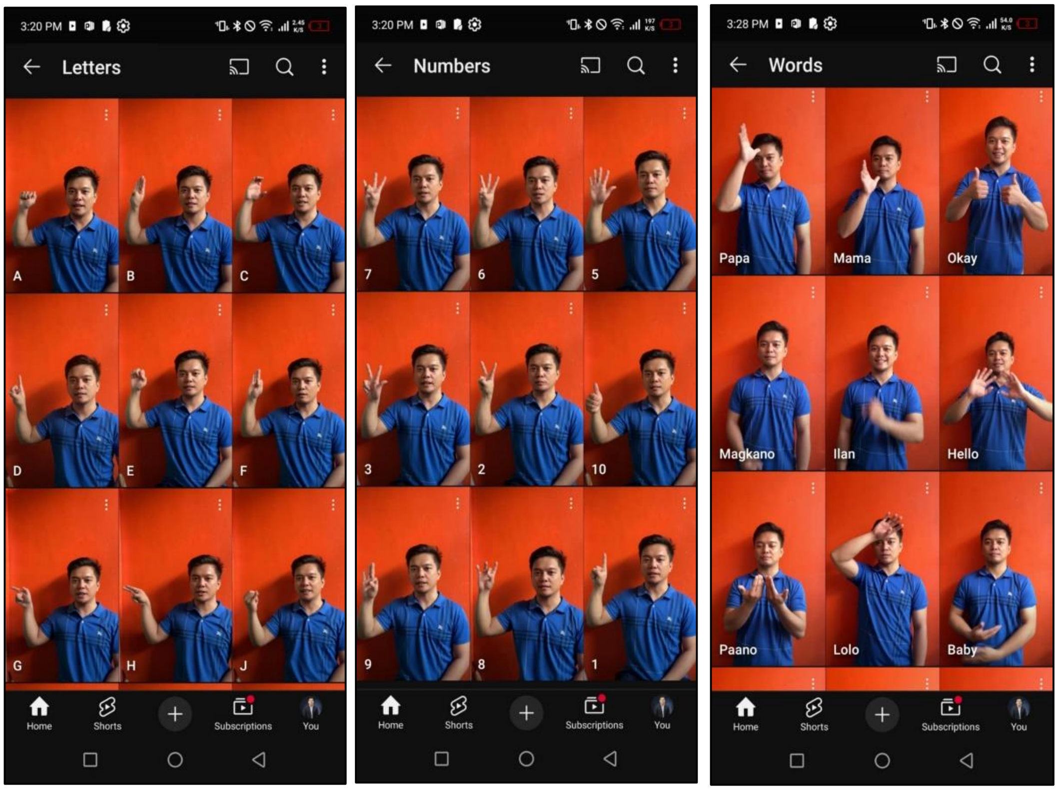


Figure 21

Tutorial video playlist for FSL words, letters, and numbers



The developed CNN-based recognition model was trained and tested using a comprehensive dataset composed of publicly available and custom-collected images. The model's performance was evaluated using standard classification metrics, including accuracy, precision, recall, and F1-score, with separate assessments for FSL letters, numbers, and words. For the FSL Letters dataset, the model achieved a precision of 97%, recall of 87%, and F1-score of 90%, demonstrating high accuracy in recognizing most letters. Some confusion occurred between visually similar hand signs such as "M" and "N," suggesting potential areas for further improvement. For the FSL Numbers dataset, the model achieved a precision of 96%, recall of 88%, and F1-score of 91%, performing reliably in identifying most number gestures, though occasional misclassifications occurred between similar hand shapes. The FSL Words dataset also yielded strong results, with a precision of 96%, recall of 88%, and F1-score of 91%, indicating the model's ability to handle more complex gestures involving common Filipino words and phrases.

Figure 22

Performance metrics of the deep learning model on FSL letters dataset

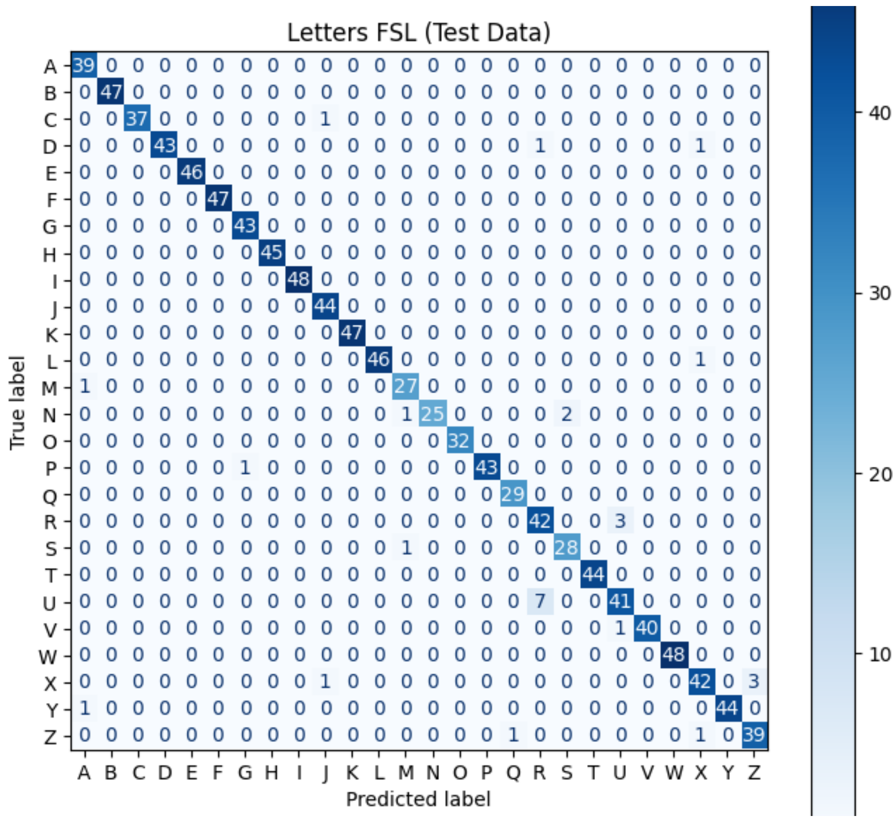


Figure 23

Performance metrics of the deep learning model on FSL numbers dataset

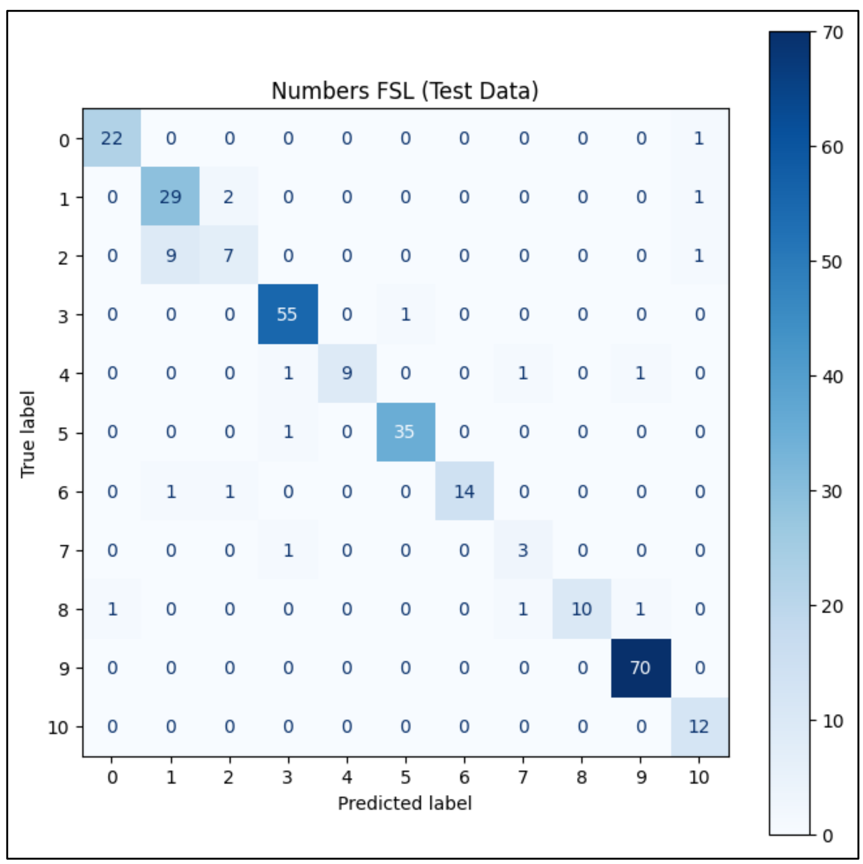
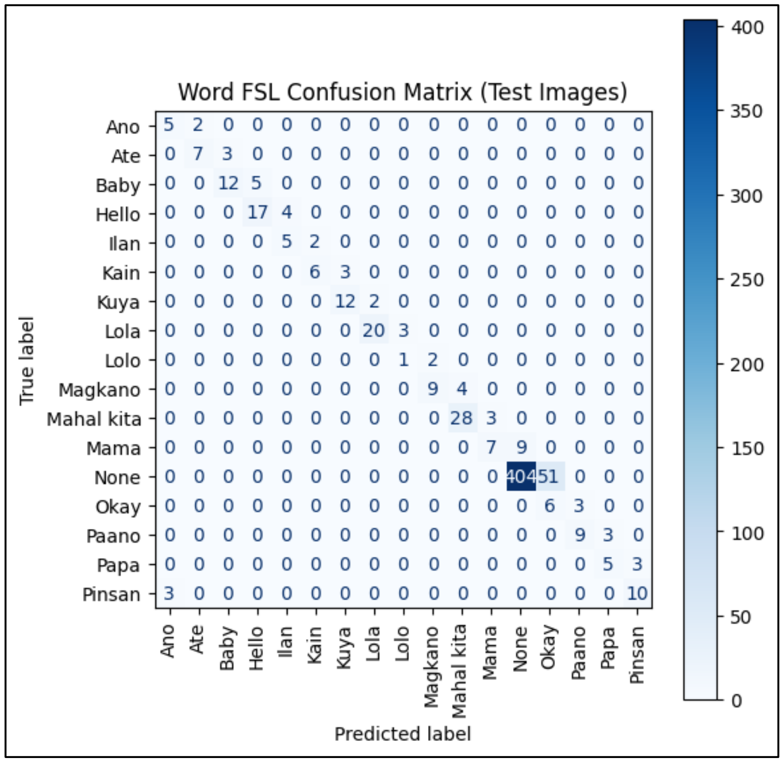


Figure 24

Performance metrics of the deep learning model on FSL words dataset



To assess the application's real-world usability and software quality, UAT was conducted using the ISO/IEC 25010:2011 framework. The UAT involved Deaf and mute individuals, hearing learners, FSL educators, IT experts, and local government representatives. The evaluation covered six software quality characteristics: functionality, reliability, compatibility, usability, efficiency, and portability.

Table 3

User Acceptance Testing results of the KamAi App based on ISO 25010 software quality characteristics

Indicators	Mean	SD	DI
Functionality	4.73	0.48	Highly Acceptable
Reliability	4.51	0.64	Highly Acceptable
Compatibility	4.71	0.47	Highly Acceptable
Usability	4.5	0.66	Highly Acceptable
Efficiency	4.59	0.61	Highly Acceptable
Portability	4.77	0.42	Highly Acceptable

Legend: 1.00-1.80 Highly Unacceptable; 1.81-2.60 Unacceptable; 2.61-3.40 Fairly Acceptable; 3.41-4.20 Acceptable; 4.21-5.00 Highly Acceptable

Across all attributes, KamAi achieved mean scores falling within the "Highly Acceptable" range. Functionality (4.73), compatibility (4.71), and portability (4.77) received the highest ratings, indicating that the application performs its intended functions accurately, operates across multiple devices and Android versions, and can be easily used on a variety of hardware platforms. Reliability (4.51), usability (4.50), and efficiency (4.59) also received high ratings, reflecting smooth operation, intuitive interface design, and minimal resource consumption. These findings confirm that the application meets user expectations for quality, accessibility, and practicality in supporting FSL learning.

5. Discussion

5.1. Development and Utility of KamAi for FSL Learning

The KamAi mobile application represents a significant step forward in the integration of assistive technology for Filipino Sign Language education. The design of the application effectively combines technological sophistication with accessibility, addressing the communication and learning needs of both Deaf and hearing individuals. The structured learning pathways—divided into letters, numbers, and common words, support a gradual acquisition of sign language skills, promoting independent learning for beginners while offering advanced learners more complex vocabulary practice. The application's real-time gesture recognition system, powered by CNN and MediaPipe, provides immediate feedback to users, reinforcing learning and increasing engagement. These design choices are consistent with prior research highlighting the importance of real-time interaction and multimedia resources in sign language learning (Montefalcon et al., 2021).

5.2. Model Performance Across FSL Data Categories

The CNN-based recognition model demonstrated high performance across all three datasets. The letter recognition model, with an accuracy of 88.38%, precision of 97%, recall of 87%, and F1-score of 90%, suggests that static hand gestures for alphabet learning are well-suited for CNN-based classification models. However, the noted misclassifications between visually similar signs like "M" and "N" highlight opportunities for refinement, potentially through deeper network architectures such as ResNet or hybrid models incorporating temporal features (Montefalcon et al., 2021).

For number recognition, the model performed exceptionally well with 91.41% accuracy and 96% precision but displayed some confusion between closely related signs, particularly between numbers "1" and "2," as well as among "6," "7," and "8." Similar challenges have been reported by other researchers when distinguishing subtle differences in finger positioning (Rahimian et al., 2021), suggesting that future models might benefit from transformer-based approaches or attention mechanisms to capture finer visual distinctions.

The word recognition model yielded an accuracy of 83.08%, precision of 96%, recall of 88%, and F1-score of 91%. These results confirm the model's capability to handle more complex multi-class classifications, which are essential for conversational vocabulary. While most frequently used words such as "Mama," "Mahal Kita," and "Lola" achieved high recognition rates, words with fewer training samples showed slightly lower performance. This observation emphasizes the importance of expanding the dataset and employing advanced augmentation techniques to further improve recognition rates for less frequently used signs (Feng et al., 2021).

5.3. Application Evaluation Based on ISO 25010 Quality Criteria

The results of the User Acceptance Testing clearly indicate that KamAi satisfies software quality standards across multiple dimensions. High scores in functionality and usability suggest that the application effectively meets the learning needs of its target audience, consistent with findings from previous usability studies in assistive learning technologies (Atanacio, 2022; Shah & Chiew, 2019). The strong ratings for compatibility and portability highlight KamAi's practical applicability across a broad range of Android devices, which is especially relevant in the Philippine context where users may rely on lower-end mobile hardware (Zulfa et al., 2020). Additionally, the efficiency and reliability scores affirm the application's ability to perform smoothly and consistently, even in resource-limited environments, making it suitable for everyday use in educational settings.

6. Conclusion

This study successfully designed, developed, and evaluated *KamAi*, an Android-based mobile application for real-time recognition of Filipino Sign Language (FSL) using deep learning techniques. The research addressed the need for accessible and inclusive digital tools that support FSL education and communication, particularly for the Deaf and mute community in the Philippines. Through the integration of Convolutional Neural Networks (CNNs) and the MediaPipe framework, the application demonstrated promising recognition capabilities across three key gesture categories: letters, numbers, and commonly used words and phrases.

The first research objective, centered on the development of an interactive and user-friendly mobile application, was effectively achieved. *KamAi* incorporates a structured and intuitive interface, offering three distinct recognition modes and immediate gesture-to-text feedback to facilitate user engagement and self-paced learning. The inclusion of tutorial video playlists, accessible both online and offline, further enriched the application's pedagogical functionality and extended its usability beyond real-time recognition tasks.

The second objective, which involved evaluating the model's performance across different datasets, produced favorable outcomes. The letter recognition model attained an accuracy of 88.38%, precision of 97%, recall of 87%, and an F1-score of 90%, indicating strong reliability in classifying static alphabetic signs. Similarly, the word recognition model achieved 83.08% accuracy, 96% precision, 88% recall, and an F1-score of 91%, demonstrating the model's capacity to handle a more complex set of FSL gestures relevant to everyday communication. Although the number recognition model recorded high accuracy (91.41%) and precision (89%), its lower recall (21%) and F1-score (28%) suggest challenges in differentiating numerically similar gestures, highlighting an area for potential improvement through further model refinement and dataset enhancement.

The third objective, which aimed to assess the application's overall quality and user acceptance using the ISO/IEC 25010:2011 software quality model, yielded highly positive results. User Acceptance Testing (UAT) revealed consistently high mean ratings across all six quality attributes: Functionality (4.73), Reliability (4.51), Compatibility (4.71), Usability (4.50), Efficiency (4.59), and Portability (4.77). These findings affirm the application's strong alignment with user expectations in terms of performance, responsiveness, cross-device operability, and ease of use. The application's high ratings across these dimensions reflect its potential as a reliable and effective assistive tool for FSL instruction and practice.

In conclusion, *KamAi* presents a viable solution for enhancing FSL learning and communication accessibility through mobile technology. It contributes to the broader discourse on inclusive education and digital equity by providing an innovative tool that bridges linguistic and social barriers. The application not only supports the linguistic needs of Deaf individuals but also promotes FSL literacy among hearing users, fostering mutual understanding and social inclusion. The findings underscore the importance of integrating user-centered design, robust model evaluation, and international software quality standards in the development of accessible learning technologies. Future work may focus on expanding the gesture library, improving recognition of similar hand signs, and incorporating adaptive learning features to further advance the application's impact and scalability.

Disclosure statement

No potential conflict of interest was reported by the authors.

Funding

This work was not supported by any funding.

References

Abougarair, A. J., & Arebi, W. A. (2022). Smart glove for sign language translation. *International Robotics & Automation Journal*, *8*(3), 109–117. https://doi.org/10.15406/iratj.2022.08.00253

Aditama, W. Y., & Hikmah, I. R. (2022, December 20). Analisis perbandingan kinerja dan kegunaan aplikasi pengelola kata sandi closed-source berdasarkan ISO/IEC 25010. *EDUMATIC: Jurnal Pendidikan, Informatika, dan Sains*, *6*(2). https://doi.org/10.29408/edumatic.v6i2.6697

Ahsan, W. B., Hossain, I., Rahman, M. M., Ahmed, T., Sultana, N., Jahan, R., Khan, M. A., & Ehsan, A. M. D. (2024, April 30). Global mobile app accessibility: A comparative study of WCAG compliance across 12 countries. *MetaArXiv*. https://doi.org/10.58947/mxrc-rzkh

Ansori, S., Hendradi, P., & Nugroho, S. (2023). Application of the Design Thinking Method in UI/UX design for the SIPROPMAWA mobile application. *Journal of Information System Research (JOSH)*, *4*(4), 1072–1081. https://doi.org/10.47065/josh.v4i4.3648

Atanacio, M. (2022). Emergency number services rescue mobile app in Sta. Cruz, Laguna. *Journal of Innovative Technology Convergence*, *4*(2), 43–52. https://doi.org/10.69478/JITC2022v4n2a05

Bender-Salazar, R. (2023). Design thinking as an effective method for problem-setting and needfinding for entrepreneurial teams addressing wicked problems. *Journal of Innovation and Entrepreneurship*, *12*, 24. https://doi.org/10.1186/s13731-023-00291-2

Bora, J., Dehingia, S., Boruah, A., Chetia, A. A., & Gogoi, D. (2023). Real-time Assamese sign language recognition using MediaPipe and deep learning. *Procedia Computer Science*, *218*, 1384–1393. https://doi.org/10.1016/j.procs.2023.01.117

Cayme, K. J., Retutal, V. A., Salubre, M. E., Astillo, P. V., Cañete, L. G., Jr., & Choudhary, G. (2024). Gesture recognition of Filipino Sign Language using convolutional and long short-term memory deep neural networks. *Knowledge*, *4*(3), 358–381. https://doi.org/10.3390/knowledge4030020

Chicco, D., Tötsch, N., & Jurman, G. (2021). The Matthews correlation coefficient (MCC) is more reliable than balanced accuracy, bookmaker informedness, and markedness in two-class confusion matrix evaluation. *BioData Mining*, *14*, 13. https://doi.org/10.1186/s13040-021-00244-z

Choo, K., Balan, R., & Lee, Y. (2019, May 2). Examining augmented virtuality impairment simulation for mobile app accessibility design. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems (CHI '19)*. ACM. https://doi.org/10.1145/3290605.3300605

Di Leo, K., Biagetti, G., Falaschetti, L., & Crippa, P. (2025). Microcontroller implementation of LSTM neural networks for dynamic hand gesture recognition. *Sensors*, *25*(12), 3831. https://doi.org/10.3390/s25123831

Doan, T.-N. (2022). An efficient patient activity recognition using LSTM network and high-fidelity body pose tracking. *International Journal of Advanced Computer Science and Applications (IJACSA)*, *13*(8). https://doi.org/10.14569/IJACSA.2022.0130827

Evangelista, C. L. L., Geli, C. J. R., Castillo, M. M. V., & Macabagdal, C. B. G. (2023). Long short-term memory-based static and dynamic Filipino sign language recognition. In *2023 IEEE International Conference on Automatic Control and Intelligent Systems (I2CACIS)* (pp. 235–240). IEEE. https://doi.org/10.1109/I2CACIS57635.2023.10193672

Ezra, D., Mastitz, S., & Rabaev, I. (2024). Signsability: Enhancing communication through a sign language app. *Software*, *3*(3), 368–379. https://doi.org/10.3390/software3030019

Familoni, B. T., & Babatunde, S. O. (2024, March 28). User experience (UX) design in medical products: Theoretical foundations and development best practices. *Engineering Science & Technology Journal*, *5*(3), 129–137. https://doi.org/10.51594/estj.v5i3.975

Feng, S. Y., Gangal, V. P., Wei, J., Chandar, S., Vosoughi, S., Mitamura, T., & Hovy, E. (2021). A survey of data augmentation approaches for NLP. *Findings of the Association for Computational Linguistics: ACL-IJCNLP 2021*, 968–988. https://doi.org/10.18653/v1/2021.findings-acl.84

Gokgoz, Z. A., Ataman, M. B., & van Bruggen, G. H. (2025). If it ain't broke, should you still fix it? Effects of incorporating user feedback in product development on mobile application ratings. *International Journal of Research in Marketing*, *42*(2), 467–486. https://doi.org/10.1016/j.ijresmar.2024.10.004

Huda, B., Sediyono, E., Hananto, A., & others. (2023, November 10). Evaluation quality of e-learning X using ISO/IEC 25010 framework and design thinking approach. In *2023 6th International Conference on Information and Communications Technology (ICOIACT)*. IEEE. https://doi.org/10.1109/ICOIACT59844.2023.10455796

Jarabese, M., Marzan, C., Boado, J., Lopez, R., Ofiana, L., & Pilarca, K. (2021, November 1). Sign to speech convolutional neural network-based Filipino Sign Language hand gesture recognition system. https://doi.org/10.1109/ISCSIC54682.2021.00036

Jha, A., & Pathak, K. (2024). Breaking barriers: Hand gesture vocalizer for the deaf and mute. *JMC Research Journal*, *12*(1), 1–20. https://doi.org/10.3126/jmcrj.v12i1.61623

Joshua, S., Abbas, W., & Kim, S. K. (2023, January 17). Trust components: An analysis in the development of Type 2 Diabetic Mellitus mobile application. *Applied Sciences, 13*(3), 1251. https://doi.org/10.3390/app13031251

Kousis, I., Perikos, I., & Virvou, M. (2022, April 19). Deep learning methods for accurate skin cancer recognition and mobile application. *Electronics, 11*(9), 1294. https://doi.org/10.3390/electronics11091294

Lipsmeier, F., Taylor, K. I., Kilchenmann, T., Wolf, D., Scotland, A., Schjodt-Eriksen, J., Cheng, W. Y., Fernandez-Garcia, I., Siebourg-Polster, J., Jin, L., Soto, J., Verselis, L., Wicks, P., Martin, W. J., Boess, F., & Lindemann, M. (2022, July 15). Reliability and validity of the Roche PD Mobile Application for remote monitoring of early Parkinson's disease. *Scientific Reports, 12*, Article 12081. https://doi.org/10.1038/s41598-022-15874-4

Lugaresi, C., Tang, J., Nash, H., McClanahan, C., Ceze, L., & Taylor, J. (2019). MediaPipe: A framework for building perception pipelines. *arXiv*. https://arxiv.org/abs/1906.08172

Madrid, G., Villanueva, R., & Caya, M. (2022). Recognition of dynamic Filipino Sign Language using MediaPipe and Long Short-Term Memory. IEEE. https://doi.org/10.1109/ICCCNT54827.2022.9984599

Martino, C., & Andry, J. (2020, January 15). Testing aplikasi Business Activity Monitoring pada Internet Service Provider menggunakan ISO 25010. *Jurnal Teknologi Informasi (JTI), 14*(1). https://doi.org/10.33365/jti.v14i1.451

Montefalcon, M. D., Padilla, J., & Rodriguez, R. (2021). Filipino Sign Language recognition using deep learning. In *Proceedings of the 2021 International Conference on Artificial Intelligence and Machine Learning* (pp. 219–225). https://doi.org/10.1145/3485768.3485783

Oliveira, A. D. A., dos Santos, P. S. H., & Eler, M. (2024, November 2). Exploring the influence of software evolution on mobile app accessibility: Insights from user reviews. *Journal of the Brazilian Computer Society, 30*, Article 43. https://doi.org/10.5753/jbcs.2024.4321

Ombai, T. O., Putri, P., Anisya, A., & others. (2024, January 19). Perancangan dan implementasi media pembelajaran gerbang logika berbasis augmented reality menggunakan standar ISO 25010. *Jurnal Teknik Informatika dan Teknologi Informasi (JUTITI), 4*(1). https://doi.org/10.55606/jutiti.v4i1.2944

Pigou, L., Dieleman, S., Kindermans, P. J., & Schrauwen, B. (2018). Sign language recognition using convolutional neural networks. In *Proceedings of the European Conference on Computer Vision (ECCV)* (pp. 572–578). https://doi.org/10.1007/978-3-319-16178-5_41

Rahimian, E., & Sohei. (2020). XceptionTime: Independent time-window XceptionTime architecture for hand gesture classification. In *Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 9054586. https://doi.org/10.1109/ICASSP40776.2020.9054586

Rösch, N., Tiberius, V., & Kraus, S. (2023). Design thinking for innovation: Context factors, process, and outcomes. *European Journal of Innovation Management, 26*(7), 160–176. https://doi.org/10.1108/EJIM-03-2022-0164

Sadueste, A. B., & Masalinto, M. L. (2023, May 27). Accessibility, usability and satisfaction of library users in the hybrid resources and services of academic libraries in the

Province of Albay. *International Journal of Innovative Science and Research Technology (IJISRT)*, *8*(5), 1584–1590. https://doi.org/10.38124/ijisrt/ijisrt23may934

Shah, U.-e.-M., & Chiew, T. K. (2019). A systematic literature review of the design approach and usability evaluation of the pain management mobile applications. *Symmetry*, *11*(3), 400. https://doi.org/10.3390/sym11030400

UP Diliman. (2023, August 9). Liberty Notarte-Balanquit presents Filipino Sign Language research. *Department of Linguistics - UP Diliman*. https://linguistics.upd.edu.ph/news/liberty-notarte-balanquit-presents-filipino-sign-language-research/

Vanacore, A., Pellegrino, M. S., & Ciardiello, A. (2024). Fair evaluation of classifier predictive performance based on binary confusion matrix. *Computational Statistics*, *39*, 363–383. https://doi.org/10.1007/s00180-022-01301-9

Wee, S. Y., Taha, N. M., & Azman, N. A. A. (2024). Behavioral intentions of higher education students toward using gadgets for mobile learning. *International Journal of Academic Research in Progressive Education and Development*, *13*(4), Article 23081. https://doi.org/10.6007/ijarped/v13-i4/23081

Zhang, Y., & Jiang, X. (2024). Recent advances on deep learning for sign language recognition. *CMES - Computer Modeling in Engineering and Sciences*, *139*(3), 2399–2450. https://doi.org/10.32604/cmes.2023.045731

Zulaikha, E., Andansari, D., & Susandari, H. (2023). Empathy-driven innovation: Analysis of five user-centered design thinking case studies. In E. Sari, M. Ghazali, & A. B. Tedjasaputra (Eds.), *Asian HCI Symposium 2023, Asian CHI 2023 - Proceedings* (pp. 1–6). Association for Computing Machinery. https://doi.org/10.1145/3604571.3604572

Zulfa, F., Munawaroh, H., & Rochimah, S. (2020, September). Portability characteristics evaluation of MyITS Mobile using ISO/IEC 25010 quality standard. In *2020 International Seminar on Application for Technology of Information and Communication (iSemantic)* (pp. 398–403). IEEE. https://doi.org/10.1109/iSemantic50169.2020.9234241