HANDSY: ALPHABET SIGN LANGUAGE TUTORIAL MOBILE APPLICATION WITH HAND RECOGNITION

JOHN BERLIN LEONOR

A RESEARCH PAPER SUBMITTED FOR REVIEW
BY THE APPCON 2024 ORGANIZERS
AND JUDGES

APRIL 2024

EXECUTIVE SUMMARY

Handsy: Filipino Sign Language Mobile Application with Hand Gesture Recognition is a mobile application operating within the Android OS that provides an accessible and intuitive platform for people to learn the official sign language of our country, Filipino Sign Language, both in the supportive premises of their school and the comfort of their homes.

Its objectives are that it would be able to integrate relevant features such as visualized lessons and gesture recognition through the device's in-built camera powered by a light-weight AI model that can operate reasonably even in low-spec devices.

Upon completion, it has been determined that a larger data set will be necessary to ensure the effectiveness of the gesture recognition feature of the application, and the improvement of it by adding phrases lessons with gesture recognition support would further improve the application in the perspective of the end users.

TABLE OF CONTENTS

LIST OF FIGURES

CHAPTER 1

INTRODUCTION

Languages that convey meaning through the visual-manual modality instead of spoken words are known as sign languages (commonly referred to as signed languages). In sign languages, lexical distinction, grammatical structure, adjectival or adverbial content, and discourse functions are all shown through manual articulation in combination with non-manual markers such as postures or movements of the body, head, eyebrows, eyes, cheeks, and mouth that are used in various combinations.

In the Philippines, a legislation or Republic Act has been passed designating Filipino Sign Language (FSL) as the official sign language of the government in all deaf-related operations and mandating its use in educational institutions, the media, and workplaces, officially recognizing it as the country's sign language.

The integration of Filipino Sign Language (FSL) into the Philippine education curriculum faces significant challenges, primarily due to its relatively recent documentation and the lack of widespread recognition by the government and society at large. The Filipino Deaf community, despite being vibrant and dynamic, remains marginalized, with FSL largely unrecognized in various domains such as schools, courtrooms, workplaces, hospitals, and mass media. This lack of recognition extends to the community itself, where both deaf and hearing Filipinos may disregard FSL as an authentic linguistic entity. The absence of a comprehensive language policy for sign language use further exacerbates these issues

Handsy: Alphabet Sign Language Tutorial Mobile Application with Hand Recognition is a mobile application designed to help people learn the country's official sign language through visualized lessons and an in-built camera that recognizes the user's gestures through the power of Artificial Intelligence. It is built to provide a complementary tool for people to be introduced and learn the sign language marginalized within our society due to the inaccessibility of its learning materials for ordinary people.

Project Context

The application is developed to address the main concern which is the development of an Android application that serves as a sign language tutorial for the Filipion Sign Language. The development of the application is built in consideration of the needs of the respondents, which are the SPED students of the Moises Padilla Elementary School, yet is built to cater everyone.

The project development begun on February 2023 and continued until April 2024 with the collaboration of the SPED teachers whose knowledge and expertise will prove invaluable for the completion of the capstone research.

Purpose of Description

Handsy: Alphabet Sign Language Tutorial Mobile Application with AI Hand Recognition or Handsy, in short, provides the following purposes for the following individual within our respondents.

The Teachers. It assists our teachers in the teaching of the Filipino Sign Language during and after classes as a supplementary tool of learning for their special needs students studying under the SPED program of DEPED.

The Users. It assists the students to learn the Filipino Sign Language through a visually appealing and interactive application accessible both in the premises of the school and the comfort of their homes.

Future Researchers. This provides reliable basis for future research efforts and endeavours towards the improvement of AI models for better gesture recognition performances on mobile phones with consideration of hardware limitations.

Objectives

The following below are the aims to that which the capstone project aims to accomplish:

1. To create a mobile application that is accessible to users especially for those who only owns low-spec mobile devices.
2. To provide interactive visual hand signal lessons for alphabets and words.
3. To provide AI functionality for hand recognition to read different hand signs (alphabets) shown by the user to the mobile camera.
4. To ensure that the application is built as an offline-access native Android application.

Scope and Limitations

The application will be created for the use of mobile devices of the Android OS platform as it presents the largest share of users that can be reached by the project, adhering with one of its objectives of user accessibility.

The code base of the Android application is built upon well-established libraries and tools such as the modern Jetpack Compose and the preferred language of choice for

Android development, Kotlin, to allow ease of development and ensure future development efforts can be done with ease.

CHAPTER II

REVIEW OF RELATED LITERATURE AND SYSTEMS

This Chapter discusses relevant literature and systems that have been published in book, journal, newspaper, and even online sources relevant towards the research and development of the project.

Foreign Literature

Machine Learning with TensorFlow (https://dl.acm.org/doi/book/10.5555/3217359)

Machine Learning with TensorFlow is a book that teaches readers about machine learning concepts and how to code with TensorFlow and Python. TensorFlow is a library created by Google for large-scale machine learning, which simplifies complex computations by representing them as graphs. The book covers classic prediction, classification, and clustering algorithms, as well as deep learning concepts such as autoencoders, recurrent neural networks, and reinforcement learning.

Tensorflow: Learning Functions At Scale (https://doi.org/10.1145/2951913.2976746)

TensorFlow is a machine learning system that works on a large scale and in diverse environments. It uses dataflow graphs with mutable state, allowing graph nodes to be assigned to different machines and devices within each machine. TensorFlow is used for a wide range of applications, with a focus on training and inference using deep neural networks. Its applications include speech recognition, computer vision, robotics, information retrieval, and natural language processing.

Machine Learning: Trends, Perspectives, And Prospects (https://www.science.org/doi/10.1126/science.aaa8415)

Machine learning is a field that focuses on building computers that can learn and improve on their own through experience. It is a rapidly growing area that intersects computer science and statistics, and is a key component of artificial intelligence and data science. The advancement of new learning algorithms and theory, along with the increase in availability of online data and low-cost computation, has led to significant progress in the field.

What Is Machine Learning? (https://doi.org/10.1002/9781119815075.ch18)

Over the last 60 years, pioneers in the industry have advanced machine learning (ML), resulting in various algorithms for different output requirements. ML algorithms generally fall under supervised or unsupervised learning. ML is widely used in software to enhance user experience and enable robots to acquire skills, such as object placement and grasping, through automated or human-guided learning.

Machine Learning: Algorithms, Real-World Applications and Research Directions (https://doi.org/10.1007/s42979-021-00592-x)

In the age of Industry 4.0, data from various sources like IOT, Cyber Security, mobile, social media, and health is abundant. The key to analyzing this data and developing smart applications is the knowledge of AI, particularly machine learning (ML).

There are various types of ML algorithms such as supervised, unsupervised, semi-supervised, and reinforcement learning, which can be used to enhance the intelligence of applications. Deep learning, a type of ML, can intelligently analyze data on a large scale. This paper provides a comprehensive view of these ML algorithms and their applicability in real-world application domains such as cybersecurity, smart cities, healthcare, e-commerce, agriculture, etc. The paper also highlights challenges and potential research directions. The study aims to be a reference point for both academia and industry professionals, as well as decision-makers, in various real-world situations and application areas.

LOCAL LITERATURE

Using Machine Learning Approaches to Explore Non- cognitive Variables Influencing Reading Proficiency in English among Filipino Learners (https://doi.org/10.3390/educsci11100628)

The Philippines ranked last in reading proficiency among all countries/territories in the PISA 2018, with only 19% meeting the minimum standard. To understand the factors contributing to this, machine learning approaches were used to identify the variables that best predict low vs. higher reading proficiency. The best model was derived using support vector machines, with 81.2% accuracy. The 20 variables with the highest impact in the model included home-related resources and socioeconomic constraints, learning motivation and mindsets, classroom reading experiences with teachers, reading self-beliefs, attitudes, and experiences, and social experiences in the school environment. The results highlight the need for a systems perspective to address poor proficiency, requiring interconnected interventions beyond the classroom.

E-tutor for Filipino Sign Language (https://ieeexplore.ieee.org/document/7581584)

This study focuses on the development of a mobile learning application that provides an interactive experience for deaf users. The app offers tutorial lessons on Filipino Sign Language (FSL), which is the natural sign language of the Philippines. The app has three modules: dictionary, illustrations, and assessment, and it covers fifty essential FSL signs through human-model video illustrations and corresponding written and visual memory aids. The app aims to help deaf and non-deaf users become familiar with commonly used signs in the environment.

Syntactic Patterns of Interrogative and Negation Constructions of Filipino Sign Language Consultants in the Philippines (https://www.proquest.com/openview/ef406cd77544f3d3acaf41092ce0f44f/1?pq-origsite=gscholar&cbl=18750&diss=y)

This project aims to investigate the syntactic position patterns in Filipino Sign Language (FSL) interrogatives and negation constructions used by Deaf Filipino consultants. The data collected consists of monologues and dialogues from 55 consultants, aged 18 to 60, in response to videos and written questions. The preliminary analysis shows that FSL has similar linguistic structures to other sign languages, with word order, manual question particles, nonmanual question signs, and scopes. The study compares data from different regions in the Philippines and shows variation in syntactic patterns. The findings suggest that further research with more consultants is needed. This study is important for policy in Deaf education and to provide more publications related to lexicon and sociolinguistics for FSL.

Utilizing Artificial Intellegence to Detect Deceit in Videos of Filipinos (https://archium.ateneo.edu/theses-dissertations/522)

Previous studies on AI models that detect lies in videos have mainly used videos of Caucasians, which may limit their generalizability to non-Caucasians due to culture-specific behaviors. This study builds a machine learning model and achieves a ROC-AUC of 0.7156 with 6-fold cross-validation. It finds that audio features are the most effective modality, followed by video features, while micro-expressions alone cannot detect lies better than chance. The study suggests that the universal-cue and specific-discrimination hypotheses also apply to AI models, but further research is needed to confirm this.

FOREIGN SYSTEM

Deep Learning-Based Sign Language Recognition System For Static Signs (https://doi.org/10.1007/s00521-019-04691-y)

Sign language for communication is efficacious for humans, and vital research is in progress in computer vision systems. The earliest work in Indian Sign Language (ISL) recognition considers the recognition of significant differentiable hand signs and therefore often selecting a few signs from the ISL for recognition. This paper deals with robust modeling of static signs in the context of sign language recognition using deep learning-based convolutional neural networks (CNN). In this research, total 35,000 sign images of 100 static signs are collected from different users. The efficiency of the proposed system is evaluated on approximately 50 CNN models. The results are also evaluated on the basis of different optimizers, and it has been observed that the proposed approach has achieved the highest training accuracy of 99.72% and 99.90% on colored and grayscale images, respectively. The performance of the proposed system has also been evaluated on the basis of precision, recall and F-score. The system also demonstrates its effectiveness over the Earlier Works In Which Only a Few Hand Signs Are Considered For Recognition.

Indian Sign Language Translator Using Gesture Recognition Algorithm (https://doi.org/10.1109/CGVIS.2015.7449921)

Sign Language is a natural language which deaf community uses for communication. Sign Language (SL) is a subset of gestures or signs made with fingers, hands, arms, eyes, and head, face etc. Each gesture in SL has a meaning assigned to it. Understanding SL is nothing but understanding the meaning of these gestures. There exists a problem in communication when a person who completely relies on this gestural SL for communication tries to converse with a person who does not understand the SL. Every country has its own developed SL. In India, this language is called as "Indian Sign Language (ISL). This paper aims to develop an algorithm that will translate the ISL into English. This paper has implemented a system named as "Indian Sign Language (ISL) Translator using Gesture recognition algorithm". The system translates gestures made in ISL into English. The gestures that have been translated include numbers, alphabets and few phrases. The algorithm first performs data acquisition, then the pre-processing of gestures is performed to track hand movement using a combinational algorithm, and recognition is done using template matching. The database used for implementation has been self-created and includes total 130,000 videos; out of which 72,000 videos were used to create the system database and remaining 58,000 videos have been tested for checking the performance of the system. The accuracy of this system is as high as 97.5%.

An Automatic Arabic Sign Language Recognition System (ArSLRS)
(http://dx.doi.org/10.1016/j.jksuci.2017.09.007)

Sign language recognition system (SLRS) is one of the application areas of human computer interaction (HCI) where signs of hearing-impaired people are converted to text or voice of the oral language. This paper presents an automatic visual SLRS that translates isolated Arabic words signs into text. The proposed system has four main stages: hand segmentation, tracking, feature extraction and classification. A dynamic skin detector based on the face color tone is used for hand segmentation. Then, a proposed skin-blob tracking technique is used to identify and track the hands. A dataset of 30 isolated words that used in the daily school life of the hearing-impaired children was developed for evaluating the proposed system, taking into consideration that 83% of the words have different occlusion states. Experimental results indicate that the proposed system has a recognition rate of 97% in signer-independent mode. In addition to, the proposed occlusion resolving technique can outperform other methods by accurately specify the position of the hands and the head with an improvement of 2.57% at $\tau = 5$ that aid in differentiating between similar gestures.

Data Gloves for Sign Language Recognition System
(https://www.ijcaonline.org/proceedings/ncetact2015/number1/20979-2009)

Communication between deaf-dumb and a normal person have always been a challenging task. About 9 billion people in the world come into this category which is quite large number to be ignored. As deaf-dumb people use sign language for their communication which is difficult to understand by the normal people. This paper aims at eradicating the

LOCAL SYSTEM

Sign Language Number Recognition (https://doi.org/10.1109/NCM.2009.357)

The purpose of the Sign Language Number Recognition System is to provide a practical solution to the current issues in handshape recognition in the deaf community. The system uses a multi-color tracking algorithm to extract important features from video input of 5000 Filipino Sign Language number files, and then learns and recognizes the numbers using Hidden Markov Model. The color-coded gloves used in the system are designed to be more efficient than existing gloves in other research. The overall accuracy of the recognizer is 85.52%, indicating that the system has the potential to be an effective tool for recognizing sign language numbers in practical applications.

Filipino Sign Language Recognition using Manifold Learning (https://doi.org/10.1109/TENCON.2012.6412231)

The purpose of this study is to explore the use of manifold learning as a method for recognizing visual signs in Filipino Sign Language. The study uses Isomap, a non-linear manifold learning algorithm, to derive a reference manifold from a training set of visual signs. The signs are then projected onto this reference manifold and transformed into trajectories which are compiled into a library. The study aims to achieve recognition of unknown signs by computing their manifold trajectories and comparing them with the trajectories found in the library using Dynamic Time Warping or Longest Common Subsequence Similarity Matching. The study reports recognition rates exceeding 80% for individual uninflected signs.

Sign to Speech Convolutional Neural Network-Based Filipino Sign Language Hand Gesture Recognition System (http://dx.doi.org/10.1109/ISCSIC54682.2021.00036)

The purpose of this study is to develop a real-time Filipino sign language hand gesture recognition system that can help deaf-mute people. The researchers used a manually gathered dataset and image pre-processing techniques to train a Convolutional Neural Network model to recognize 20 different gestures. The study shows that retraining the model with all layers using an imbalanced dataset was more effective, achieving up to 95% accuracy. The researchers also developed a sign language recognition application using the Rapid Application Development model and evaluated its usability by target users. The study aims to provide an effective and efficient solution for sign language recognition to help deaf-mute people communicate more easily.

Filipino Sign Language Recognition using Deep Learning (https://doi.org/10.1145/3485768.3485783)

This project aims to use computer vision and Convolutional Neural Network (CNN) ResNet architecture to build an automated Filipino Sign Language (FSL) recognition model. The purpose is to bridge the communication gap between the deaf community and the hearing majority in the Philippines. The project uses static images of Filipino number signs as its dataset and has achieved a validation accuracy of 86.7% with the fine-tuned ResNet-50 model. The future work involves implementing real-time FSL recognition and collecting more data to enable recognition of Filipino alphabets, basic phrases, and common greetings.

SimboWika: A Mobile and Web Application to Learn Filipino Sign Language for Deaf Students in Elementary Schools (https://doi.org/10.1109/R10-HTC49770.2020.9357056)

The communication between deaf and hearing people is hard especially when you consider the lack of technology and knowledge on language that the deaf people use. The population of the Philippines is increasing and so are the deaf people. The communication gap between these people is also expanding. English is the commonly use language by most people with hearing disabilities. American Sign Language (ASL) is considered as a renowned sign language in the Philippines but this is not the official language of the country. Most Filipinos are not aware of the existence of Filipino Sign Language (FSL). This is an American Sign Language-inspired and designed specifically for Filipinos. Deaf students in one elementary school in Manila were confused in learning Tagalog words because the sign language that had been established back then was American Sign Language (ASL). With this drawback, the researchers aimed to develop “SimboWika”, a mobile application to help the deaf students in elementary schools to learn Filipino Sign Language. It provides illustrations to practice FSL and assess student's learning. The teachers can also keep track of the student's progress using its web application. Based on findings, the SimboWika app got a Very Satisfactory result from the users and technical evaluation.

CHAPTER III

METHODOLOGY

In this chapter, the methodology adhered by the Researcher and developer used in the development project will be discussed here.

Requirements Specification

These are the (hardware/software), development environment, design and implementation features, and general application features the project has within the development phase and its final p roduct.

Hardware Requirements

- Android 8.1 OS (Oreo) and later versions. Versions below the specified one will cause fatal run-time operational problems due to library incompatibility, rendering app unusable.
- Device should have camera

Software Requirements

The application does not require other software to run properly within the user's Android OS device.

Development Environment

The Android application codebase is built upon Kotlin programming langugage and modern Jetpack Compose libraries. It is developed with the use of Android Studio Iguana (2023.2.1) utilizing Java Software Development Kit 17. The codebase is then compiled and built with the use of the Gradle Build tool, the standard for Android development.

The model training used for gesture recognition utilizes the MediaPipe Solutions framework as it provides pre-trained models that can be fine-tuned for our use case, which is for our case is the Filipino Sign Language alphabet gestures. The MediaPipe Model Maker were used for the model training together with limited use of TensorFlow and Keras. The implementation of the gesture recognition within the application is developed with the use of MediaPipe Tasks API for Android.

The project dependencies of the application were taken care of the Gradle itself, and its proper versioning is handled by a Central Dependency Management file within the codebase for ease of maintenance and development.

Design Constraints

The Handsy application user interface and user experience is built based upon the concepts and principles of the Material 3 Design theme as its design language of choice due to it being the standard of modern Android development.

Developer Workflow

This developer work flow refers to the sequence of activities that we followed to build, test, and integrate each of the unique structural components of the application until it culminates into one single entity.

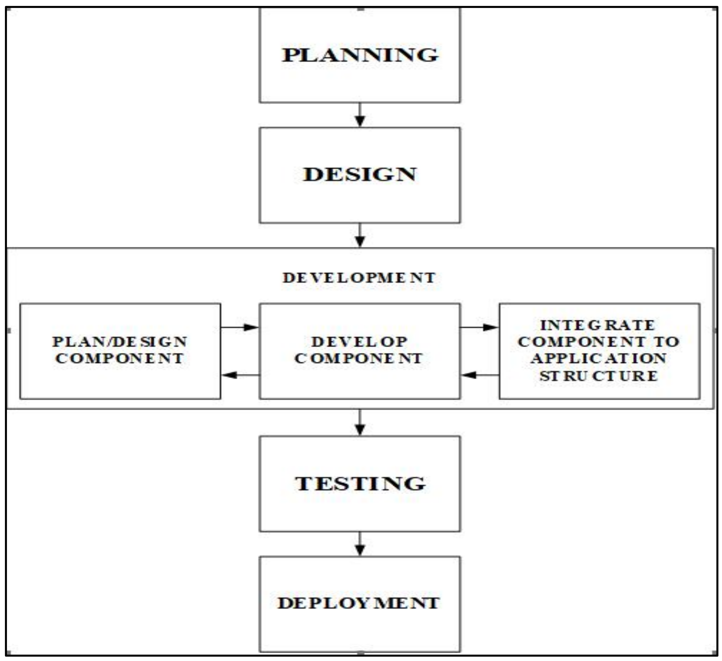

Figure 1: Developer Workflow

As properly illustrated above, the development phase of the application follows a sequential flow, wherein each component are individually designed, developed, and integrated to the component.

This means that for example the camera screen will not be developed until the lessons list screen of the application is not developed completely.

Development

Handsy as an Android project utilizes the concepts and principles declared within the Modern Android Development skills and documentation. To understand the architectural design of the application's codebase, it will be discussed within this section.

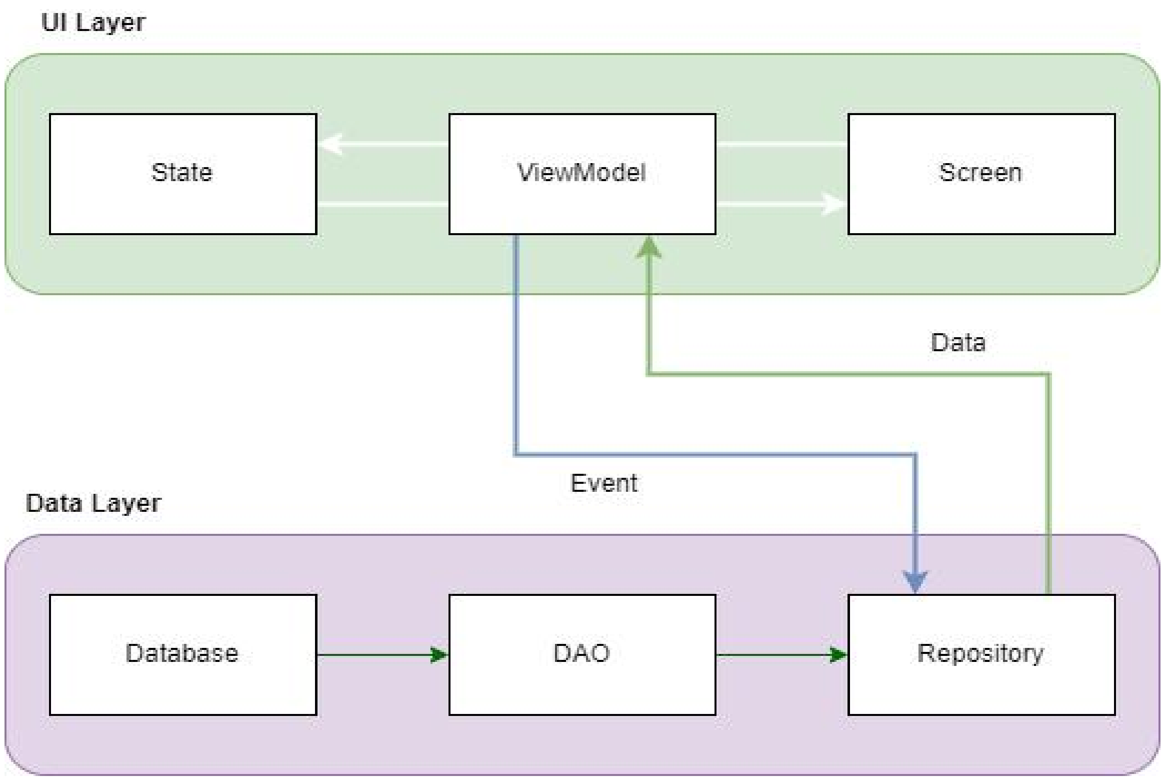

Figure 2. Architectural Overview of the Application

The overall architecture is designed as visualized in Figure 2. The application is divided into two layers - the UI layer, and the Data Layer. The UI Layer is responsible in providing the functionality to display data for the user to see and interact with, while the Data Layer is responsible in the storage and manipulation of the data within the database.

It is important to note that the database of the application does not contain any information relevant to the organization's function, but instead the information necessary for display of the data. The layers is separated with each other, and can be tested individually.

The application is built upon the concepts and principles found within the Modern Architectural Design (MAD) provided by the official documentations as provided by Google. This is adhered to by the developer in order to ensure that the application performs as well as other industry - grade applications of similar performance requirements.

The visual design language is based upon the principles founded within the Material 3 Design language developed by Google, to ensure user accessibility within the user interface and user experience aspect of the mobile application.

The Artificial Intelligence capabilities of the application is built using the MediaPipe Solutions framework. It provided ready - made solutions that can be built upon by developers for specific use cases as required by the application or project being worked on.

The developer built the application with compliance and adherence to proper industry standards, such as clean code principles and proper design pattern implementation to ensure the prototype meets its objective while providing seamless performance and a scalable codebase.

The UI Layer is designed in such a way that each component within it specifically the State, ViewModel, and Composable Screens has only one role to comply with and adheres to the principle of UnidirectionalFlow (UDF) in the overall operation of the layer. This means that while the Composable Screen provides the UI to which the user can interact with the data, it does not contain any codebase for the logic of how to display this data. The ViewModel component is the one responsible for the logic of the UI in which how data is displayed for the user while the State component is responsible as the storage for the information necessary for bothl the ViewModel and Composable screen to work on.

The Data Layer operates in a similar way. Each component operates for one role only with Database concerned only in the storage of the data, the DAO as the interface to which the application can manipulate or access this data, and the Repository operating as a higher-level component to utilize the DAO interface to transform such data into a format usable by the UI Layer.

Within the application, the user can navigate between screens to work on different features offered by the application. This is achieved by the Compose Navigation Library to which the application's navigation component is based upon.

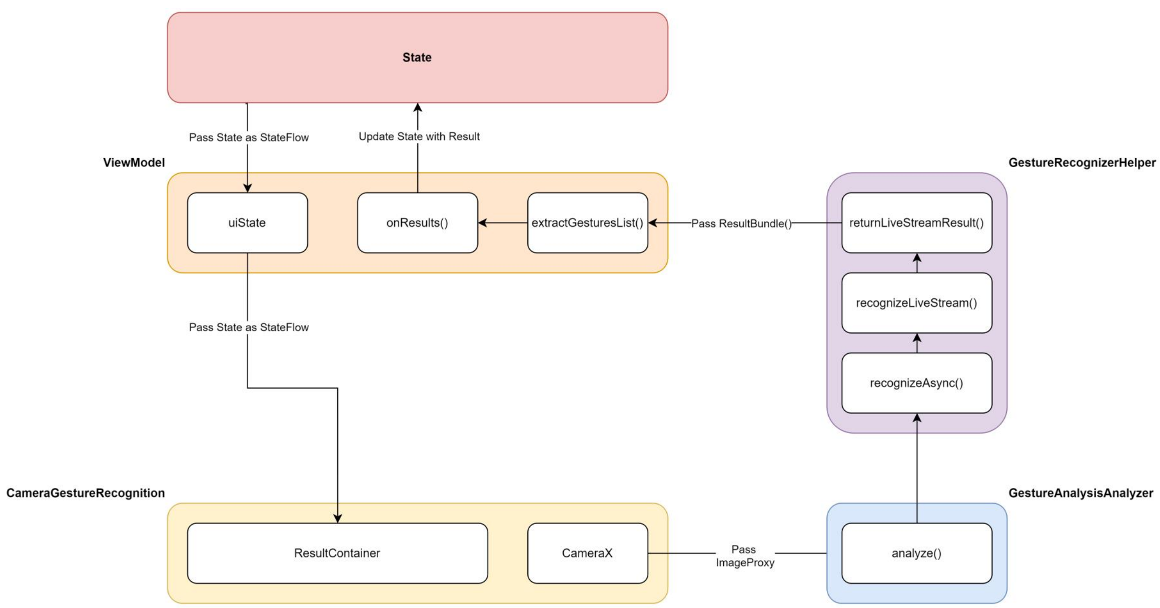

Figure 3. Gesture Recognition Flow

The figure above shows the overall logic of the gesture recognition feature of the application.

To begin with, the CameraX component contained within the CameraGestureRecognition Composable provides an ImageProxy - an interface that handles image frames captured by the camera - to an ImageAnalysis.Analyzer named GestureAnalysisAnalyzer to begin analysis and recognition.

The GestureAnalysisAnalyzer serves only as a bridge between the CameraX Image Analysis use case and the actual Kotlin class that handles recognition using MediaPipe Tasks API, and its function analyze() passes this ImageProxy to the GestureRecognizerHelper for recognition.

The GestureRecognizerHelper is responsible for the actual role of transforming the ImageProxy to readable format for the model to interpret with. It utilizes the MediaPipe Tasks API to achieve this. Its function recognizeLiveStream() is responsible for performing transformation and manipulation of the ImageProxy into a readable MPImage compatible for the AI model to recognize. It passes this MPImage to recognizeAsync() which runs this format through the model for recognition. returnLiveStreamResult() extracts the result returned by the AI model and passes this as a ResultBundle to the ViewModel for transformation into something the Screen UI can interpret.

The ViewModel's onResult() transform the ResultBundle into different information with the help of extractGesturesList() and update the State of the Composable screen. Note that the actual gesture is a collection of a specific type that requires different operations before a String format info can be extracted from it. This is due to the fact that the raw result is actually a list of detected gestures and each gesture contains first an outer list of different hands detected, and its inner list containing the actual information for the gesture.

The State, adhering to the principle of Modern Android Development of Unidirectional Flow is passed as a StateFlow through the ViewModel to the CameraGestureRecognition Composable screen to be used to update the ResultContainer within it.

CHAPTER V

CONCLUSIONS AND RECOMMENDATION

This chapter provides information regarding the final product's conclusive performance.

Conclusion

The application performed well and was able to recognize most of the letters within the alphabet of the Filipino Sign Language, with exception of a few letters that are visually similar such as E with S, and M with N which the model struggles to properly identify. Gesture - related letters J and Z that has to be signed in a specific manner of steps cannot be identified by the model since it can only detect single frames.

Proper grammar cannot be taught also since the lessons provided is limted with the use of the alphabet of the Filipino Sign Language.

Reception with a few test users are satisfactory regardless due to the innovation provided the application.

Recommendation

The development of Handsy proved difficult due to the fact that resources are limited and expertise from sign language experts are not utilized properly, relying mostly from online references. It is recommended that further developments must integrate with proven and licensed experts within the field to ensure the effectiveness of its teaching.

It is further recommended to integrate achieved progress by other projects rather than working from scratch, such as developments achieved by a well-known research under the title of SimboWika. This would allow the application to be more effective towards its purpose rather than start from scratch in developing a working product if there are already proven results achieved by other initiatives.