Department of Software Engineering and Information Technology

Graduation Project 2023

Video Chat Application with Bilateral Arabic Sign Language Translation

Group Members:

Alaa Osama Yahya

Alaa Tarek Abbas

Farah Muhammed Hamdy

Haidy Ahmed Ibrahim

Nathalie Nader Nabil

Supervised by:

Dr. Mohamed Taha

Acknowledgement

We would like to thank Dr. Mohamed Taha and Dr. Soha Safwat for their support, supervision, and patience with us as well as T. A Rahma Amin and T. A. Ghada Adel for always being willing to help and giving advice.

Abstract

Successful communication requires the efforts of all people involved in a conversation. Communication with a hearing disability can be challenging if one doesn't know sign language. They always need someone to translate for them [14]. Sign language allows them to learn, work, access services, and be included in their communities, but it also limits their communication with those who do not know sign language. Out of the Middle East’s 350 million people, over 11 million have disabling hearing loss, comparable to other regions. But services for deaf citizens have not kept pace. The dearth of sign language interpreters is partly a problem of inadequate recruitment and training, and partly a reflection of broader obstacles for deaf citizens [13]. By 2050 nearly 2.5 billion people are projected to have some degree of hearing loss and at least 700 million will require hearing rehabilitation [15].

In developing countries, children with hearing loss and deafness often do not receive schooling. Adults with hearing loss also have a much higher unemployment rate. Among those who are employed, a higher percentage of people with hearing loss are in the lower grades of employment compared with the general workforce [15]. Everyone should be able to access information equally and communicate freely.

In this document, we propose a video chat application with a bilateral Arabic sign language system to translate the sign language of the impaired person into Arabic text and translate the Arabic language voice into sign language using an avatar to simulate the sign, so hearing impaired people can communicate without being limited by other people's knowledge of sign language. To apply the system, we developed our dataset containing 30 Arabic words, 100 videos for each word, at the beginning to train and test the model. To get high accuracy, we will combine Convolutional Neural Network (CNN) and Long Short-Term Memory (LSTM). They will be trained on the key points extracted from each frame using mediapipe.

Table of Contents

List of Figures...5

Chapter 1...6

Introduction……………………………………………………………………………6

1.1 Motivation……………………………………………………………………..6

1.2 Problem Definition…………………………………………………………… 6

1.3 Objective………………………………………………………………………7

1.4 Document Organization……………………………………………………....7

Chapter 2.. 9

Background……………………………………………………………………………9

2.1 Project Terminologies and Concepts……………………………………...9

2.1.1 Deep learning………..…………………………………………………..9

2.1.2 Convolutional neural network (CNN)…………………………………...9

2.1.3 Long shrot-term memory (LSTM)……………………………………..…..9

2.1.4 Mediapipe……………………………………………………………………10

2.2 Related Work……………………………………………………………..10

2.3 Project Decription……………………………………………………...11

2.4 Functional Requirements………………….………………………………..12

2.5 Non-functional Requirements …………………..………………………..13

2.6 System Users…………………………………………………………...13

A. Intended Users.....……………………………………………………..13

B. User Characterisitics.………………………………………………….13

Chapter 3...14

Analysis and Design……………………………………………………………………14

3.1 System Overview…………………………………………………..…..14

3.1.1Use case Diagram…………………………………………………… .14

3.1.2 Activity Diagram…………………………………………………..…… 15

3.1.3 Sequence Diagram…………………………………………………….16

3.1.4 State Diagram………………………………………………………….17

3.1.5 Class Diagram…………………………………………………………18

3.1.6 FLow Charts...…………………………………………………………19

3.1.6.1 Register and Login…………………………………………………..19

3.1.6.2 Add or Delete Contact……….……………………………………..19

3.1.6.3 Start Video Call..……………………………………………...…..20

3.1.6.4 Answer Video Call.....…………………………………………………21

3.1.6.5 Edit Profile………..……………………………………………………21

3.1.6.6 Search Contact…........………………………………………………22

3.1.6.7 Logout……………....……………………………………………………22
3.1.6.8 video call……………....………………………………………………23
3.1.7 Block Diagram……………....…………………………………………..…24
3.1.8 System Architecture……………...……………………………………..…24
Application design interface 25
References... 29

List of Figures

Figure **Page**

FIGURE 1 USE CASE DIAGRAM 14
FIGURE 2 ACTIVITY DIAGRAM 15
FIGURE 3 SEQUENCE DIAGRAM 16
FIGURE 4 STATE DIAGRAM 17
FIGURE 5 CLASS DIAGRAM 18
FIGURE 6 REGISTER AND LOGIN FLOW CHART 19
FIGURE 7 ADD OR DELETE CONTACT FLOW CHART..20
FIGURE 8 START VIDEO CALL FLOW CHART..20
FIGURE 9 ANSWER VIDEO CALL FLOW CHART 21
FIGURE 10 EDIT PROFILE FLOW CHART..22
FIGURE 11 SEARCH CONTACT FLOW CHART 22
FIGURE 12 LOGOUT FLOW CHART 23
FIGURE 13 VIDEO CALL FLOW CHART 23
FIGURE 14 BLOCK DIAGRAM 24
FIGURE 15 SYSTEM ARCHITECTURE 24
FIGURE 16 LOGIN AND REGISTER PAGE 25
FIGURE 17 REGISTER PAGE 25
FIGURE 18 LOGIN PAGE..26
FIGURE 19 SEARCH AND ADD CONTACT 26
FIGURE 20 NAVIGATION DRAWER 27
FIGURE 21 EDIT PROFILE PAGE 27
FIGURE 22 CONTACT INFO PAGE..28
FIGURE 23 VIDEO CALL PAGE 28

Chapter 1

Introduction

1.1 Motivation

Sign language is the language used by deaf people to communicate with one another and with common people. This enables them to be involved in the traditional educational and pedagogic process. An interpreter is typically required when an ordinary person wishes to communicate with a deaf person. Visual sign language recognition is a dynamic area of research in computer vision [16]. The law of multiple proportions states that 5% of the populations who speak Arabic struggle with hearing impairments which is a significant number. This demonstrates why Arabic Sign Language is significant (ArSL).

Arabic sign language is divided into two parts: The first part is a complete language where each word is represented by a sign. In the second part, the Arabic sign language alphabets are used to represent each letter of the Arabic alphabet (ArSLA) [17]. Due to the importance of Arabic sign language, which was previously discussed, it is an important area of study. It helps to eliminate barriers between the deaf and the community, so researchers and practitioners have become interested in the challenge of creating Arabic sign language recognition systems. However, all the current solutions are just basic translation apps either for words or letters, but that still limits hearing-impaired community's interactions and social life experiences. A functional sign language recognition (SLR) system can provide an opportunity for a deaf person to communicate with non-signed persons without the need for an interpreter. It can be used to produce speech or text that makes the mute more autonomous. As a result, we have to think of developing a video chat application with real-time bilateral Arabic sign language translation. This will give hearing-impaired people a better opportunity to express themselves and communicate with the world around them. In the meantime, spoken people would be able to naturally communicate with them as there would be an avatar translating what they are saying into signs.

1.2 Problem Definition:

Singing has always been an important part of human communication. Deaf people use it to communicate with one another, but the problem arises when a deaf person tries to communicate with a normal person who does not understand sign language. This limits deaf people's interactions and social life experiences because they can only communicate with a few people. According to WHO, currently more than 1.5 billion people (nearly 20% of the global population) live with hearing loss and 11 million of the 350 million people living in the Arab world suffer from hearing loss [13]. That's why researchers have tried to concentrate more on this issue and look for a solution that would allow the use of AI and deep learning to help facilitating the interaction of the deaf community with those who do not understand sign language. Despite their best efforts most of sign language recognition research is simulation-based and does not work for real-time or it is just a basic translation apps which still not good enough for effective communication these interpreter’s performance can be affected by small differences such as the performer himself as the interpreter can be designed to be user-dependent or user-independent. In the first case, the interpreter depends on the person while in the latter one, the user is not a problem anymore.

This is due to so many challenges that face researchers starting from finding a good dataset that considered one of the biggest challenges. To collect a good and clean dataset, you have to overcome these challenges:

- Viewpoint Difference: as people can take the same sign with different hand kinematics and poses.
- Environment: The background, lighting, landmarks, and other elements can exist in the video all these factors may affect the quality of the dataset.
- Facial Expressions: A facial expression can be part of a sign. The system may be implied by the face, which may have accessories like earrings and glasses [18].

Researchers have tried to solve those challenges using mainly two approaches sensor-based or vision-based. In sensor-based, the user wears sensors, such as colored or special gloves and a motion capture system captures the sign, but this method has different drawbacks as it was ineffective in real-world scenarios because it requires the user to wear such sensors that depend on the continuous power supply, wires, and other requirements. In the vision-based approach, the system captures videos using mobile or laptop camera, there is no need for any external camera or device then, predicting the processed data using machine learning and deep learning techniques [18]. Convolutional neural networks, artificial neural networks, hidden-Markov models, and recurrent neural networks (RNN) are some examples of these techniques. This method is obviously low cost as it is not dependent on any other device. That's why we decided to continue with this approach.

1.3 Objectives

Assistive technologies for the Deaf community in Arab countries are numerous and limited. The complexity of Arabic and ArSL is the main reason for this limitation. Arabic language is much wealthier than English and is considered as one of the most complex natural language. As for ArSL, you can find a standard dictionary with 1600 of the basic and most common signs [19]. Nevertheless, each country has its own modified version of this dictionary that holds new signs for its colloquial language and modified signs for most of the existing words. Even more, you can find for the same word a different sign in distinct cities of the same country. As a conclusion, deaf individuals in developing Arabic countries can communicate with vocal people only through an ArSL human interpreter who is very hard to find and in all the cases will break the conversation privacy.

Our goal is to develop a video chat application with real time bilateral Arabic sign language translation to fill the gap between the deaf and vocal people and help them communicate with each other.

To reach our goal, we are expecting to deliver the following:

- A new dataset of 30 words with 100 videos per word and with the possibility to add more.
- A deep learning model trained on this dataset with a high accuracy rate.
- A video chat application that uses this model to translate their signs and helps deaf people communicate easily with vocal people.

1.4 Document Organization

- Chapter 2 introduces:
    - Project terminologies and concepts

- The related work done.
- A detailed description of the project.
- Functional and non-functional requirements.
- System users

- Chapter 3 introduces:
    - System overview

Chapter 2

Background

2.1 Project Terminologies and Concepts:

In this section we will clarify the important concepts used in this document.

2.1.1 Deep learning:

Deep learning is a subset of machine learning, which is essentially a neural network with three or more layers. These neural networks attempt to simulate the behavior of the human brain—albeit far from matching its ability—allowing it to “learn” from large amounts of data. While a neural network with a single layer can still make approximate predictions, additional hidden layers can help to optimize and refine for accuracy [20].

Deep learning drives many artificial intelligence (AI) applications and services that improve automation, performing analytical and physical tasks without human intervention. Deep learning technology lies behind everyday products and services (such as digital assistants, voice-enabled TV remotes, and credit card fraud detection) as well as emerging technologies (such as self-driving cars).

2.1.2 Convolutional neural network (CNN):

A Convolutional Neural Network (ConvNet/CNN) is a Deep Learning algorithm that can take in an input image, assign importance (learnable weights and biases) to various aspects/objects in the image, and be able to differentiate one from the other. The pre-processing required in a ConvNet is much lower as compared to other classification algorithms. While in primitive methods filters are hand-engineered, with enough training, ConvNets can learn these filters/characteristics.

The architecture of a ConvNet is analogous to that of the connectivity pattern of Neurons in the Human Brain and was inspired by the organization of the Visual Cortex. Individual neurons respond to stimuli only in a restricted region of the visual field known as the Receptive Field. A collection of such fields overlaps to cover the entire visual area [21].

2.1.3 Long short-term memory (LSTM):

LSTM used in the field of Deep Learning. It is a variety of recurrent neural networks (RNNs) that are capable of learning long-term dependencies, especially in sequence prediction problems. LSTM has feedback connections, i.e., it is capable of processing the entire sequence of data, apart from single data points such as images. This finds application in speech recognition, machine translation, etc. LSTM is a special kind of RNN, which shows outstanding performance on a large variety of problems [22].

2.1.4 Mediapipe:

MediaPipe is a Framework for building machine learning pipelines for processing time-series data like video, audio, etc. This cross-platform Framework works on Desktop/Server, Android, iOS, and embedded devices like Raspberry Pi and Jetson Nano [23].

2.2 Related Work

Research work on sign language recognition is reported for many languages such as Arabic [1], Indonesian [2], American [3], Indian [4] and Thai [5]. Sign language recognition systems are based on one of two ways to detect sign languages 'gestures. They are sensor-based recognition systems and image-based recognition systems [25]. In sensor-based systems, sign language recognition is based on sensors that detect the hand's appearance. For this kind of system, two types are considered, which are the glove-based systems [27] and the Kinect-based systems [29]. Glove based systems [27] use electromechanical devices to recognize hand gestures. Hearing and speech impaired signers are required to wear a glove that is linked to some sensors that gather information [26]. Although this technique can offer good results, it can be inconvenient to the signers [26]. For the second category, Kinect sensors are used to detect sign language gestures. Originally, these sensor devices were developed by Microsoft for their Xbox game as an input device to interact with video games without using any remote controllers. Nowadays, the use of this device is expanding to include recognition systems like sign language recognition. On the other hand, image-based systems use images or videos along with image processing and machine learning techniques to recognize sign language gestures [26]. These systems fall into two categories. The first depends on using gloves containing visual markers to detect hand gestures, such as colored gloves [25]. However, this method prevents sign language recognition systems from being natural, where naturalness is expected from similar HCI systems [25]. The second category depends on images capturing hand gestures of the sign language [26]. When using these image-based recognition systems, hearing and speech impaired do not need to use any sensors or gloves with visual markers, which eliminates any inconvenience and makes the system more convenient [25].

In the past 20 years, deep learning have been used in sign language recognition by researchers from all around the world. Convolutional Neural Networks (CNNs) have been used for video recognition and achieved high accuracies last years.

- B. Garcia and S. Viesca at Stanford University proposed a real-time ASL recognition with CNNs for classifying ASL letters, described in Reference [28] . After collecting the data with a native camera on a laptop and pre-training the model on GoogLeNet architecture, they attained a validation accuracy of nearly 98% with five letters and 74% with ten.

- CNNs have also been used with Microsoft Kinect to capture depth features. Reference [29] proposed a predictive model for recognizing 20 Italian gestures. After recording a diverse video dataset, performed by 27 users with variations in surroundings, clothing, lighting and gesture movement, the model was able to generalize on different environments not occurring during training with a cross-validation accuracy of 91.7%. The Kinect allows capture of depth features, which aids significantly in classifying ASL signs.

- A.E.E.El Alfi, M.M.R.El Basuony, S.M.El Atawy [24] introduce an Arabic speech to ArSL intelligent conversion system. The system is based on a knowledge base to resolve some of the Arabic language problems (e.g. derivational, diacritical and plural). According to the authors' evaluation, the system has accuracy up to 96%. Conversely, the system has two main problems. It is a desktop application and the system output is a sequence of still images without any motion. Therefore, it is more suitable for educational purposes, not for real-time translation.

- Keskin et al. [7] developed a Random Decision Forests (RDFs) model trained on a 3D hand models that represented the hand with 21 different parts. RDF classified each pixel and assign it to a hand part. Then they estimate the joint locations for the hand part using a local mode finding algorithm They also developed a support vector machine (SVM) model to recognize the Arabic sign language digits using the same technique. They achieved a high recognition rate on live depth images in real-time.

- Mehdi et al. [8] developed a sensors-based method. He used 7 sensor gloves of the 5DT company to get the input data of the hands' movements. Then an artificial neural network (ANN) used to recognize the signs' gestures. This approach achieved an accuracy rate of 88%. Lopez- ´Noriega et al. [11] followed their same approach and also offered a graphical user interface made with ".NET".

- Starner et al.[12] used Hidden Markov Model (HMM) based that worked effectively in continuous and real-time sign language recognition tasks. They used color gloves to capture hand shape, orientation, and trajectory and then fed it to HMM-based system to recognize the sentence-level ASL. This system achieved high word accuracy results.

- Nandy et al. [10] proposed a new India sign language dataset and classify based on the feature of a direction histogram that appealed for illumination ad orientation invariance using two different approaches for recognition, the Euclidean distance and K-nearest neighbor metrics.

Most of these systems achieved high accuracy rate, but they are simulation-based and does not work for real-time applications. However, with the advent of mobile Apps, true deployment of sign language translation systems is profiling. For instance, a Vision-based Android App for Indonesian sign language using the OpenCV computer vision library is proposed in [2]. The App recognizes alphabets of the Indonesian sign language in real-time. Likewise, a mobile App for recognizing alphabets of the American sign language is proposed in [6]. The App uses a vision-based approach for the sign language recognition.

2.3 Project Description

In this document we propose a video call application available for all users, beneficial for communicating between an impaired and a vocal person. As in this call, the sign language of the impaired person will be translated into Arabic text. This application will use our proposed deep learning model that is a combination of CNN and LSTM. This model does not depend on any sensors nor gloves a it is only depend on recognizing the face and hand gestures, this allows our model to be more convenient and allows mobility and usability of the system. Our model is not user dependent as it will not be affected when different users use the system, the model will normally recognize the user. We used mediapipe solution to get a high-fidelity body pose tracking inferring 33 3D landmarks and background segmentation mask on the whole body from RGB video frames [30].

Many available datasets in Arabic sign language that focus on letters or words are based on specific conditions such as: (i) the user must wear gloves or (ii) many images refer to static words. So that, the major goal, which is the independence of unnecessary features related to specific users or the surrounding environment, cannot be achieved. That's why we created our dataset with a main objective of being a user dependent. As almost everyone has his own smartphone with camera, we collected our dataset using our smartphones without using any stabilization tool. The dataset videos are captured in different environments with different background and different resolution. This dataset will contain 30 words with 100 videos for each word.

We will start by extracting video frames then, extracting key points by using mediapipe. Mediapipe is a cross-platform library developed by google that offers fast and accurate results for a lot of things including the sign language recognition, since it recognizes the gestures made and integrates separate models for the pose, face, and hand component [23].

After that, we will pass those key points for the model the is a combination of CNN and LSTM to compute the final classification.

• CNN does not require human supervision to identify the important features. It is made up of layers of neurons that have learnable weights and biases. Adjusting these weights properly, the layers are able to extract features from the input image. They are very accurate at image recognition, also minimizing computation in comparison with a normal neural network, and also make use of the same knowledge across all image locations. Convolutional neural networks are great for a 1 to 1 relation; given an image of a sign, it generates fixed-size label, like the class of the sign in the image. However, What CNNs cannot do is accept a sequence of vectors. That’s where Recurrent Neural Networks (RNNs) are used

•LSTM is a special type of RNN. it works on the present input while taking into consideration the previous output and storing it in its memory for a short period of time, so it allows us to understand the context of a video frame, relative to the frames that came before it.

So the pre-trained CNN for feature extraction from input data along with LSTMs for sequence prediction. This model will be used in our final goal which is a video call application that is capable of translating sign language into Arabic language to help in the communication between an impaired and a vocal person

2.4 Functional Requirements

- **User can:**
  - If a user wants to register, he can do so by creating a new account or by signing up using a Google or Facebook account.
  - User can sign in if he already has an account.
  - User can add new contact.
  - User can select whom to call from his contacts.
  - User can delete any contact from his contact list.
  - User can search for any contact from his contact list.
  - User can start a video call with any of his contact list.
  - User can answer any incoming video call.
  - User can edit his profile picture and username.
  - User can reset his password

- **System should:**
  - During the video call, System can translate the Arabic sign language into text.
  - System can also simulate the speech into Arabic sign language

2.5 Non-Functional Requirements

- Performance: the system should return the result as fast as possible.
- Scalability: the system performance should not change with higher workloads
- Reliability: the extent to which the system run without a failure for a given period of time under predefined conditions
- Availability: accessibility of the system to a user at a given point in time.
- Security: assuring all data inside the system will be protected against malware attacks or unauthorized access.
- Usability: the system should be easy to use and easy to learn Privacy: protecting personal information and undesired access to personal space.

2.6 System Users

A. Intended Users:

- Any person who wants to communicate with a deaf person through video call.

B. User characteristics:

- Basic Knowledge of how to use the phone and make a video call with another person

Chapter 3

Analysis and Design

3.1 System Overview

In this section we will discuss a high-level description of the system through these 8 diagrams.

3.1.1 Use Case Diagram

The following figure represent the actions that the user can directly perform on the system namely:

- Registering if he has no account
- Logging in if he already has an account
- Editing his profile by editing one or more of the following: (username, password, profile pic).
- Starting a video call
- Answering an incoming video call
- Adding a new contact
- Searching any contact from his contacts
- Deleting any contact from his contacts
- Logging out his account.

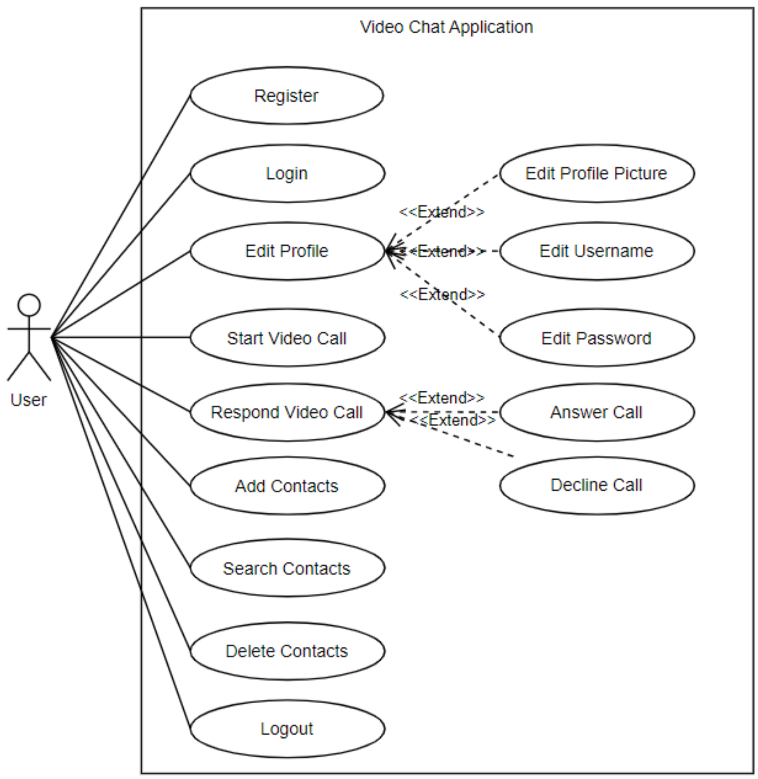

Figure 1: Use case diagram

3.1.2 Activity Diagram

The following figure represent the flow of activities that the user can do:

- It shows the flow of Login and Register activities, user can choose whether to login if he has account or register to create new account to access the home page. User will not be able to access the home page without verifying.
- After logging in user can edit profile, add contact, delete contact, search contact, start video call, answer video call and logging out his account.

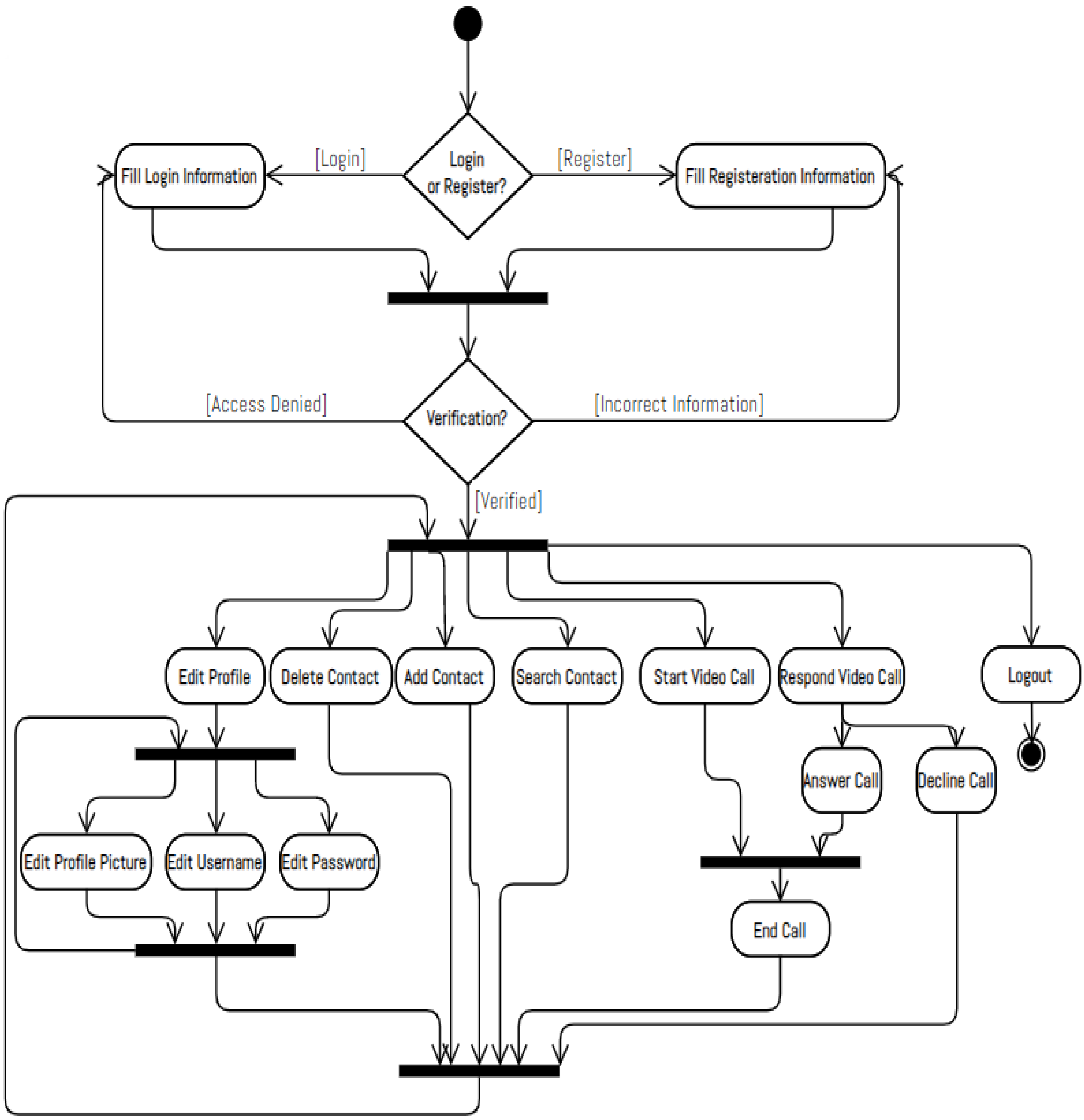

Figure 2: Activity diagram

3.1.3 Sequence Diagram

This diagram describes the sequence of actions that a user can make starting by logging in if he already has an account. After providing his username and password, the system will check with the database if they are valid or not. On the other hand, if he is a new user, he can also register and definitely the system will add his account to the database if it is valid. Once the user logged in successfully, he can edit his profile picture, password, or username and the system will update them in the database. He also can add any contact and if this contact has an account on the system the database will add this contact to his contact list. Moreover, the user can search for any contact and if he is already in his list, the database will return his profile and the user can start a video call or delete him. Furthermore, the user can answer any incoming call or end a call.

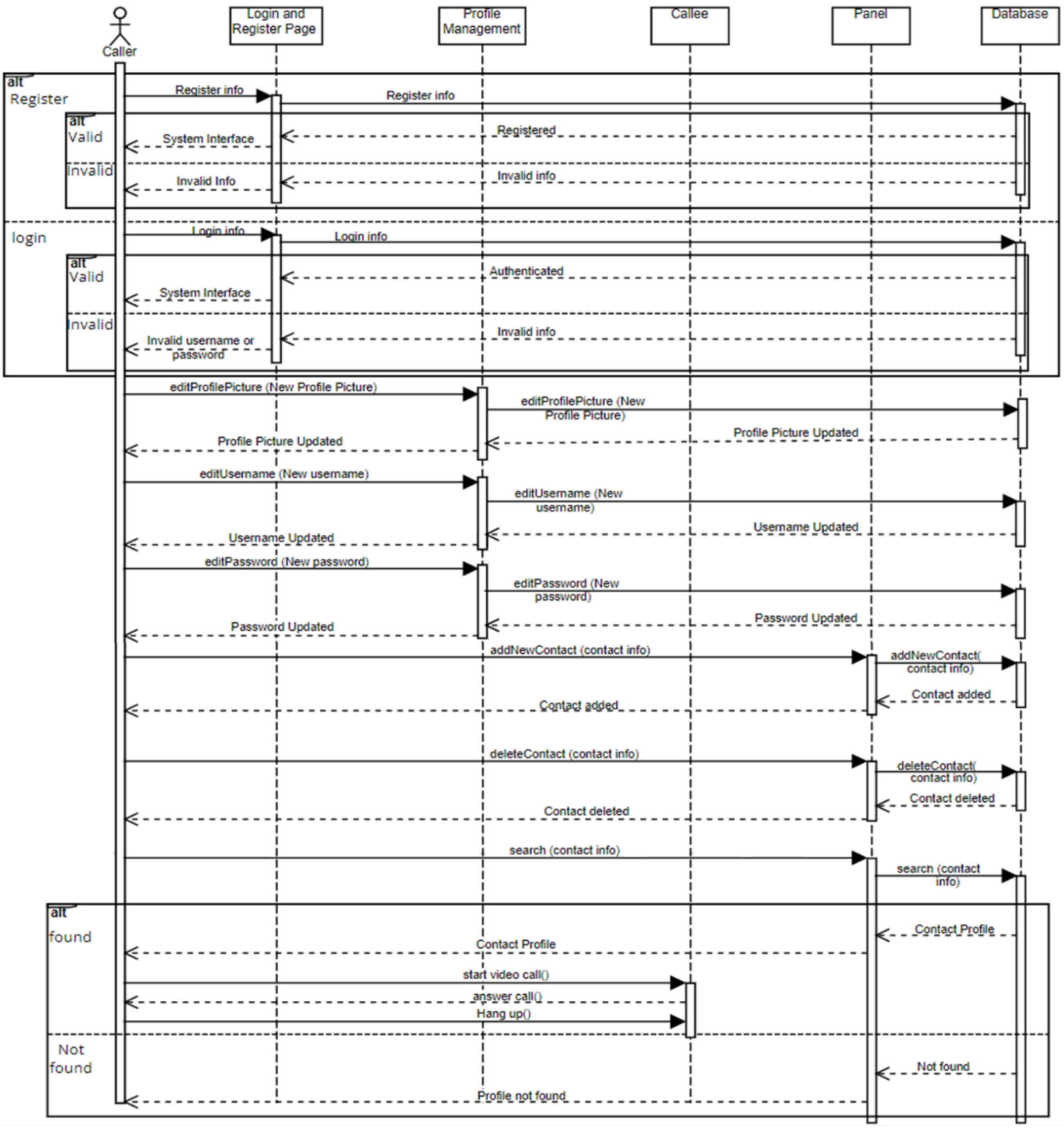

Figure 3: Sequence diagram

3.1.4 State Diagram

This Diagram shows the response of all the classes in the system to the user actions.

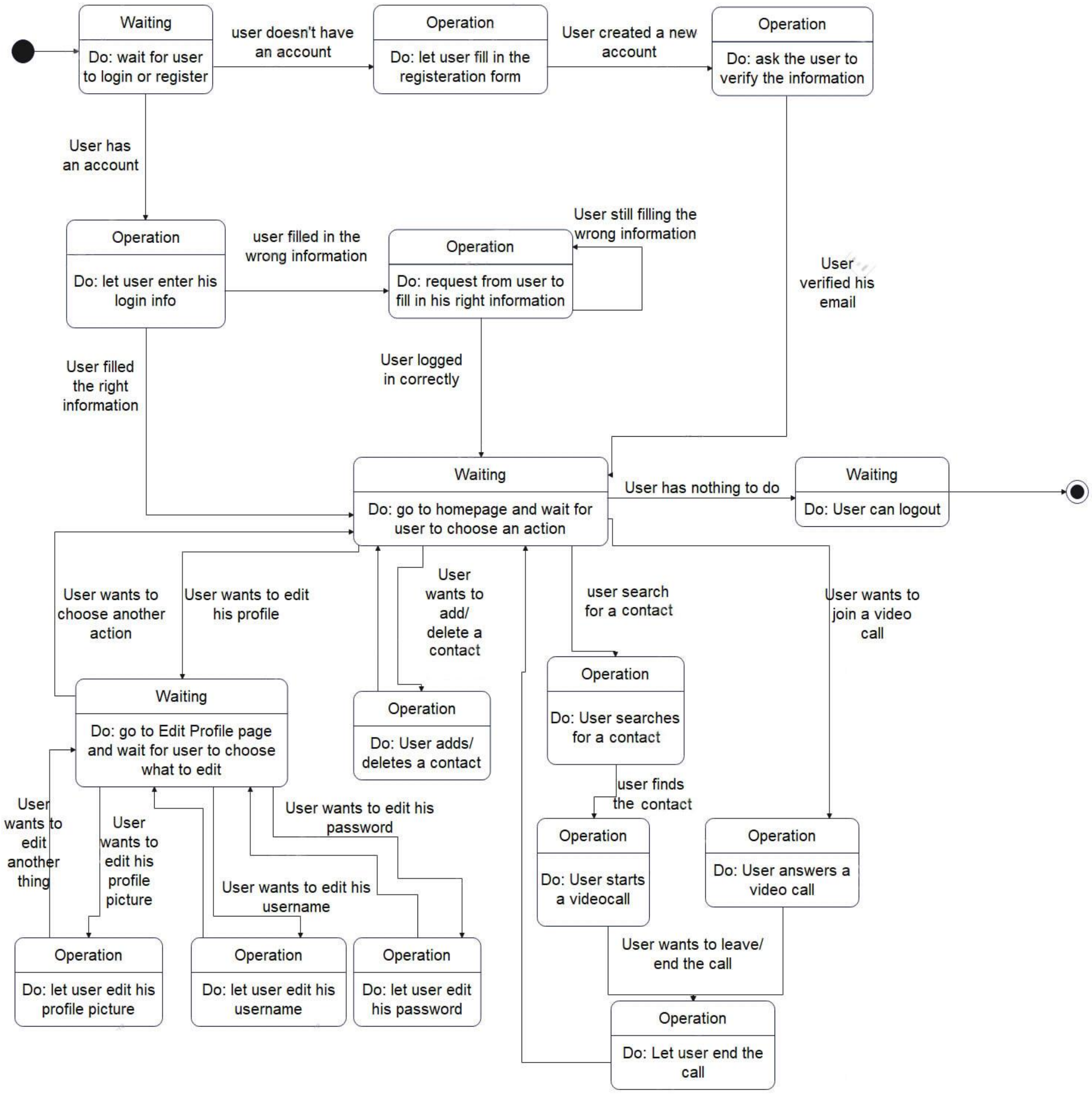

Figure 4: State Diagram

3.1.5 Class Diagram

The following figures represents the functions and attributes of the four main classes that the system use:

- User
- Profile
- Contact
- Panel

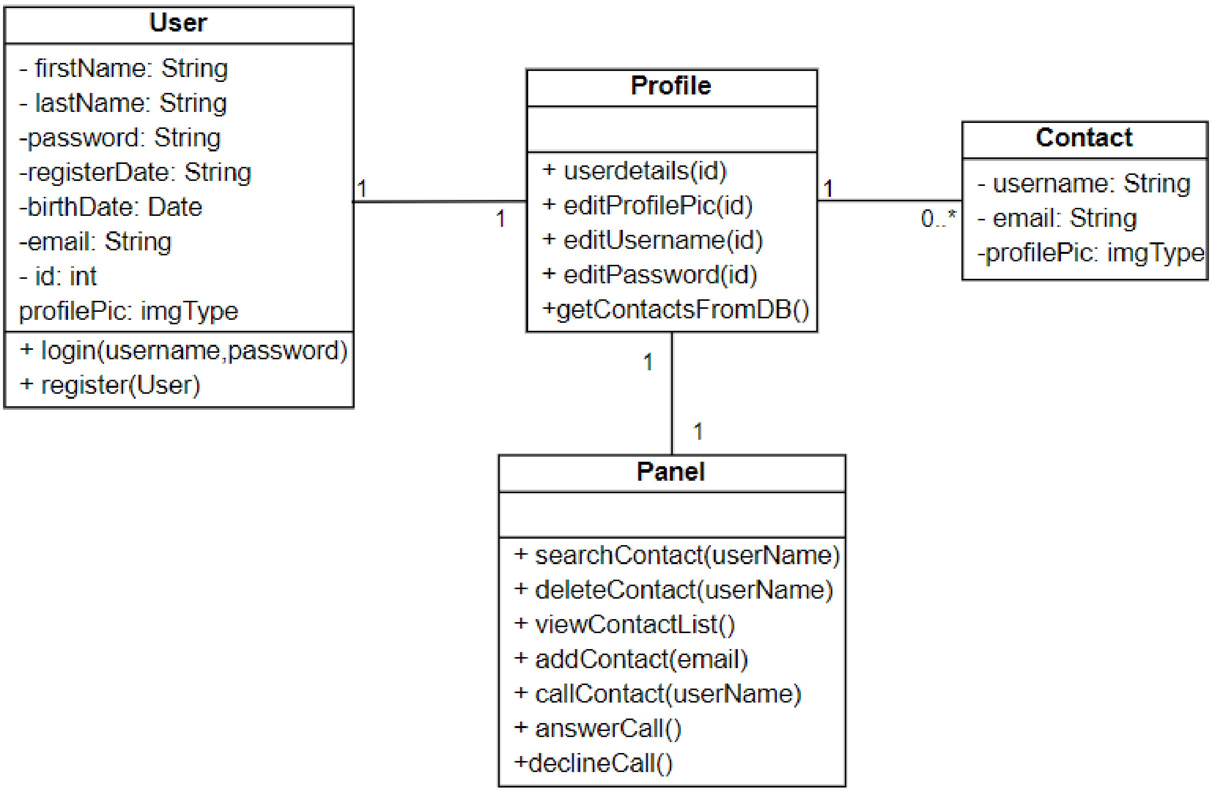

Figure 5: Class diagram

3.1.6 Flow Chart

Each one of the following flow charts represents a function in our system in details.

3.1.6.1 Register and Login

The following figure shows the steps of register and login function:

- The user can either register to create new account or login by his account.
- User can login by entering username and password.
- The system will let the user access the home page if the information is valid.

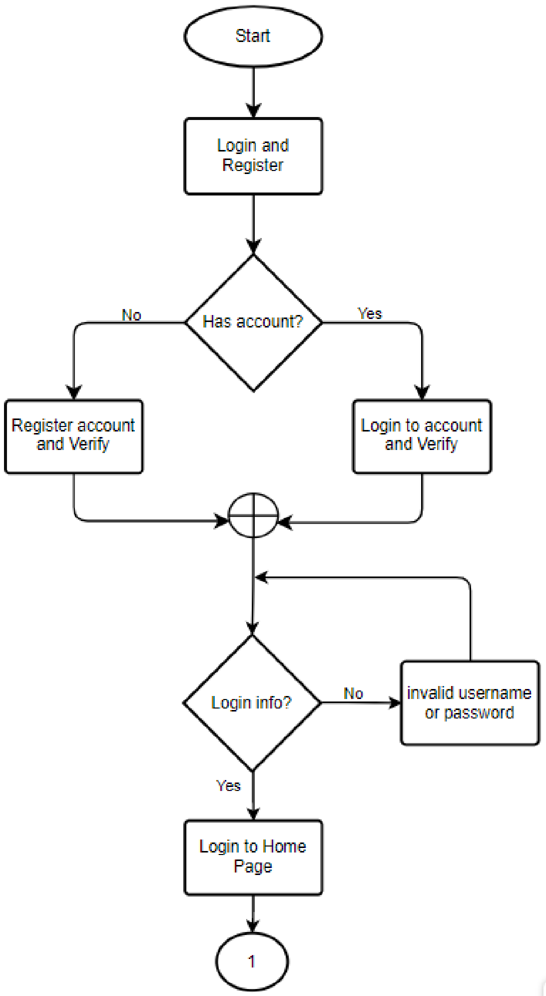

Figure 6: Register and Login flow chart

3.1.6.2 Add or Delete Contact

The following figure shows the steps of add and delete contact function:

When user wants to add a new contact or delete an existing contact, he can search for the contact by entering his email or username if the contact found user can add it to the contact list or delete it from his contact list.

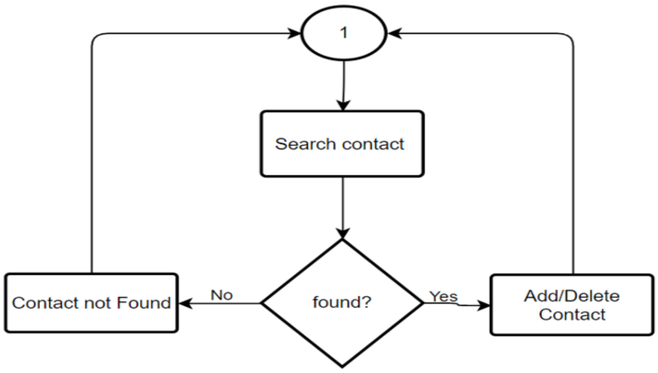

Figure 7: Add or Delete Contact flow chart

3.1.6.3 Start Video Call

The following figure shows the steps of starting a video call:

- When the user wants to start a video call, he can select any contact from his list then starting the video call.
- If the call ended user will return to the home page.

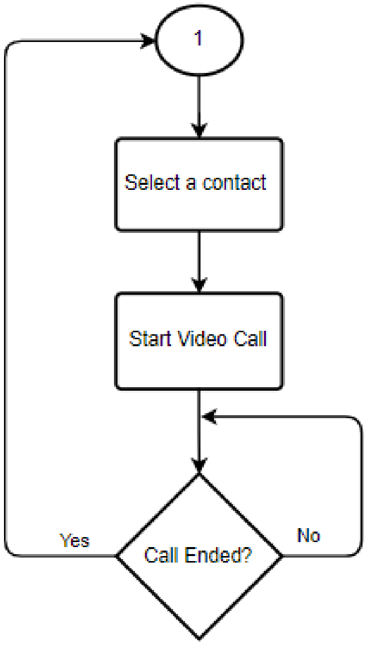

Figure 8: Start Video Call flow chart

3.1.6.4 Answer Video Call

This figure shows the workflow of answering video call function:

When the user has an incoming video call, he has the option to answer it and during the call he can end it, or he can hang it up in either case the user will return to the homepage again.

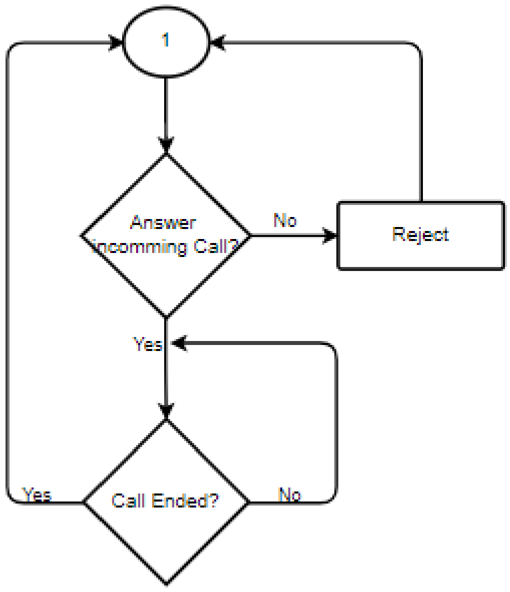

Figure 9: Answer Video Call flow chart

3.1.6.5 Edit Profile

This figure shows the workflow of edit profile function:

- When the user chooses to edit his profile, he can edit his password, pofile picture , or password.
- If his edits are valid the system will return the user to the homepage and if there was an error or the edits are invalid, the system will return the user to the edit profile screen again.

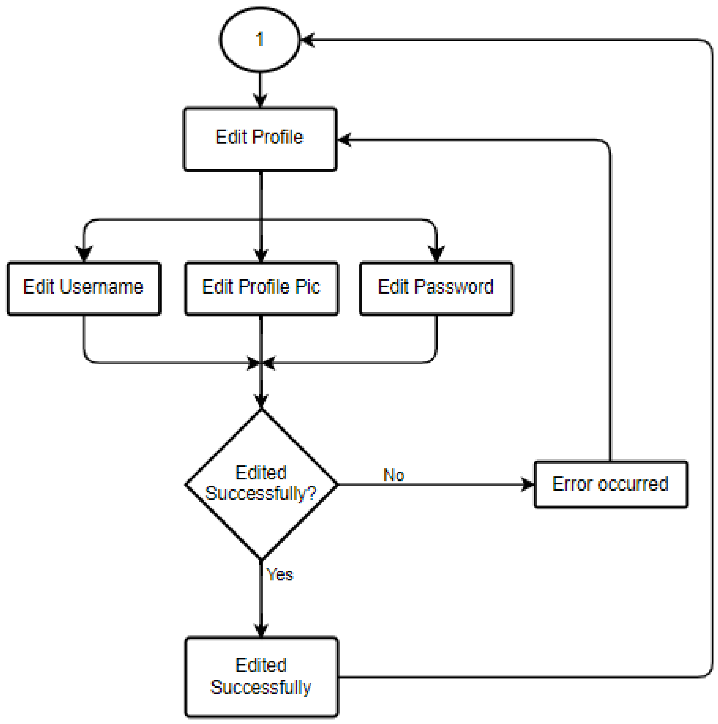

Figure 10: Edit Profile flow chart

3.1.6.6 Search Contact

This figure shows the workflow of search contact function:

- When user wants to search for a contact, he can type his username or email in the search bar.
- If the database found it in his contact list, the system will return his profile and contact information.
- If the contact was not found in the database the system will return to the homepage.

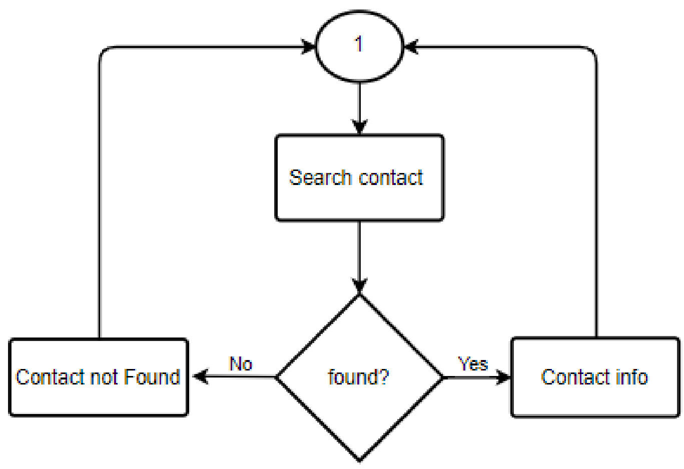

Figure 11: Search Contact flow chart

3.1.6.7 Logout

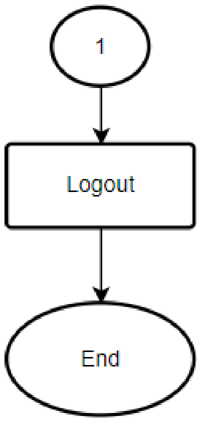

Figure 12: Logout flow chart

3.1.6.8 Video call

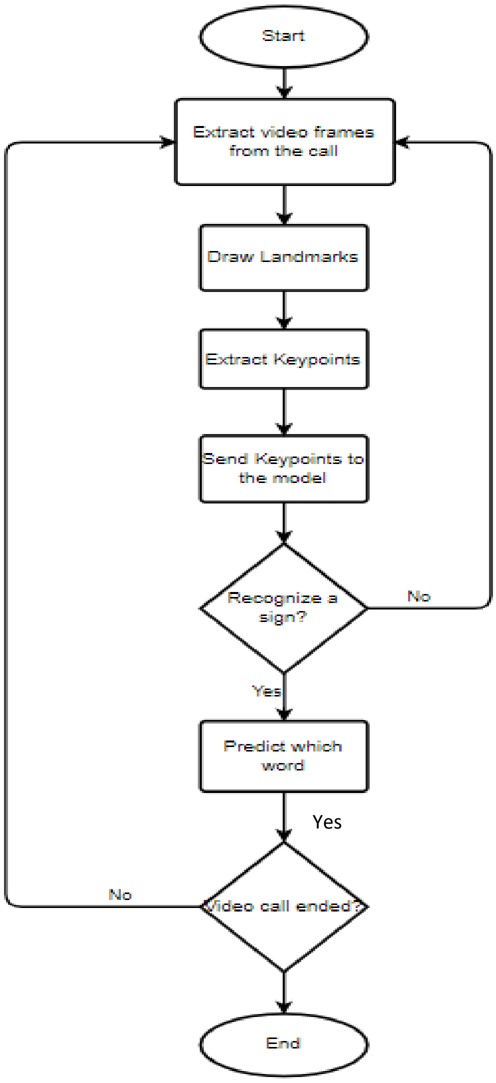

Figure 12: Video call flow chart

3.1.7 Block Diagram

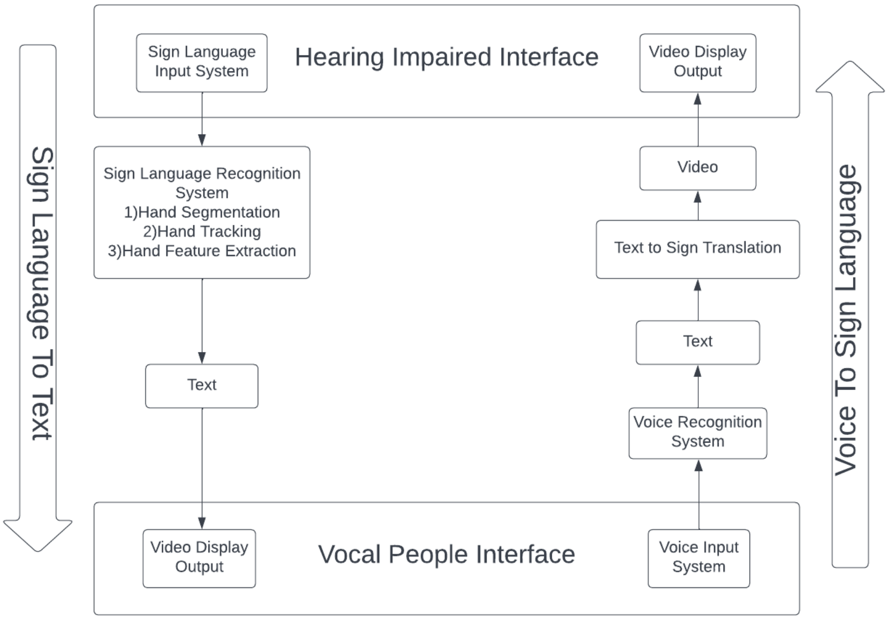

Figure 14: Block Diagram

3.1.8 System Architecture

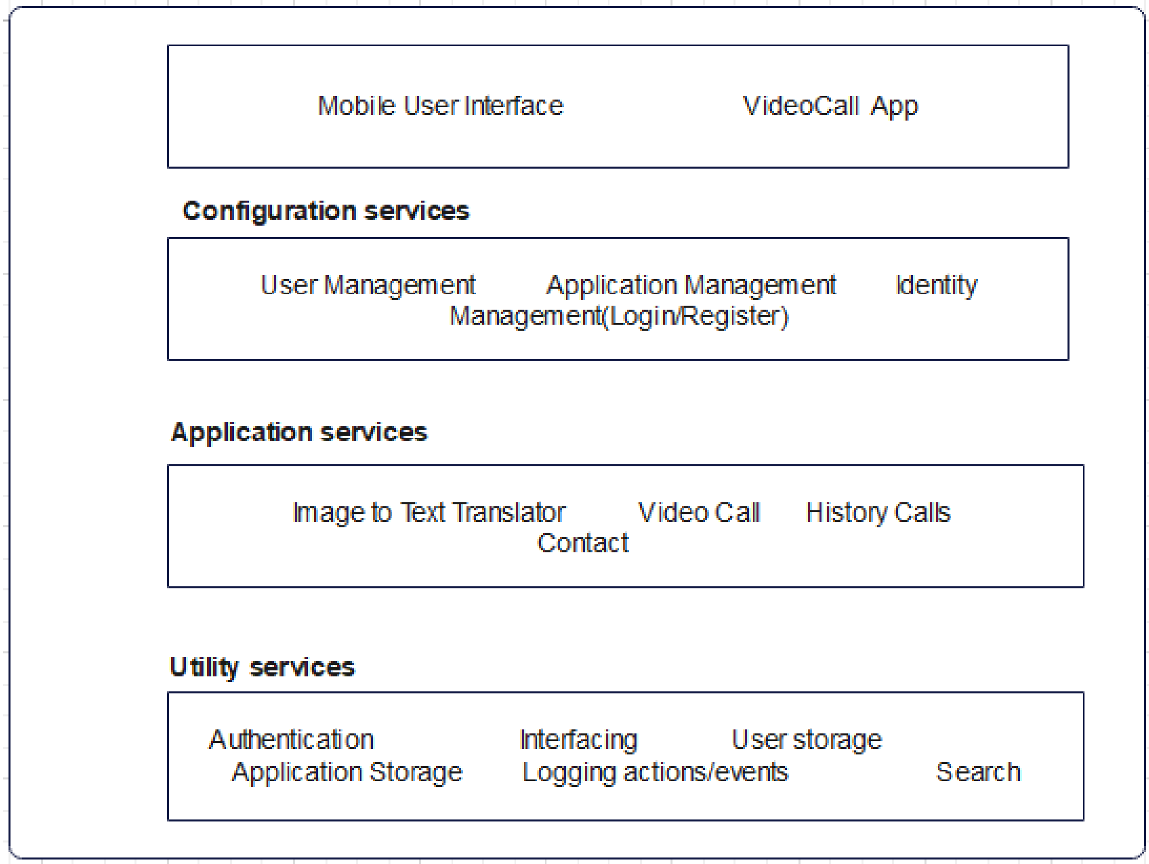

Figure 15: System Architecture

Application Design Interface

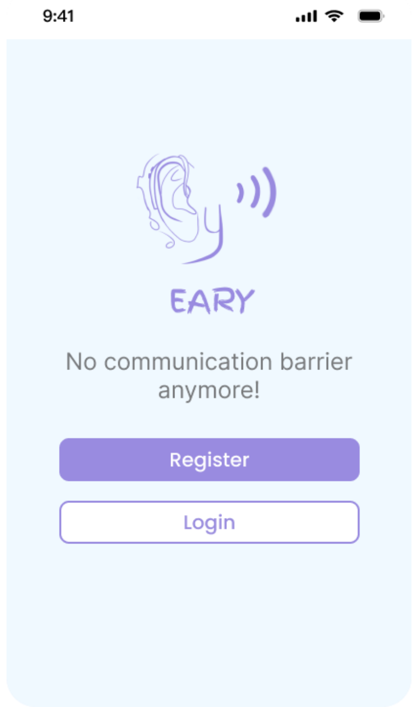

Figure 16: Login and register page

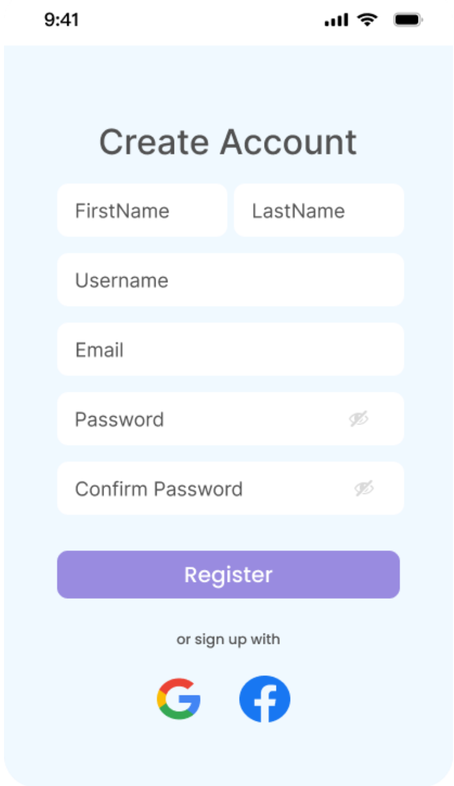

Figure 17: Register page

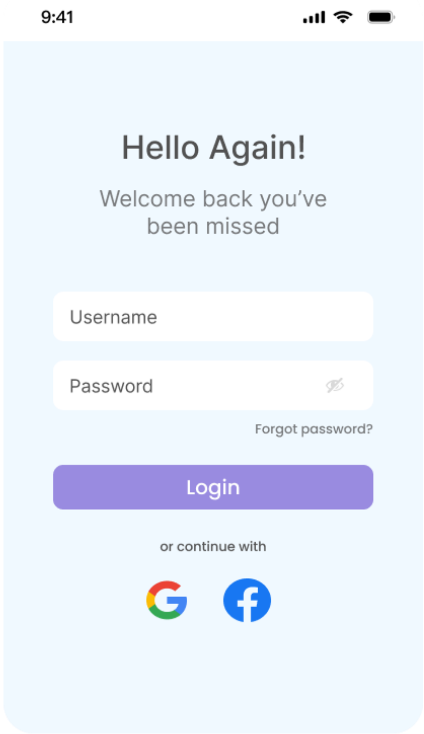

Figure 18: Login page

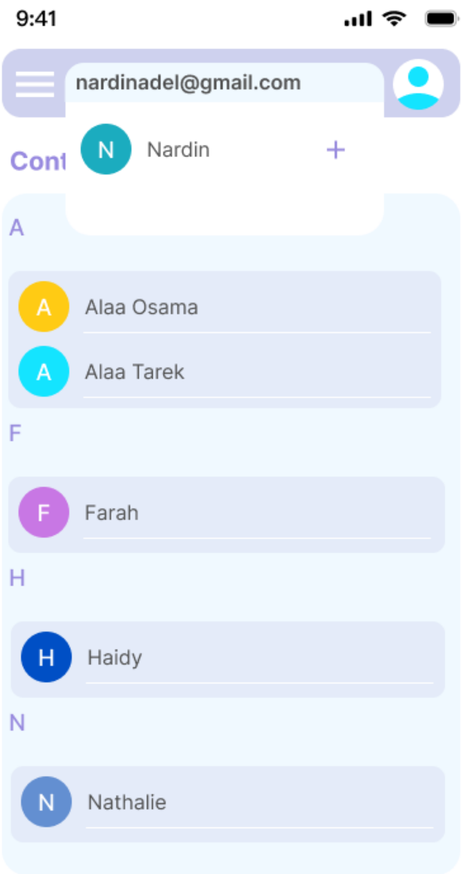

Figure 19: Search and add contact page

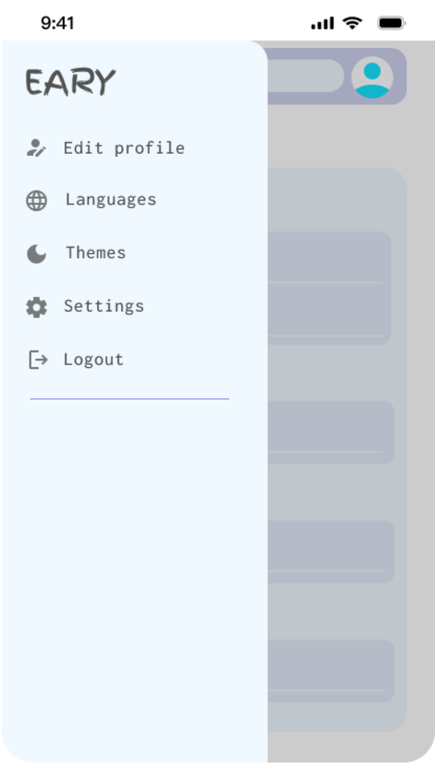

Figure 20: Navigation drawer

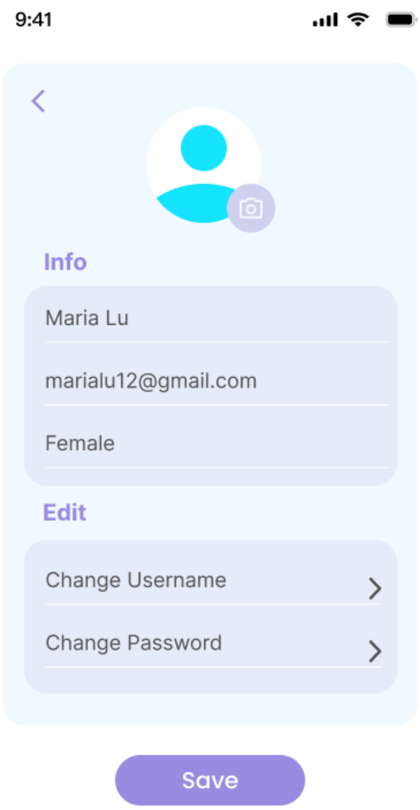

Figure 21: Edit profile page

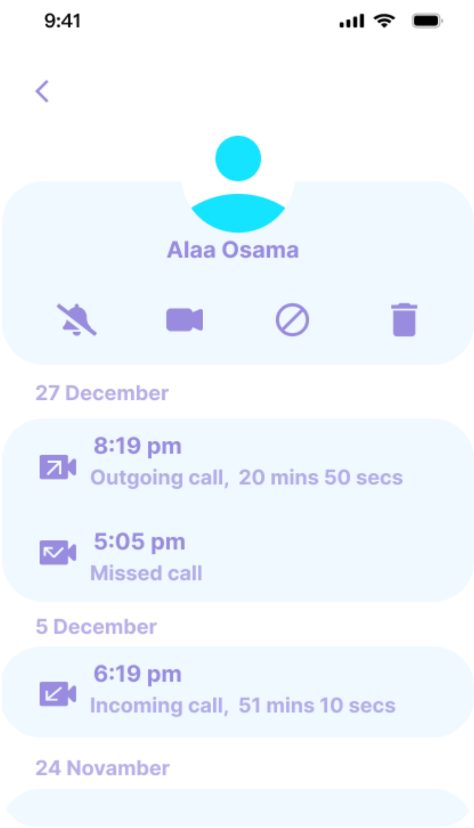

Figure 22: Contact info page

Figure 23: Video call page

References

[1] T. Shanableh, K. Assaleh and M. AL–Rousan, "Spatio–Temporal feature extraction techniques for isolated Arabic sign language recognition," IEEE Transactions on Systems, Man and Cybernetics Part B, 37(3), June, 2007.

[2] I. Kryvonos, I. Krak and W. Wojcik, "Information technologies applications for sign languages investigations," Computer Science and Information Technologies (CSIT), 2015, Yerevan, 2015, pp. 148-150.

[3] J. Wu, Z. Tian, L. Sun, L. Estevez and R. Jafari, "Real-time American Sign Language Recognition using wrist-worn motion and surface EMG sensors," 2015 IEEE 12th International Conference on Wearable and Implantable Body Sensor Networks (BSN), Cambridge, MA, 2015, pp. 1- 6.

[4] N. Yadav, S. Thepade and P. H. Patil, "Noval approach of classification based Indian sign language recognition using transform features," 2015 International Conference on Information Processing (ICIP), Pune, 2015, pp. 64-69

[5] C. Chansri and J. Srinonchat, "Reliability and accuracy of Thai sign language recognition with Kinect sensor," 2016 13th International Conference on Electrical Engineering/Electronics, Computer, Telecommunications and Information Technology (ECTI-CON), Chiang Mai, 2016, pp. 1-4.

[6] C. M. Jin, Z. Omar and M. H. Jaward, "A mobile application of American sign language translation via image processing algorithms," 2016 IEEE Region 10 Symposium (TENSYMP), Bali, 2016, pp. 104-109

[7] Keskin C, Kırac¸ F, Kara YE, Akarun L (2013) Real time hand pose estimation using depth sensors in consumer depth cameras for computer vision. Springer, pp 119–137

[8] Mehdi SA, Khan YN (2002) Sign language recognition using sensor gloves. In: Proceedings of the 9th international conference on neural information processing, 2002. ICONIP'02, vol 5. IEEE, pp 2204– 2206

[9] Yang S, Zhu Q (2017) Continuous chinese sign language recognition with cnn-lstm. In: Ninth international conference on digital image processing (ICDIP 2017). (International Society for Optics and Photonics), vol 10420, p 104200F

[10] Nandy A, Prasad JS, Mondal S, Chakraborty P, Nandi GC (2010) Recognition of isolated indian sign language gesture in real time. In: International conference on business administration and information processing. Springer, pp 102–107

[11] Lopez-Noriega JE, Fern ´ andez-Valladares MI, Uc-Cetina V (2014) Glove-based sign language recognition solution to assist communication for deaf users. In: 2014 11th International conference on electrical engineering, computing science and automatic control (CCE). IEEE, pp 1–6

[12] Starner TE (2008) Visual recognition of american sign language using hidden Markov models. Massachusetts Inst Of Tech Cambridge Dept Of Brain And Cognitive Sciences. Technical report

[13] Reading the signs: Diverse arabic sign languages. Reading the Signs: Diverse Arabic Sign Languages | Center for Strategic and International Studies. (n.d.). Retrieved December 26, 2022.

[14] UCSF Health. (2022, June 24). Communicating with people with hearing loss. ucsfhealth.org. Retrieved December 26, 2022.

[15] World Health Organization. (n.d.). Deafness and hearing loss. World Health Organization. Retrieved December 26, 2022.

[16] Alselwi, G., & Tasci, T. (n.d.). (PDF) arabic sign language recognition: A review - researchgate. Retrieved December 26, 2022.

[17] Alsaadi, Z.; Alshamani, E.; Alrehaili, M.; Alrashdi, A.A.D.; Albelwi, S.; Elfaki, A.O. A Real Time Arabic Sign Language Alphabets (ArSLA) Recognition Model Using Deep Learning Architecture. Computers 2022, 11, 78.

[18] Balaha, M.M., El-Kady, S., Balaha, H.M. et al. A vision-based deep learning approach for independent-users Arabic sign language interpretation. Multimed Tools Appl (2022).

[19] M.El-Gayyar, M., S.Ibrahim, A., & M.E.Wahed. (2016, May 18). Translation from Arabic speech to arabic sign language based on cloud computing. Egyptian Informatics Journal. Retrieved December 26, 2022.

[20] What is deep learning? IBM. (n.d.). Retrieved December 26, 2022.

[21] Saha, S. (2022, November 16). A comprehensive guide to Convolutional Neural Networks-the eli5 way. Medium. Retrieved December 26, 2022.

[22] says, K. L. (2022, December 13). What is LSTM - introduction to long short term memory. Intellipaat Blog. Retrieved December 26, 2022.

[23] Kukil. (2022, November 18). Introduction to MediaPipe. LearnOpenCV. Retrieved December 26, 2022.

[24] A.E.E.El Alfi, M.M.R.El Basuony, S.M.El Atawy Intelligent Arabic text to arabic sign language translation for easy deaf communication Int J Comput Appl, 92 (2014), pp. 22-29

[25] O. Al-Jarrah and A. Halawani, “Recognition of gestures in arabic sign language using neuro-fuzzy systems,” Artif. Intell., vol. 133, no. 1–2, pp. 117–138, Dec. 2001

[26] A. Tharwat, T. Gaber, A. E. Hassenian, M. K. Shahin, and B. Refaat, “SIFT-Based Arabic Sign Language Recognition System,” vol. 334, Nov. 2014.

[27] N. Tubaiz, T. Shanableh, and K. Assaleh, “Glove-based continuous arabic sign language recognition in user-dependent mode,” IEEE Trans. Hum.-Mach. Syst., vol. 45, no. 4, pp. 526–533, Aug. 2015.

[28] B.Garcia, S. Viesca, “Real-time American Sign Language Recognition with Convolutional Neural Networks” 2016.

[29] L. Pigou, S. Dieleman, P. Kindermans, B. Schrauwen. “Sign Language Recognition using Convolutional Neural Networks” 2015.

[30] Pose. mediapipe. (n.d.). Retrieved December 27, 2022.